



MRC Integrative
Epidemiology
Unit



University of
BRISTOL

School of Psychological Science
12a Priory Road
Bristol BS8 1TU
United Kingdom

kd16662@bristol.ac.uk

www.bristol.ac.uk

STUDY PROTOCOL
**WOULD CONSUMERS BE CONFUSED BY THE TERM ‘MILK’ ON MILK
SUBSTITUTE LABELS?
AN ONLINE EXPERIMENTAL STUDY.**

Katie De-loyde, Marcus Munafò, Angela Attwood, Oliva Maynard

Background

Milk substitutes (e.g., plant-based milks) are products that replicate the taste, texture and function of animal-based dairy products, but contain no animal-based ingredients, and are instead made using plant-based ingredients. Due to concerns over health, animal welfare, and the environment, it is now estimated that almost a quarter (23%) of the British public drink milk substitutes (Mintel, 2019). Moreover in 2019, 37 researchers from 16 countries collaboratively published the Eat-Lancet report, which advocates for global adoption of a predominantly plant-based diet and significant reductions in consumption of animal products such as dairy (Willett *et al.*, 2019).

As the popularity of milk substitutes increases, so does the disagreement about how these products should be labelled and described. In Europe (Court of Justice of the European Union, 2017) and the UK (Department for Environment: Food & Rural Affairs, 2013), existing regulation states that terms like ‘milk’, ‘cheese’ and ‘yogurt’ can only be used to describe products that originate from animals. This means that combining words like ‘oat’ and ‘soya’ with ‘milk’ to describe a product on its labelling is not allowed. In contrast, in some other countries (e.g., Australia) terms such as ‘soya milk’ are legally permitted on labelling, and in the USA the Food and Drug Administration (FDA) also allow milk substitute products to use the term ‘milk’; however, the dairy industry has petitioned the FDA and USA Congress for several decades to prohibit milk substitutes from using the term ‘milk’ on labels (Food and Drug Administration, 2018; Welch, 2019).

Supporters of regulations that prohibit the term ‘milk’ on milk substitute labelling argue that people have traditionally associated terms like ‘milk’ with animal products, and therefore, when these terms appear on plant-based products they could mislead consumers about the products’ use, ingredients, and/or nutritional value (Food and Drug Administration, 2018; Gleckel, 2020). In contrast, opponents of such regulations insist that using terms like ‘milk’ on product labels does not confuse consumers. They maintain that, on the contrary, these terms are necessary to prevent confusion by more precisely communicating the uses of products. Plant-based food companies argue that the animal agriculture sector wants to hide

the fact that these milk substitutes can be used for milk substitution (Gleckel, 2020), while the dairy industry argues that regulation of the term ‘milk’ on milk substitute packing is needed to protect consumers (Welch, 2019).

Research in this area is limited. However, a survey of 1,000 adults by the international Food Information Council in October 2018 found that the majority (~75%) of Americans understand which products contain dairy milk and which do not when shopping for products that carry the label ‘milk’ (International Food Information Council Foundation, 2018). In fact, less than 10% of respondents believed that coconut, soy, cashew, and rice ‘milk’ contained animal-based ingredients (International Food Information Council Foundation, 2018). A small experimental study in 2020 (Gleckel, 2020) also showed that consumers were no more likely to think that plant-based products come from an animal if the product’s name incorporates words traditionally associated with animal products (such as butter) than if it does not. However, although this study assessed both meat and dairy confusion regarding plant-based food labelling, it used butter substitutes and not milk substitutes.

In a previous observational study (manuscript currently in preparation), we found that labelling of milk substitutes in UK supermarkets is varied and inconsistent. In fact, across supermarkets, 14-30% of milk substitutes were labelled only with their brand or main ingredient. Furthermore, although ‘milk’ was often used by the supermarket in the online description of milk substitutes, in a real-world supermarket setting any clarity that comes from this description would rarely be available to consumers.

This current study aims to address this gap in the literature. The primary aim is to assess whether adding the term ‘milk’ to labels of milk substitutes would: 1) more accurately communicate the use of milk substitutes to consumers, and 2) confuse consumers about which products come from an animal source. A secondary aim is to assess whether adding the term ‘milk’ to milk substitute labelling would, as per the Eat-Lancet report (Willett *et al.*, 2019), encourage intended use of these products. Assessing knowledge of nutritional or ingredient variations between dairy milk and milk substitutes is not within the scope of this study.

Definitions

Dairy milk: (fluid) milk from an animal source

Milk substitute: (fluid) milk not from an animal source (from a plant-based source).

Study Hypotheses

Primary hypothesis:

Participants in the Milk Labelling Condition will correctly identify more milk substitutes as being a product that could be added to a cup of tea or coffee compared to participants in the No Milk Labelling Condition. This reflects the view of supporters of using the term ‘milk’ on milk substitute labelling.

Secondary hypothesis:

Participants in the Milk Labelling Condition will misidentify more milk substitutes as animal source products compared to participants in the No Milk Labelling Condition. This reflects the view of opposers of using the term ‘milk’ on milk substitute labelling.

Tertiary hypotheses:

Participants in the Milk Labelling Condition will be faster at identifying milk substitutes as products that could be added to a cup of tea or coffee compared to the No Milk Labelling Condition. This reflects the view of supporters of using the term ‘milk’ on milk substitute labelling.

Participants in the Milk Labelling Condition will be slower at identifying milk substitutes as non-animal source products compared to the No Milk Labelling Condition. This reflects the view of opposers of using the term ‘milk’ on milk substitute labelling.

Participants in the Milk Labelling Condition will be more likely to them themselves choose a milk substitute to put in their cup of tea / coffee or a bowl of cereal in the future compared to participants in the No Milk Labelling Condition. This reflects the view of supporters of using the term ‘milk’ on milk substitute labelling.

Study Design

This study will be an experimental study conducted online.

Participants will be equally randomised to one of two conditions in a between-subjects design:

No Milk Labelling Condition: Where milk substitutes are not labelled with the term milk and instead have examples of current UK labelling found in supermarkets from a previous observational study.

Milk Labelling Condition. Where milk substitutes are labelled with the term milk.

See Appendix 1 for further details on carton design and labelling.

Study Site

The study will be administered from the School of Psychological Science at the University of Bristol, but data will be collected online. The study will be advertised, and participants recruited via Prolific (<https://www.prolific.co/>), an online testing platform. Data will be collected and stored on the Qualtrics online survey platform (<http://www.qualtrics.com/>).

Registration

The study has been registered on the Open Science Framework: <https://osf.io/c74ka/>

Participants and Recruitment

Participants will be recruited through the Prolific Academic online platform, which will advertise the study to existing members, who will be reimbursed after task completion through their Prolific accounts.

Participants who want to take part will read an information statement before giving their consent to participate. Participation is expected to take approximately 5 minutes to complete, and participants will be reimbursed £0.63 on completion in line with the current hourly rate suggested by Prolific Academic.

Inclusion criteria includes:

- ≥ 18 years old
- A current resident in the UK
- Fluent in English.

Withdrawal of participants

Participants will be informed that they are able to withdraw from the experiment by closing their internet browser. None of their data will be saved if they do this. However, participants who withdraw will not be reimbursed. Data collected will be anonymous; therefore, participants will not be able to withdraw their data at a later point. Participants will be made aware of this in the information statement.

Randomisation

The study will be delivered on the Qualtrics platform, which will be used to pseudo-randomise participants into one of the two conditions. Study participants will be recruited to achieve balanced numbers in each of the two conditions

Measures and Materials

All participants will see 20 drinks cartons during the study which will be made up of ten milk substitute cartons and ten 'other drinks' cartons (Table 1). Images of all cartons will be taken from currently available products sold in the top four UK supermarkets chains (based on customer base and sales). Milk substitute cartons in the Milk Labelling Condition will be modified to include the term 'milk' on the packaging. All other cartons will remain unmodified. See Appendix 1 for images.

Table 1. Drinks cartons between study conditions

Condition	Carton labelling		
	Milk substitute cartons (10 cartons)	Dairy milk cartons (5 cartons)	Other drink cartons (5 cartons)
<p>No Milk Labelling Condition:</p> <p>Where milk substitutes are not labelled with the term milk.</p>	<ul style="list-style-type: none"> -Oatly. Oat drink -Alpro Soya -Soya drink -Rice -Almond drink -Mylk -Oat drink -Soya unsweetened -Cashew unsweetened -Dream rice 	<ul style="list-style-type: none"> -Whole milk -Skimmed milk -Whole milk -Semi-skimmed milk -Skimmed milk 	<ul style="list-style-type: none"> -Orange juice -Apple juice -Banana drink -Chocolate drink -Strawberry drink
<p>Milk Labelling Condition:</p> <p>Where milk substitutes are labelled with the term milk</p>	<ul style="list-style-type: none"> -Oatly. Oat milk -Alpro Soya milk -Soya milk -Rice milk -Almond milk -Mylk milk -Oat milk -Soya milk unsweetened -Cashew milk unsweetened -Dream rice milk 		

Procedures

Participants will be recruited using the Prolific Academic online platform, which will provide only pre-screened participants (those fluent in English, those who are a resident of the UK and those who are ≥ 18 years old), using the Prolific demographic pre-screening tool, with a link to the study on the Qualtrics platform. The sample will be balanced between sex (male and female) using the Prolific balanced sample tool. The task will be optimised for viewing on small devices (e.g., mobile phones) as well as desktops and laptops.

Participants will first be shown an information statement explaining the study and what they will be required to do. As a cover story, participants will be told the study is about breakfast consumption in the UK. Participants will be informed that they are able to withdraw from the study by closing their Internet browser. Before commencing the study, participants will complete a tick-box consent page.

To provide credibility to the cover story, participants will then be asked two questions about their breakfast consumption: What do you have for breakfast? How many cups of coffee do you drink at breakfast time?

Participants will then be presented with a study question, before both conditions are presented with 20 drinks cartons, listed in a random order, on one continuous page (the participant will need to scroll down to see all products).

These 20 drinks cartons will be made up of ten milk substitute cartons and ten ‘other drinks’ cartons (Table 1). The milk substitute cartons will vary in the labelling depending on the condition assigned to each participant: the No Milk Labelling Condition will have the original unmodified design, which is currently available in UK supermarkets, while the Milk Labelling Condition will have a modified version which includes the term ‘milk’ on the packaging (Table 1 and Appendix 1).

The ten ‘other drinks’ cartons will be made up of five cartons representing dairy milk, and five cartons which will represent other drink products (e.g., orange juice) (Table 1). These ten cartons will be the same for both conditions and will be used for filler purposes only and will only be reported descriptively between study conditions. Having fewer of these filler cartons compared to milk substitute cartons will lower the participant burden and time taken to complete the task. See Appendix 1 for images of all cartons.

For each of the 20 drink cartons, participants will be asked to answer the study question by selecting either yes, no or unsure. The participant will only be able to move onto the next study question once all products have either a yes, no or unsure answer. After the participant has completed this, the same 20 drinks cartons for their condition will appear, again in a random order, and a further study question will be presented. There will be three study questions in total (see below). Questions 1 and 2 will be timed (in seconds), which participants will be aware of as indicated by a display before these questions:

“You are being timed for this question. Please answer accurately and as quickly as possible”.

Study questions:

1) Please select whether you think someone would add each of the products to a cup of tea or coffee. This includes those products which would be added to tea or coffee by someone using the product to replace dairy milk. (yes, would be added / no, would not be added / unsure).
-100% correct would be 10 x yes to milk substitute cartons

2) Please select whether you think each product listed below comes from an animal source or not (yes, comes from an animal source / no, does not come from an animal source / unsure).
-100% misidentified would be 10 x yes to milk substitute cartons

3) Please select whether you think, in the future, you may purchase each product below to add to your tea or coffee (or for another such use, for example, to be added to a smoothie or bowl of cereal) (yes, I would purchase / no, I would not purchase / unsure).
-Number of yes answers out of the 10 milk substitutes.

Answers will be recorded for each product for each question. Participants will then complete demographic questions and additional measures (see Additional Measures section below). Participants will then be asked an open question about what they thought the purpose of the study was. We will also include a free text box at the end where participants can provide any additional feedback (these data will not be formally analysed, but we find this is helpful for identifying any issues participants had and their feedback can be useful for interpreting the

results). After all questions have been answered, participants will be debriefed, which will include giving them information about the real purpose of the study. They will also be given the Principal Investigator's details if they wish to contact them. Participants will be redirected back to Prolific for reimbursement.

Attention checks

Given concerns regarding the attention participants may pay in unsupervised settings, two attention check items will be embedded within the questionnaire to check attention. Participants will be asked 'How often do you eat cheese in a week? Regardless of your answer, please write the word zero' (open ended) and 'Please select the number 5 from this list' (list of numbers 1 – 5). One question will be presented during the task, and one will be presented at the end of the study. Any answers that are \neq 'zero' to the first question, and \neq 5 to the second question will be deemed an attention check fail. These participants will still be reimbursed but will be excluded from the final analysis. These participants will therefore be replaced in the sample so that the calculated sample size will be met. Participants will be given the following message before answering the three questions:

"Please note: Data quality is of the utmost importance for this survey, please be aware that the task will include some questions intended to test whether you are paying full attention."

Study purpose check

Participants will be asked about the purpose of the study as an open-ended question. This will be coded after study completion and any participant who was deemed to guess the true purpose of the study will be excluded for a sensitivity analysis of the primary outcome.

Outcomes

The primary outcome will be:

The number of milk substitutes the participant correctly identifies in question 1 (out of 10).

100% correct would be 10 x yes to milk substitute cartons (i.e., a no or unsure answer would be considered an answer that does not correctly identify milk substitutes as being a product that could be added to a cup of tea or coffee).

The secondary outcome will be:

The number of milk substitutes the participant misidentifies in question 2 (out of 10).

100% misidentified would be 10 x yes or unsure to milk substitute cartons (i.e., a no answer would be considered the only answer that does correctly identify a milk substitute as not being from an animal source).

Tertiary outcomes will be:

3.1 The time it takes to complete question 1 (in seconds).

3.2 The time it takes to complete question 2 (in seconds).

3.3 The number of milk substitutes the participant selects in question 3 (out of 10).

3.4 Were any milk substitutes selected in question 3 (yes or no)

Additional measures

Participant gender, age, highest education level, long-term dietary restrictions (i.e., following a vegan or dairy-free diet) baseline dairy milk consumption (Question: how many days a week do you consume dairy milk in an average week) and baseline milk substitute consumption (Question: How many days a week do you consume a milk substitute (i.e. a plant-based milk) in an average week) will also be collected so that balance between conditions can be viewed.

Sample size determination

We are unaware of sufficiently similar previous research that could inform any power calculation. We have therefore been guided primarily by results from (Gleckel, 2020) who showed that more participants understood, on a five-point scale, that “Vegan Butter” could be used on pasta, for baking biscuits and on toast, than “Vegan Spread”, resulting in Cohen’s d effect sizes of 0.5, 0.8 and 0.2 respectively.

However, as we are using outcome measures which have not previously been used, we also conducted a pilot study to estimate the pooled standard deviation (SD) for the primary and secondary outcome. The pilot study recruited 50 participants, of which 41 (82%) passed the attention check and were included in the following calculations.

Primary hypothesis

We consider that a meaningful difference between study conditions on the primary outcome would be one more correct answer in the Milk Labelling condition compared to the No Milk Labelling condition. Given a pooled SD of 2.88, estimated from the pilot study, this equates to an effect size of Cohen’s $d = 0.35$.

Therefore, using G*Power 3, we calculate that a total of 346 participants would allow us to detect an effect of this magnitude or greater, for a two-group independent t-test of the primary outcome, to test our primary hypothesis, with an alpha level of 5% and power of 90%. To allow for any attention check fails we will recruit an additional 20% equating to 416 participants (208 in each study condition).

Secondary hypothesis

We consider that a meaningful difference between study conditions on the secondary outcome would be one more misidentified milk substitute in the Milk Labelling condition compared to the No Milk Labelling condition. Given a pooled SD of 1.25, estimated from the pilot study, this equates to an effect size of Cohen’s $d = 0.8$.

Using G*Power 3, we calculate that a total of 68 participants would allow us to detect an effect of this magnitude or greater, for a two-group independent t-test of the secondary outcome, to test our secondary hypothesis, with alpha level of 5% and power of 90%.

We will therefore recruit 416 participants so that the study is powered to test both the primary and secondary hypotheses.

Statistical Plan

A CONSORT diagram will be produced showing the flow of participants through the study (see example in Appendix 2). A table will be produced showing additional measures between the study conditions. A second table will also be produced showing the mean number of correct answers given for each of the three different product categories (dairy milk, milk substitutes and other drinks), for each of the three questions, between study conditions.

Responses (yes, no and unsure) for each question will be presented descriptively for each milk substitute label (i.e., how many participants correctly identified 'soya drink' and how many correctly identified 'soya milk').

A GLM (general linear model), or similar, will be used to compare the means of the primary outcome (the number of yes answers given to milk substitutes [out of 10]) between the two study conditions. The difference in means with 95% CI (confidence interval) of the mean, t statistics and p-values will be reported. This will test the primary hypothesis.

The secondary outcome (the number of yes or unsure answers given to milk substitutes [out of 10]) and outcome 3.3 (the number of yes answers the participants gives to whether they themselves would choose a milk substitute to put in their tea or coffee in the future [out of 10]) will also use a GLM, or similar, to compare the means of each outcome between the two study conditions.

Pre-specified covariates of baseline dairy milk consumption and baseline milk substitute consumption will be added to each GLM.

Secondary outcomes 3.1 and 3.2 will be compared descriptively between study conditions. For each condition, the time it took to answer these questions will be reported separately for those who were 100% correct and those who were not 100% correct.

Secondary outcome 3.4 will be compared between study conditions using a logistic regression model to predict whether participants select a milk substitute (or not). The odds ratio (OR), along with a 95% and p value will be reported.

If any outcome is considered to be non-normally distributed, a bootstrapping p value with 95% CI, from 1000 bootstrap samples, will also be reported. Non-linear models will be considered, depending on the distribution of model residuals.

For sensitivity analyses, the primary outcome analyses will be repeated four times: 1. Excluding participants that guessed the true nature of the study, 2. Excluding those that reported being on a restricted diet (diary-free or vegan), 3. Excluding any participant that said yes to adding a 'other drink' (i.e., a fruit juice) to their tea or coffee', and 4. Excluding anyone who answered zero to the 'number of times the participants consume dairy milk in an average week'.

Any outliers for each outcome will be identified using range checks, scatter plots and histograms. True outliers will be defined as any value where the median absolute deviation exceeds 3. Any true outliers will be included in the primary analysis but, if deemed necessary, a sensitivity analysis will be completed without any true outliers to compare results.

Ethical Considerations and Informed Consent

Ethics approval has been obtained from the School of Psychological Science Research Ethics Committee at the University of Bristol (Approval Code: 10313). The participant will receive information at the start of the task. The study will be closed online once the required number of participants have been recruited. Participants will be given sufficient time to read the information, consider any implications, and raise any questions with the investigators prior to deciding to participate. Consent will then be obtained. Participants will be informed that they are free to withdraw at any time by closing their browser.

The study will be conducted according to the revised Declaration of Helsinki (2013) and the 1996 ICH Guidelines for Good Clinical Practice E6(R2). The participant information statement will explain the nature, purpose and risks of the study to the participant.

Safety

As this is an online experiment, we do not foresee any risks to participants.

Data Management

All aspects of the General Data Protection Regulation (GDPR) and the Data Protection Act (DPA) 2018 and Freedom of Information Act 2000 will be adhered to. All personal data will be treated as confidential.

Anonymised study data

Anonymous study data (i.e., participant Prolific IDs will be deidentified) will be shared with collaborators for the purposes of analysis and results interpretation under appropriate collaboration agreements. At the end of the study, anonymous electronic study data (including finalised data sheet) will be locked and made open using the University of Bristol Research Data Repository. Study data will be kept for a minimum of 20 years.

Quality Control and Quality Assurance

The investigators will be responsible for data quality. After approximately 10% of data collection has been completed, the study will undergo an in-house quality assessment by a second researcher.

Insurance

The University of Bristol has Public Liability Insurance (<http://www.bristol.ac.uk/secretary/insurance/liability-insurance/>) to cover the liability of the University to research participants.

Publication Policy

The findings from this research study may be published in an appropriate scientific journal (and made available open access), and/or presented at an appropriate meeting. Study data will be collected and held by the study investigators and may be used to inform future research. The data will be made available for sharing via a University of Bristol online data repository.

Study Personnel

Katie De-loyde
School of Psychological Science
12a Priory Rd
Bristol BS8 1TU
Email: kd16662@bristol.ac.uk

Marcus Munafò
School of Psychological Science
12a Priory Rd
Bristol BS8 1TU
Email: marcus.munaf@bristol.ac.uk

Angela Attwood
School of Psychological Science
12a Priory Rd
Bristol BS8 1TU
Email: angela.attwood@bristol.ac.uk

Olivia Maynard
School of Psychological Science
12a Priory Rd
Bristol BS8 1TU
Email: olivia.maynard@bristol.ac.uk

Funding Source

Personal research budget from School of Psychological Science provided to Olivia Maynard

Conflicts of Interest

Katie De-loyde follows, and advocates for, a plant-based diet. The other study investigators have no known conflicts of interest to declare.

References

Court of Justice of the European Union. (2017). Purely plant-based products [Online]. Available: <https://curia.europa.eu/jcms/upload/docs/application/pdf/2017-06/cp170063en.pdf>.

- Department for Environment: Food & Rural Affairs. (2013). Food standards: labelling and composition [Online]. Available: <https://www.gov.uk/guidance/food-standards-labelling-durability-and-composition>.
- Food and Drug Administration. (2018). Food and Drug Administration Extends Comment Period on Use of the Names of Dairy Foods in Labeling Plant-Based Products [Online]. Available: <https://www.fda.gov/food/cfsan-constituent-updates/fda-extends-comment-period-use-names-dairy-foods-labeling-plant-based-products>.
- Gleckel, A. (2020). Are Consumers Really Confused by Plant-Based Food Labels? An Empirical Study [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3727710.
- International Food Information Council Foundation. (2018). What's in a Name? Survey Explores Consumers' Comprehension of Milk and Non-Dairy Alternatives [Online]. Available: https://foodinsight.org/wp-content/uploads/2018/10/Milk-Nomenclature_PDF_1.pdf.
- Mintel. (2019). Milking the vegan trend: A quarter (23%) of Brits use plant-based milk [Online]. Available: <https://www.mintel.com/press-centre/food-and-drink/milking-the-vegan-trend-a-quarter-23-of-brits-use-plant-based-milk>.
- Welch, P. (2019). Dairy Pride Act [Online]. Available: <https://www.congress.gov/bill/116th-congress/house-bill/1769/text?q=%7B%22search%22%3A%5B%22dairy+pride%22%5D%7D&r=1&s=1>.
- Willett, W., Rockström, J., Loken, B., *et al.* (2019). Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems. *Lancet*. 393, 447-492.