

**ANALYSIS PLAN**

**The Impact of Health Warning Labels on Alcohol Selection: Study Two**

Contents

[Study design 3](#_Toc2772205)

[Hypotheses 3](#_Toc2772206)

[Aims 3](#_Toc2772207)

[Primary aim 3](#_Toc2772208)

[Secondary A aim 3](#_Toc2772209)

[Secondary B aim 4](#_Toc2772210)

[Outcomes 4](#_Toc2772211)

[Primary 4](#_Toc2772212)

[Secondary A 4](#_Toc2772213)

[Secondary B 4](#_Toc2772214)

[Additional measures 4](#_Toc2772215)

[Data collection 5](#_Toc2772216)

[Creating variables 5](#_Toc2772217)

[Using SPSS syntax 5](#_Toc2772218)

[Outliers 5](#_Toc2772219)

[Missing data 5](#_Toc2772220)

[Violations of assumptions 6](#_Toc2772221)

[Descriptive statistics 6](#_Toc2772222)

[Outcome analysis 6](#_Toc2772223)

[Primary aim 6](#_Toc2772224)

[Secondary A aim 6](#_Toc2772225)

[Secondary B aim 7](#_Toc2772226)

[Additional measures 7](#_Toc2772227)

[Exploratory analysis 7](#_Toc2772228)

# Study design

A between-subjects 2 (image: present vs. absent) x 2 (text: present vs. absent) factorial experimental design. Participants will be randomised to one of 4 possible experimental groups (Table 1).

**Table 1. Label Design Table**

|  |  |  |
| --- | --- | --- |
|  | **Image of adverse health consequence** | |
| **Health consequence** | Present | Absent |
| **Health consequence** | Group 1  Image and text  V:\Behaviour Change by Design\3. Studies\Study 8a10a - Online labelling\8. Other\Warning Label design\NewWordingalcoholLabels\NewWordingAlcohlLabels.Seperate\label-livercancer2.jpgPictorial HWL | Group 2  V:\Behaviour Change by Design\3. Studies\Study 8a10a - Online labelling\8. Other\Warning Label design\NewWordingalcoholLabels\NewWordingAlcohlLabels.Seperate\label-livercancer2.jpgText-only HWL |
| **No text** | Group 3  Image and no text  Pictorial HWL | Group 4  No image  No health consequence |

# Hypotheses

HWLs (health warning labels) presented as (a) text-only or (b) text and image, will decrease selection of alcoholic drinks.

# Aims

## Primary aim

To estimate the impact on selection of alcoholic beverages of HWLs warning of the adverse health consequences of excessive alcohol consumption presented as (a) text (b) images and (c) text and images.

## Secondary A aim

To assess the impact of each of the HWLs on emotional and cognitive responses (see secondary outcomes) to alcoholic beverages.

## Secondary B aim

To investigate the impact of cognitive resource (time pressure group vs. no time pressure) on selection of drinks with HWLs.

# Outcomes

## Primary

Proportion of participants selecting alcoholic beverages (beer or wine) in the selection task, i.e. 0 = did not select an alcoholic drink. 1 = did select an alcoholic drink.

## Secondary A

Negative emotional arousal. Assessed using the mean average of a four-item measure, with scores ranging from 1 (not at all afraid / worried / uncomfortable / disgusted) to 7 (very afraid / worried / uncomfortable / disgusted).

Reactance to the labels. Participants are asked to what extent the labels are annoying. Assessed with one item on a scale from 1 (not annoying) to 7 (extremely annoying).

Avoidance of the labels. Participants are asked to what extent they are likely to avoid the labels. Assessed on a scale from 1 (not at all likely) to 7 (extremely likely).

Perceived disease risk relating to drinking the alcoholic beverage (cancer, liver disease, perceived healthiness of the drink) assessed using a three-item measure, with the combined score being 3 (low risk) to 21 (high risk).

Acceptability of health warning labels. Assessed using a single-item measure, with a score of 1 (strongly oppose) to 7 (strongly agree). For this outcome, group 4 participants will be re randomised to one of the other three groups, and only participants from these three conditions will be analysed for this outcome.

## Secondary B

Selection in relation to cognitive resource manipulation. In a second between-subjects selection task, participants will be randomised to one of two cognitive resource manipulation groups (time pressure [3 seconds] group and no time pressure [60 seconds] group). This will make a 2 (image: present vs. absent) x 2 (text: present vs. absent) x 2 (time pressure vs. no time pressure) factorial design in total.

During this selection task participants will be asked six times to select a drink they would prefer to consume (between two choices: alcoholic drink vs. non- alcoholic drink).

Scores will range from 0 (no alcoholic drinks selected or neither product selected) to 6 (all 6 alcoholic drinks selected), calculated by the total number of times (out of six) participants select an alcoholic drink.

# Additional measures

Demographics: age, gender, ethnicity, education (highest level), household income, BMI (body mass index).

Weekly unit consumption, participants will be asked to enter the number of drinks they have consumed over the previous 7 days.

Heavy and binge drinking behaviours (AUDIT-C), three questions to detect heavy and binge drinking behaviour in a general population, with a total score of 0 (low risk) to 12 (high risk).

# Data collection

All data will be sent directly to the researcher completing the data analysis at the end of the study in an Excel spreadsheet. This data will already have excluded those participants who failed the attention check. A data dictionary will also be sent which includes all coding and ranges.

# Creating variables

## Using SPSS syntax

IBM SPSS version 24 will be used to create two variables (an image variable and a text variable). The basis of this code can be seen in Table 2.

**Table 2. New variables.**

|  |  |  |
| --- | --- | --- |
| **Variable to be created** | **Code to be used** | **Study arm coding (see Table 1 for study grouping)** |
| Image vs. no image | Yes image = 1  No image = 0 | Group 1 = 1  Group 2 = 0  Group 3 = 1  Group 4 = 0 |
| Text vs. no text variable | Yes text = 1  No text = 0 | Group 1 = 1  Group 2 = 1  Group 3 = 0  Group 4 = 0 |

## Outliers

Any outliers will be identified using range checks, scatter plots and histograms. Outliers should be minimal due to the design of the study but, if any are identified further checks will be performed by the research team to ensure they are not the result of data entry errors.

Any 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.

# Missing data

Any missing data will be coded as an impossible value (-999).

Data that are missing due to participants choosing ‘Rather not say’ option are coded as an impossible value (-888).

If an excessive amount of missing data is identified (>10%), the research team will be notified so that checks can be made.

The data for all outcomes should be fully complete due to the nature of the study. However, if missing data is >10% for any outcome variable, and is missing at random, multiple imputation will be used to manage missing data. Results will be reported for complete case data and imputed data. Otherwise (if missing data is <10% for the outcome variables), analysis will be on a complete case basis.

# Violations of assumptions

Normality for all secondary outcomes will be assessed using normality plots and statistical tests (Kolmogorov-Smirnov and Shapiro-Wilk).

If strong non-normality is demonstrated in any of the secondary outcomes the analysis plan will not change but bootstrap p values will be reported. If there is evidence of significance, then bootstrap confidence intervals (CI) of the mean difference (MD) will also be reported.

# Descriptive statistics

Demographic data of the participants, and descriptive statistics for the primary dichotomous outcome and all the secondary outcomes (and additional measures), will be reported in a table with four columns, one for each of the four study HWL groups.

Numbers and percentages will be presented for categorical variables, including the primary outcome. Means and standard deviations (SD) will be presented for continuous variables. The median will also be reported for variables with clearly non-normal distributions (see *Violations of Assumptions*).

# Outcome analysis

All analysis will be done in IBM SPSS version 24. Analysis will be coded in syntax and will be reproducible at any time (this will be added as Appendix A after the analysis is complete).

## Primary aim

Logistic regression will be used to assess the relative odds of selecting an alcoholic beverage when a text or image warning is present, compared to when these warnings are absent.

The model will utilise the 2 x 2 design of the study with two independent variables: 1) image vs. no image and 2) text vs. no text. The dependent variable will be the primary outcome (was an alcoholic drink selected).

The initial model will include main effects of text and image respectively, and the interaction between them. Unless the interaction term attains evidence of a high degree of statistical significance (*p* < 0.01), it will be dropped in favour of a single model which includes only the two main effects, which will then be estimated with greater precision. To assess the effect of text warning, it will then be possible to pool data from those who saw the image and those who did not see the image. Similarly, to assess the effect of image warning, it will also be possible to pool data from those who saw the text and those who did not see the text.

If a significant interaction is detected, we will use the model including the interaction term to assess (i) the effect of the text warning in the absence of an image (ii) the effect of the image in the absence of the text warning (iii) the effect of the text and image warning in conjunction with each other. Each effect will be reported as an odds ratio (OR) with 95% CI, along with an associated p-value.

## Secondary A aim

For four of the five secondary A outcomes (negative emotional arousal, reactance, avoidance, disease risk) a general linear model, using a 2 x 2 design, will be used to assess differences between the HWL groups. This will involve four separate general linear models (two-way ANOVA [analysis of variance] models) ‑ one for each of the four secondary A outcomes (negative emotional arousal, reactance, avoidance, disease risk). The dependent variable will be the secondary A outcome, and there will be two independent variables: 1) image vs. no image and 2) text vs. no text.

As with the primary outcome, the initial models will be applied which includes an interaction term – if its p-value exceeds 0.01, a simpler model only including main effects will be applied. To assess the effect of text warning, it will then be possible to pool data from those who saw the image and those who did not see the image. Similarly, to assess the effect of image warning, it will also be possible to pool data from those who saw the text and those who did not see the text.

If a significant interaction is detected, we will then use the model including the interaction term to assess (i) the effect of the text warning in the absence of an image (ii) the effect of the image in the absence of the text warning (iii) the effect of the text and image warning in conjunction with each other

Each effect will be reported as a difference in means with 95% CIs, F statistics and p-values. An effect size (Cohen d) will also be calculated and presented alongside 95% CI of the effect size.

For analysis of the remaining secondary A outcome (acceptability) a one-way ANOVA will be conducted. The dependent variable will be acceptability and there will be one independent variable: HWL group (image and text, text-only, image-only). The reference category will be the text only group.

## Secondary B aim

A general linear model, using a 2 x 2 x 2 design, will be used to assess the differences in the secondary B outcome between the two time pressure groups (time pressure vs. no time pressure) and the four HWL groups.

This will involve a general linear model representing a three-way ANOVA. The dependent variable will be the secondary B outcome (the amount of times out of six an alcoholic drink or no drink was selected), and there will be three independent variables: 1) image vs. no image; 2) text vs. no text; and 3) time pressure vs. no time pressure.

The effect of the time pressure will be pooled across participants who are or are not exposed to the text warning or image. The effect size will be reported as a difference in means, reported with 95% CIs, F statistics and p-values. An effect size (Cohen d) will also be calculated and presented alongside 95% CI of the effect size.

## Additional measures

All additional measures will be reported in the demographic table to be sure they are adequately balanced between groups, but will not be used during statistical analysis.

## Exploratory analysis

Where differences in the primary outcome are observed between study arms, mediation analysis (using a bias-corrected bootstrapping approach) will be conducted to assess the mediating role of negative emotional arousal between the label (independent variable) and selection (the outcome).