

Health warning labels and alcohol selection: a randomised controlled experiment in a naturalistic shopping laboratory

Natasha Clarke¹ , Anna K. M. Blackwell² , Katie De-loyde² , Emily Pechey¹ , Alice Hobson¹ , Mark Pilling¹ , Richard W. Morris³, Theresa M. Marteau¹  & Gareth J. Hollands¹ 

Behaviour and Health Research Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK,¹ Tobacco and Alcohol Research Group, School of Psychological Science, University of Bristol, Bristol, UK² and Bristol Medical School, University of Bristol, Bristol, UK³

ABSTRACT

Background and aims Health warning labels (HWLs) on tobacco products reduce smoking. There is an absence of evidence concerning the impact of alcohol HWLs on selection or purchasing in naturalistic settings. Using a commercial-standard naturalistic shopping laboratory, this study aimed to estimate the impact on selection of alcoholic drinks of HWLs describing adverse health consequences of excessive alcohol consumption. **Design** A between-subjects randomised experiment with three groups was conducted: group 1: image-and-text HWL; group 2: text-only HWL; group 3: no HWL. **Setting** A commercial-standard naturalistic shopping laboratory in the United Kingdom. **Participants** Adults ($n = 399$, 55% female) over the age of 18 years, who purchased beer or wine weekly to drink at home. **Interventions** Participants were randomised to one of three groups varying in the HWL displayed on the packaging of the alcoholic drinks: (i) image-and-text HWL ($n = 135$); (ii) text-only HWL ($n = 129$); (iii) no HWL ($n = 135$). Participants completed a shopping task, selecting items from a range of alcoholic and non-alcoholic drinks, and snacks. **Measurement** The primary outcome was the proportion of alcoholic drinks selected. Secondary outcomes included HWL ratings on negative emotional arousal and label acceptability. **Findings** There was no clear evidence of a difference in the HWL groups for the percentage of drinks selected that were alcoholic compared to no HWL (44%); image-and-text HWL: 46% (odds ratio [OR] = 1.08, 95% confidence interval [CI] = 0.82, 1.42); text-only HWL: 41% (OR = 0.87, 95% CI = 0.67, 1.14). Concordant with there being no difference between groups, there was extreme evidence in favour of the null hypothesis (Bayes factor [BF] < 0.01). Negative emotional arousal was higher ($P < 0.001$) and acceptability lower ($P < 0.001$) in the image-and-text HWL group, compared to the text-only HWL group. **Conclusions** In a naturalistic shopping laboratory, there was no evidence that health warning labels describing the adverse health consequences of excessive alcohol consumption changed selection behaviour.

Keywords Alcohol, choice architecture, graphic warnings, health warning label, labelling, pictorial health warning label.

Correspondence to: Natasha Clarke, Behaviour and Health Research Unit, Institute of Public Health, University of Cambridge, Cambridge, UK.

E-mail: ncc42@medschl.cam.ac.uk

Submitted 3 September 2020; initial review completed 15 December 2020; final version accepted 7 April 2021

INTRODUCTION

Excessive consumption of alcohol causes a significant population health burden, increasing the risk of a range of diseases including liver disease, heart disease and many cancers [1,2]. An objective of the "Global Strategy to Reduce the Harmful Use of Alcohol" is to reduce harmful use by 10% by 2025 [3]. One potential

intervention to discourage excessive alcohol consumption is the use of health warning labels (HWLs) on packaging. There is strong evidence that HWLs on tobacco are both effective and a feasible population-level intervention [4,5], with larger effects of image-and-text HWLs than text-only HWLs [6–8], and similar effects evident among those in more and less deprived groups [9]. Given this, their potential application to alcohol

products has been suggested by academics and public health bodies [10–13], although such bodies have also highlighted that caution should be applied when comparing evidence across products [11].

In the United Kingdom, the only mandatory health information required on alcohol product labels is alcohol strength by volume (ABV), the net quantity of the drink, and allergens [14]. There are further best-practice recommendations to include three key elements: unit content, pregnancy warnings and a link to the industry-funded Drinkaware website [15]. However, even if this voluntary information is included labels are typically small, placed on the back of product packaging and may attract only minimal consumer attention in their current form [16]. These factors probably contribute to the ineffectiveness of current alcohol labelling [17]. Elements that may increase label effectiveness include larger labels and health warnings beyond those aimed at pregnant women [18]. Providing more highly visible health risks on labels may increase the currently low awareness of harms, such as the link between alcohol and cancer [19].

Evidence from online studies suggests that HWLs can reduce hypothetical selection of alcoholic drinks, with the addition of an image increasing effectiveness, an effect which may be mediated by negative emotional arousal [20]. There is limited evidence available in laboratory settings for the effectiveness of HWLs, with only one study to date looking at consumption [21]. Initial studies suggest that both image-and-text and text-only HWLs are effective at decreasing consumption rate [21], with image-and-text HWLs exerting larger effects on consumption intentions [22]. However, HWLs with images are also the least acceptable [20], and evidence suggests they may be less acceptable on alcohol products than they are on food products [23]. A lack of awareness of the health risks associated with alcohol may contribute to the low acceptability of HWLs, but acceptability can be increased by raising awareness of harms associated with drinking [24].

Importantly, there is an absence of evidence on the impact of both text-only and image-and-text HWLs on alcohol selection or consumption in naturalistic settings more likely to reflect real-world alcohol-related behaviour [25]. Assessment of this is particularly key, given that effect sizes observed within preliminary online studies may not translate to laboratory and field settings [25]. Naturalistic settings—such as bar laboratories—are frequently used in alcohol consumption research to mimic typical drinking environments [26]. Recently, artificial shopping laboratories have been used to study tobacco [27,28] and sugar-sweetened beverage purchasing behaviour, with HWLs reducing the likelihood of purchasing these drinks [29]. Such settings provide greater ecological validity than studies conducted online. Moreover, they offer an alternative for testing interventions — such as aversive product

labels — where a lack of palatability to commercial stakeholders is a barrier to identifying viable field settings.

The current study is the first to use a naturalistic shopping laboratory (Shopper Lab) to investigate the impact of alcohol HWLs on selection behaviour. The primary aim of the study was to estimate the impact on selection of alcoholic drinks displaying image-and-text or text-only HWLs describing adverse health consequences of excessive alcohol consumption. It was hypothesised that image-and-text and text-only HWLs placed on bottles and cans of beer and wine would reduce their selection compared to no HWLs. Secondary aims were (1) to assess the impact of image-and-text and text-only HWLs on alcohol quasi-purchasing behaviour (i.e. total and proportional spend on alcohol and alcohol units bought) and (2) to describe emotional responses to, and acceptability of, image-and-text and text-only HWLs.

METHODS

The study protocol and a detailed statistical analysis plan were pre-registered (ISRCTN18275963; OSF:<https://osf.io/hbdg3/>).

Design

The study used a between-subjects design. Participants were randomly allocated to one of three possible experimental groups: Group 1: image-and-text HWL; Group 2: text-only HWL; Group 3: no HWL (Box 1).

Randomisation

Stock in the shopping laboratory could only be changed at most twice a day, and could not be re-configured for every participant. Participants were therefore randomised to attend an experimental session consistent with their group allocation. Sessions were allocated using stratified randomisation and participants were allocated using simple randomisation. The statistician conducting the randomisation was blind to group assignment. Allocation was double-blind, with neither the recruiter nor the participants aware of the intervention assignment; although the recruiter was aware that the participant was in group A, B or C, they were unaware of to which label group this corresponded. See Supporting information, S1, Figs S1 and S2) for further randomisation details.

Setting

The study was conducted in a naturalistic shopping laboratory setting in February 2020 (Blue Yonder Shopper Lab). This is a 4 × 7.6 m room, designed to mimic a real supermarket store, and is used specifically for commercial research purposes. Products were branded and prices

Box 1. Study design.

Label type		
<p>Group 1 Image-and-text HWL</p> 	<p>Group 2 Text-only HWL</p> 	<p>Group 3 No HWL</p> 

N.B. HWLs were displayed on branded products on top of usual product packaging (see supplementary material Figure S5 for more example images).

consistent with those in UK stores. The Shopper Lab displayed drinks (alcoholic and non-alcoholic) and snacks with product selection based on current supermarket stock in English stores, except for non-alcoholic beer and wine, for which more options were available than are usually provided. Shelves were re-stocked after each participant session. Further details and photographs of the set-up can be found in the Supporting information, S3, Figs S3 and S4.

Participants

Participants were adults (18+) residing in the United Kingdom, who consumed beer or wine regularly (i.e. at least once a week), purchased supermarket products (food and drink) for their household and shopped in a physical store at least monthly. Participants attended the study in person and were recruited via Blue Yonder Research Ltd; the agency has a pool of approximately 10 000 members with a wide range of social characteristics.

Based on previous research on alcohol HWLs [20], a sample was recruited to detect a difference of at least 16 percentage points for the proportion of participants selecting an alcoholic beverage between no HWL and the HWL groups [80% power, $\alpha = 0.025$ (applying Bonferroni adjustment for two comparisons to the no HWL group)]. At least 130 participants per group was required (390

participants in total). Given uncertainty concerning the translation of effect sizes seen in online settings to real-world laboratory settings, a larger sample size would have been preferable. However, there was no prior comparable study to aid us in estimating effect sizes, and available resources also dictated that this represented the largest possible sample size.

Interventions

Health warning labels

Depending on the allocated group, the alcoholic products either displayed HWLs (image-and-text or text-only) or no HWLs, applied directly on the product. The same products were used in each group. Within the HWL groups, different variants of HWLs (i.e. type of health consequence described on the HWL) were used to increase variety, maximise engagement and increase potential impact of at least one HWL across our sample. This was also informed by tobacco guidelines specifying that rotating warnings are most effective. Image-and-text and text-only HWL variants described the same six health consequences. For example, both included a breast cancer HWL and a liver cancer HWL. The specific warnings used were created by a graphic designer and were based on previous studies highlighting these HWLs as being most effective in

increasing negative emotions [23] and decreasing the odds of selecting alcohol [20]. Illustrative examples of labelled alcohol products are included in the Supporting information, S5, Fig. S5 and the full range of labels are available at <https://osf.io/gye5n/>.

Labelling interventions are classed as 'Information' interventions that focus on the 'Product' of interest in the TIPPME intervention typology (Typology of Interventions in Proximal Physical Micro-Environments [30]).

Outcomes

Primary outcome

This was the proportion of total drinks selected that were alcoholic.

Secondary outcomes

Alcohol unit indices.

- *Mean alcohol units from all drinks selected.* This was assessed by dividing the total number of alcohol units selected by the total number of drinks selected [e.g. if two bottles of wine were selected (each 10 units) and two bottles of soft drink were selected: $20/4 = 5$ mean units of alcohol].
- *Number of alcohol units selected.* This was the total number of alcohol units contained within the selected alcohol products.

Alcohol quasi-purchasing indices.

- *Total spend (£) on alcoholic drinks.* This was the total spend on alcoholic drinks.
- *Spend (£) on alcoholic drinks as a proportion of total spend.* This was the total spend on alcoholic drinks.

Alcohol selection.

- *Proportion of participants selecting at least one alcoholic beverage.* This was a binary outcome of 'did not select an alcoholic beverage' or 'did select at least one alcoholic beverage' (0 or 1, respectively).

Post intervention measures (assessed after the shopping task).

- *Negative emotional arousal.* This was assessed using the mean of a four-item measure, previously used to assess the impact of warning labels on cigarette packages [31] and adapted for alcohol HWL studies [20,23]. Responses were rated on 7-point scales: 'How (afraid/worried/uncomfortable/disgusted) does the label on this drink make you feel?' [1, not at all (afraid/worried/uncomfortable/disgusted) to 7, very (afraid/worried/uncomfortable/disgusted)].
- *Acceptability of health warning labels.* This was assessed using one item on a 7-point scale, adapted from previous research assessing the impact of sugar tax [32] and alcohol HWLs [20]: 'Do you support or oppose putting this

label on alcoholic drinks?' (Strongly oppose—neither oppose nor support—strongly support). Ratings past the scale mid-point (indicating neither acceptable nor unacceptable), i.e. > 4 , indicated that the label was acceptable.

Procedure

Ethical approval was granted by the Cambridge Psychology Research Ethics Committee (PRE.2019.058).

A pilot study was conducted ($n = 10$) to ensure that the setting and task were representative of a real shopping experience and that the task wording was clear.

Participants were selected from the research agency's existing pool of participants and were first invited by e-mail to complete an initial screening questionnaire. To disguise the true aim of the study, it was described as an experiment on 'shopping habits and behaviour' and the screening questionnaire contained filler questions on general shopping behaviours and snack consumption. If deemed fully eligible, participants were randomised to a specific HWL group and an appropriate time for study participation was arranged (see Randomisation section for details). Upon arrival, after providing informed consent, eligible participants answered questions regarding their demographic characteristics [age, gender, ethnicity, education (highest level), income], before completing the shopping task.

Participants were given the instructions: 'Imagine you are going to do your weekly shop with your own money, including buying drinks and snacks. Shop how you would usually shop for drinks and snacks, and if there is something that you would not be happy to buy, you do not have to buy it'. Participants were informed that all prices reflected typical market value and to imagine that they were spending their own money when making their selection (no minimum or maximum spend), but that they were not expected to purchase the items. Participants were given a shopping basket to browse the Shopper Lab aisles and then selected the products they wanted to purchase (this took no longer than 10 minutes). Depending on their allocated condition, the alcoholic drinks displayed an image-and-text HWL, text-only HWL or no HWLs (Box 1). In the HWL groups each alcoholic drink displayed one of six (image-and-text or text-only) different warnings on their packaging.

After the shopping task, participants left behind their selected items and were asked if they noticed any labels on the products. Participants who noticed labels were asked to describe them. In a separate room, all participants were then randomised to view one of the six text-only and one of the six image-and-text HWLs (in a counterbalanced order) on a beer or wine bottle (depending on their preference) and rated each HWL on negative emotional arousal

and acceptability. Further demographic details—height, weight and drinking characteristics [Alcohol Use Disorders Identification Test (AUDIT-C) [33], weekly consumption]—were recorded. Participants were debriefed, answered questions on the study aims and were given the opportunity to provide feedback. Participants were reimbursed £20 for their time.

Statistical analysis (Supporting information, material S6)

The primary outcome was the proportion of drinks selected that were alcoholic [number of alcoholic drinks divided by the total number of drinks (alcoholic plus non-alcoholic)]. This was analysed using a binomial generalised linear mixed model. The Bayes factor was calculated from the Bayesian information criteria for models with and without the experimental condition [34].

For the secondary outcomes of mean units of alcohol selected, total units of alcohol selected, total spend on alcoholic drinks and spend on alcoholic drinks as a proportion of total spend a generalised linear mixed model was used.

For the secondary outcome of the odds of selecting at least one alcoholic beverage, a logistic mixed model was used.

Negative emotional arousal and acceptability were analysed as difference scores using two separate general linear models. Adjustment was made for previous study group allocation by adding this variable as a covariate.

General free text comments about the HWLs or the study were coded and thematically analysed following the six steps outlined by Braun & Clarke (2006) [35]. Themes were agreed between all authors.

RESULTS

In total, 399 participants were randomised, all of whom completed the study. Figure 1 shows the flow of participants through the study and Table 1 their characteristics across groups. Just over half the sample were female (55%) and the mean age was 39.9 years [standard deviation (SD) = 13.7; median = 46; range = 24–61]. Groups were well balanced on most characteristics. Due to concerns regarding differences between study groups of weekly consumption of alcohol (Table 1), and because it was

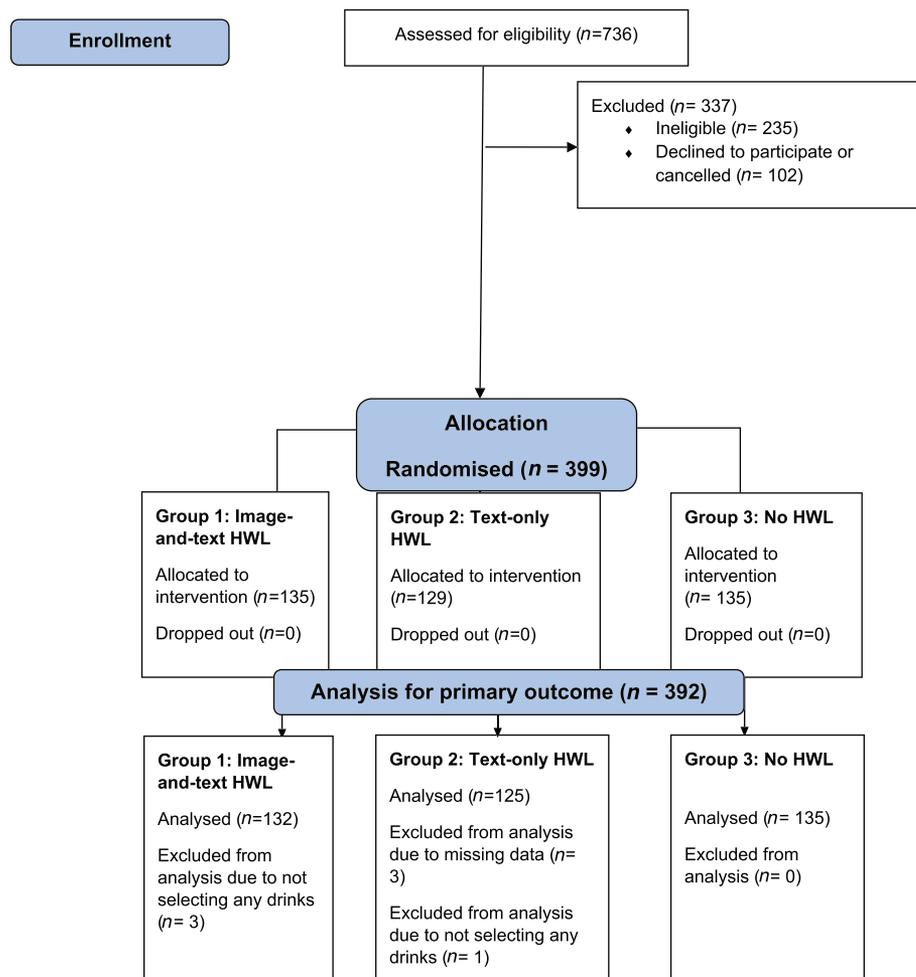


Figure 1 Flow of participants through study

Table 1 Participant characteristics (*n* %, unless otherwise stated).

	Study condition		
	Group 1: Image-and-text HWL <i>n</i> = 135	Group 2: Text-only HWL <i>n</i> = 129	Group 3: Control (no HWL) <i>n</i> = 135
Weekly alcohol consumption (units) ^a			
Mean (SD)	15.6 (11.9)	17.5 (11.9)	19.4 (15.9)
0–14	82 (61%)	68 (53%)	62 (46%)
15–30	36 (27%)	45 (35%)	50 (37%)
31–50	16 (12%)	13 (10%)	16 (12%)
51+	1 (<1%)	3 (2%)	7 (5%)
AUDIT-C ^b score (mean ± SD)	6.2 (2.1)	6.7 (2.1)	6.7 (2.2)
Age, mean (SD)	40.4 (13.6)	39.0 (13.9)	40.3 (13.5)
18–39 years	61 (45%)	67 (52%)	63 (47%)
40–59 years	55 (41%)	41 (32%)	53 (39%)
60 and over	19 (14%)	21 (16%)	19 (14%)
Gender			
Male	59 (44%)	64 (50%)	57 (42%)
Female	76 (56%)	65 (50%)	78 (58%)
Ethnicity			
White	131 (97%)	120 (93%)	127 (94%)
Mixed	3 (2%)	2 (2%)	5 (4%)
Asian	0 (0%)	4 (3%)	2 (2%)
Black	1 (<1%)	3 (2%)	1 (<1%)
Highest qualification			
No qualifications	4 (3%)	7 (5%)	3 (2%)
Up to 4 GCSEs	24 (18%)	34 (26%)	41 (30%)
5 or more GCSEs or 1 A-level	43 (32%)	31 (24%)	32 (24%)
2 or more A-levels	23 (17%)	19 (15%)	20 (15%)
Bachelor's degree	25 (18%)	20 (16%)	26 (19%)
Post-graduate degree or qualification	16 (12%)	18 (14%)	13 (10%)

SD = standard deviation; HWL = health warning label. ^aAll participants in the sample explicitly reported drinking at least once a week in the screener questions. A further weekly drinking measure recorded the amount of alcohol consumed in the previous week as an overall indication of the volume of alcohol consumed weekly. ^bHeavy and binge drinking behaviours [Alcohol Use Disorders Identification Test (AUDIT-C) [33]], 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) [33].

highly related to the outcome, this measure was added as a continuous covariate to the primary outcome analysis, although its inclusion did not alter conclusions. Three participants were excluded from primary and secondary outcome analysis due to missing data.

Primary outcome (see Tables 2 and 3)

Compared to no HWL, modelling suggested no clear evidence of a difference in the HWL groups for the percentage of drinks selected that were alcoholic: image-and-text HWL: 46%; text-only HWL: 41%; no HWL: 44% (adjusted values, see Table 2 for unadjusted means). Concordant with there being no difference between groups in this specific study context, there was extreme evidence in favour of the null hypothesis [Bayes factor (BF) < 0.01].

Secondary outcomes

Raw secondary outcome data are presented in Table 2 and modelled estimates in Table 3. Compared to no HWL, modelling suggested no clear evidence of a difference in the HWL groups for any of the secondary outcomes.

HWL ratings (Table 4)

Negative emotional arousal ratings were higher in the image-and-text HWL group compared to the text-only HWL group. Acceptability was lower in the image-and-text HWL group compared to the text-only HWL group.

Based on a score of 5 or more on the 7-point scale, 53% (*n* = 213) of participants indicated that they supported text-only HWLs and 41% (*n* = 163) supported image-and-text HWLs.

Table 2 Primary and secondary outcomes [unadjusted mean^a (SD) and *n* %].

	Group 1: Image-and-text HWL <i>n</i> = 135	Group 2: Text-only HWL <i>n</i> = 126	Group 3: Control (no HWL) <i>n</i> = 135
	Mean (SD)		
Primary outcome: proportion of alcoholic drinks ^a	0.50 (0.29)	0.46 (0.26)	0.45 (0.28)
Secondary outcomes			
Mean number of alcohol units selected ^b	3.33 (2.71)	2.99 (2.54)	3.12 (3.05)
Total number of alcohol units selected	12.23 (9.93)	11.93 (8.92)	12.33 (12.04)
Total spend on alcoholic drinks, in £	9.09 (7.03)	10.52 (11.68)	9.56 (9.87)
Total spend on alcoholic drinks as a proportion of total spend ^a	0.68 (0.30)	0.68 (0.28)	0.66 (0.32)
	<i>n</i> (%)		
Number of participants who selected at least one alcoholic drink	117 (87)	114 (91)	114 (84)

SD = standard deviation; HWL = health warning label. ^aAdjusted means are reported in the manuscript. ^bFor these outcomes, four participants are excluded (three in the image-and-text HWL group, one in the text-only HWL group), as they did not select any drinks, i.e. these outcomes would have been calculated as 0/0).

Table 3 Estimated mean differences (95% CI) or odds ratios (95% CI) for HWL groups compared to no HWL (*n* = 135) for primary and secondary outcomes.

	Group 1: Image-and-text HWL <i>n</i> = 135	Group 2: Text-only HWL <i>n</i> = 126
	OR (95% CI), <i>P</i> -value	
Primary outcome: proportion of alcoholic drinks ^{a,b}	1.08 (0.82, 1.42), <i>P</i> = 0.55	0.87 (0.67, 1.14), <i>P</i> = 0.31
Secondary outcomes	MD (95% CI), <i>P</i> -value, effect size (Cohen's <i>d</i>) ^c	
Mean number of alcohol units selected ^a	0.22 (−0.47, 0.91) <i>P</i> = 0.53, <i>d</i> = 0.07	−0.13 (−0.82, 0.56) <i>P</i> = 0.71, <i>d</i> = −0.04
Total number of alcohol units selected	−0.09 (−2.65, 2.48) <i>P</i> = 0.95, <i>d</i> = −0.01	−0.37 (−2.90, 2.16) <i>P</i> = 0.77, <i>d</i> = −0.04
Total spend on alcoholic drinks, in £	−0.43 (−2.41, 1.56) <i>P</i> = 0.67, <i>d</i> = −0.05	1.15 (−1.43, 3.72) <i>P</i> = 0.38, <i>d</i> = 0.11
Total spend on alcoholic drinks as a proportion of total spend ^a	0.03 (−0.04, 0.11) <i>P</i> = 0.36, <i>d</i> = 0.09	0.01 (−0.06, 0.09) <i>P</i> = 0.73, <i>d</i> = 0.03
	OR (95% CI), <i>P</i> -value	
Number of participants who selected at least one alcoholic drink	0.85 (0.42, 1.70), <i>P</i> = 0.64	0.63 (0.28, 1.42), <i>P</i> = 0.27

CI = confidence interval; HWL = health warning label; MD = mean difference; OR = odds ratio. ^aFor these outcomes, four participants are excluded (three in the image-and-text HWL group, one in the text-only HWL group), as they did not select any drinks, i.e. these outcomes would have been calculated as 0/0). ^bThis model also included covariates for time of study completion: (i) weekend and (ii) afternoon. Model Akaike's information criterion (AIC) = 1196.5. ^cEstimated means and robust standard errors reported in Supporting information, Table S7c.

Table 4 Estimated mean differences (95% CI) for negative emotional arousal and acceptability ratings—within subjects ($n = 399$).

	Raw mean (SD)	Estimated MD (95% CI) ^a	P-value ^a	t-statistic (d.f.) ^a
Negative emotional arousal				
Image-and-text HWL	4.29 (1.67)	1.04 (0.87, 1.22)	< 0.001	$t(396) = 11.72$
Text-only HWL	3.25 (1.46)	–	–	
Acceptability				
Image-and-text HWL	3.64 (1.95)	–0.68 (–0.91, –0.41)	< 0.001	$t(396) = -5.15$
Text-only HWL	4.32 (1.74)	–	–	

CI = confidence interval; d.f. = degrees of freedom; HWL = health warning labels; MD = mean difference; SD = standard deviation. ^aEstimated using a general linear model of the difference scores between image-and-text and text-only, with adjustment for previous study group allocation and clustering (18-level session variable) as a nested random variable.

Table 5 Themes identified in free-text comments.

Theme	Subtheme	Description	Examples
Theme 1: Drinkers' position on the acceptability of HWLs	General support for HWLs	Health warning labels (HWLs) were received positively and described as a good idea; for example, participants mentioned that they were necessary or even thought they should be a legal requirement. Some said they should be used because they have worked on tobacco. Some thought only the text-only HWLs were acceptable	'They are a great idea'
	General opposition to HWLs	HWLs were received negatively and described as a bad idea. Generally, images were less accepted than text-only HWLs. References to nanny-state and that drinking should be an individual choice. Terms such as 'unnecessary', 'patronising' and 'scaremongering' used	'I personally do not like them and feel we are in danger of living in a nanny state'
	Unintended consequences	HWLs may have unintended consequences, such as a negative impact on children exposed to them, increased stigma or judgement potentially leading to depression, shame or anxiety around drinking. The potential negative impact on sales in pubs and small alcohol businesses were also mentioned as an adverse effect of the labels	'I feel products should have warning labels on but the images are too far, as a parent I would not feel comfortable with my children seeing images like this in the supermarket'
	Drinking context and culture	References to the value of alcohol in society, as a part of culture. HWLs as suitable for certain types of drinks (such as cheap, strong, bulk buys), but not for premium products (such as wine, real ales or an expensive whisky). Drinking was highlighted as a social and enjoyable activity, which participants did not want ruined by the labels. Many participants expressed a sense of unfairness that their responsible and 'civilised' drinking would be punished, when in their eyes it is specific contexts of drinking, e.g. binge drinking, which should be targeted. Similarly, participants described that alcohol is often given as gifts or consumed as social events in pubs or bars	'I find them distasteful. A bottle of wine is a beautiful thing and is spoiled by them'
Theme 2: Potential impact of the HWLs	How people behave	Participants highlighted they perceived HWLs as being effective in deterring themselves or	'I think labels would make me drink less'

(Continues)

Table 5. (Continued)

Theme	Subtheme	Description	Examples
		other people from purchasing or drinking alcohol	
	How people feel	Feelings of alarm, anxiety, anger, discomfort, disgust or fear, either expressed at an individual level or as a reference to how others might react to HWLs. References to negative emotional arousal particularly common	'Yes they are disgusting but I have my own mind and they would not put me off'
	How people think	HWLs would make participants think twice, or increase knowledge or awareness	'It might make you think twice about how much alcohol you drink/units'
	Message saliency	HWLs described as bold, hard-hitting, graphic, ugly or unpleasant, and attention to HWLs highlighted	'Very strong and hard hitting'
	Ineffectiveness of labels	Participants said they did not notice or attend to the labels, or would avoid them. They also raised concerns that there would be wear-out effects over time, or uncertainty as to any potential impact of the HWLs	'As with warnings on cigarette packets I do not think having warnings on alcohol bottles will stop people drinking'
Theme 3: Alcohol-related risk perception	Reflections on health communication	Comments related to the believability of the health messages. Suggestions or reflections on how these messages should be best communicated were made. There was some acknowledgment of an existing awareness of the health risks. In terms of the HWLs, there were suggestions that the message could be improved by adding statistics or risk information to the health message, or adding alternative information, such as %ABV, nutritional or unit information. These comments also included suggestions for alternative methods, such as information campaigns on TV or billboards, advertising, education and comparisons with tobacco approaches, such as plain packaging	'I am acutely aware of the consequences of my actions consuming adult beverages' 'I do not believe there is any real evidence that alcohol causes cancer or heart disease'
	Personal relevance of health information	The relevance of the information and individual susceptibility to health risks was questioned. Participants often included references to their own consumption levels and purchasing (as responsible drinkers). There were suggestions that the information might only be suitable for heavier drinkers—i.e. 'alcoholics'	'Makes you think. Although I am not worried about my alcohol intake'

Additional analyses

See Supporting information, S7 for additional analyses including sensitivity analyses and thematic analysis (themes in Table 5). Supporting information, Table S7d shows a model that includes participant characteristics as covariates, with conclusions unchanged.

DISCUSSION

In a naturalistic shopping laboratory, health warning labels (HWLs) describing the adverse health consequences of excessive alcohol consumption did not change selection

behaviour. There was no clear evidence of a difference between groups for the primary outcome—the proportion of alcoholic drinks selected—or for any of the secondary outcome measures.

These findings are inconsistent with prior studies that show promising effects of alcohol HWLs on selection in online settings using identical HWLs [20], on sales in field settings [13] and on consumption in laboratory settings [21]. They are also not consistent with evidence that suggests HWLs on other products, such as sugary drinks, can change purchasing behaviour in a similar shopping laboratory context [29].

There are four possible explanations for the current findings. First, HWLs could impact selection in a real-world setting, but the shopping laboratory setting and study task lacked ecological validity. While the quasi-purchasing task used in the current study was equivalent to the standard task used widely in the commercial sector and the majority of participants indicated that the shopping task felt at least somewhat typical compared to their usual shop, we are unaware of any evidence in the public domain that directly tests its validity. In contrast to other studies conducted in similar settings [29], participants did not exchange money or keep selected items. This was not possible in the present study, as alcohol cannot be sold in unlicensed premises [36]. The exchange of money could equally have introduced other biases, such as choices based on product prices. For example, another study using a similar design [29] provided an incentive to each participant to purchase products and the participant kept the remaining money. Depending on the price of the products and the value participants attached to cash, this might have increased or decreased purchasing of the study products.

A second possibility related to the study task is that HWLs may impact more opportunistic behaviours, such as selecting an item for immediate consumption, as typically measured in online settings [20,37], rather than as part of a typical weekly shop. This may have encouraged relatively routine or planned shopping behaviour, where participants were motivated to select products that they usually purchased (other studies showing effects on sugary drink purchasing have not specified a weekly shop in task wording [29,38]). Such behaviour may be less likely to be influenced by environmental cues such as packaging.

A third possible explanation for the null effect in the current study is that a brief exposure to HWLs placed on real alcohol products at a single time-point does not affect selection or has an effect that the current study was underpowered to detect. While the majority of participants indicated that they noticed the HWLs, future studies could assess differences in responses to HWLs in greater detail by measuring viewing time. The evidence that HWLs on sugary drinks can have an immediate impact on selection in a similar setting [29] raises the possibility that responses to HWLs on alcohol may differ from responses to HWLs on other products [39]. We should be cautious when comparing across products, as they are different substances associated with specific behaviours, even though alcohol shares some properties with both sugary drinks (foodstuff) and tobacco (drug) [11]. Moreover, our findings do not obviate the possibility that repeated exposure to HWLs could impact behaviour over the longer term, as has been observed in multiple studies in tobacco [6,40,41]. It is also possible that HWLs could affect other drinking behaviours, such as the rate at which alcohol is consumed [21].

Finally, the ineffectiveness of the HWLs in changing selection in the current study could be due to participants being relatively light drinkers: just over half the participants in the current study reported drinking within the low-risk drinking guideline amount of up to 14 units per week [42]. However, this is a similar sample in terms of drinking characteristics to our previous online study, where HWLs reduced selection [20] and AUDIT-C scores of the majority of the sample (83%) are indicative of more risky drinking (score of 5+). An exploratory analysis also found that including participants' characteristics as covariates—including alcohol consumption—did not change model conclusions. While drinkers at all levels of risk should be aware of potential harms [43], there may be differences in responses to HWLs by consumption or perceived relevance of health harms, warranting further study.

Despite there being no impact upon selection behaviour, HWLs affected emotions and cognitions. Negative emotional arousal was higher, and acceptability lower, in the image-and-text HWL group, compared to the text-only HWL group, mirroring findings from online settings [20]. Negative emotional arousal has been shown to be a potential mediator of the effect of HWLs on selection of alcohol and other products in online settings [20,37,44]. In the current setting it may be that increases in negative emotional arousal were insufficient, or other conditions are required, for an effect upon selection behaviour to be realised.

Approximately half the sample (53%) perceived text-only HWLs as acceptable, and this was lower (41%) for image-and-text HWLs. In free-text comments, some participants advocated their implementation and others thought they were unacceptable, in accordance with other studies that suggest mixed acceptability for alcohol HWLs [23,45]. Acceptability of alcohol policies may be higher with an increase in knowledge of health harms, specifically that alcohol can cause cancer [24]. Acceptability of an intervention, such as HWLs, is also higher when evidence for its effectiveness is communicated [46].

Thematic analysis of free-text comments highlighted that participants perceived HWLs as having potential effects on other cognitions including awareness, attention and knowledge. The evidence as to whether these changes lead to consistent changes in behaviour is variable. Previous research suggests that alcohol HWLs may improve recall and health-related knowledge [47], and such cognitions may be precursors to behaviour change [40,48]. However, labelling on other products has been found to change cognitions (e.g. implicit attitudes), but these changes are not necessarily associated with changes in behaviour [49,50]. Future research should investigate whether, when and under which conditions such cognitions and emotions translate to behaviour change. Furthermore, knowledge of the alcohol-cancer link is low [19,51], and many participants in the current study expressed surprise at the depicted

health risks. As such, changes in cognitions are potentially positive outcomes of HWLs in their own right.

Free-text comments also highlighted that the HWLs used in the present study may be suitable for heavy or problem drinkers, but not for 'sensible' drinkers such as themselves—an 'othering' effect which has been observed in other qualitative studies of alcohol interventions [52]. This emphasises the need to communicate health consequences associated with drinking at all consumption levels, together with low-risk drinking guidelines. References to alcohol's particular cultural significance to participants, and the pleasure obtained from drinking, were also common. This aligns with previous research highlighting the importance of considering the culture of drinking and different drinking practices when assessing the effectiveness of potential interventions [53]. Additional in-depth qualitative work could further illuminate the apparent ineffectiveness of the HWLs.

Strengths and limitations

This pre-registered, randomised study is the first to investigate alcohol HWLs in a commercial-standard naturalistic shopping laboratory setting, with HWLs displayed across a wide selection of real products. It also has some limitations.

First, although real products were used, participants did not keep the products or exchange money. Although commonplace in such settings for commercial research, as discussed above, this may have decreased ecological validity. In addition, participants were aware that they were being observed, which may have further affected ecological validity and potentially altered behaviour. Secondly, given that our sample size was based on effect sizes from an online setting [20], it is possible the study was underpowered. However, the results do not suggest effects in the expected direction and the Bayes factor suggests extreme evidence for no difference between groups.

Implications for future research

Our findings provide limited evidence that HWLs on alcohol may not be effective in altering selection in a naturalistic setting. While actual purchasing may elicit its own biases, on balance we judge that stronger evidence would be provided by studies that use objective purchasing measures—i.e. involving exchange of participants' own money—with products selected for consumption.

This study provides further evidence that effects of HWLs on real products may differ to effects seen in online studies using images of products [25]. Testing these interventions in naturalistic or field settings comes with clear challenges, as recently evidenced in Canada, where researchers were forced to withdraw cancer labels from alcohol products due to industry pressure [54]. HWLs are

likely to be unpalatable to commercial stakeholders and wider industry pressures can be expected if such labels are considered for implementation.

Should further research confirm that HWLs are likely ineffective in the immediate term, they still have potential to be effective in the longer term, both directly through repeated exposure to product packaging, or indirectly, such as by resulting in changes to industry behaviours, e.g. through increased availability of lower strength or alcohol-free products that would not require HWLs. Such industry responses have been seen, for example, in relation to the UK sugar levy resulting in widespread re-formulation of sugary drinks to avoid crossing the taxable threshold for sugar content [55] and following the implementation of food labelling interventions [56].

CONCLUSIONS

In a naturalistic shopping laboratory, there was no evidence that health warning labels reduced selection of alcoholic drinks. These types of labels applied to alcohol products may be ineffective in changing selection behaviour, or the study as designed was unable to detect their effect.

Clinical trial registration

ISRCTN18275963; OSF: <https://osf.io/hbdg3/>.

Declaration of interests

None.

Acknowledgements

This research was funded in whole, or in part, by the Wellcome Trust (ref: 206853/Z/17/Z). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Data availability statement

The dataset generated and analysed during the current study are available on the Open Science Framework project page: <https://osf.io/j9b63/> and the University of Cambridge Research Repository: <https://doi.org/10.17863/CAM.66856>.

References

1. Rehm J., Guiraud J., Poultais R., Shield K. D. Alcohol dependence and very high risk level of alcohol consumption: a life-threatening and debilitating disease. *Addict Biol* 2018; 23: 961–8.
2. Sheron N., Gilmore I. Effect of policy, economics, and the changing alcohol marketplace on alcohol related deaths in England and Wales. *BMJ Clin Res* 2016; 353: i1860.

3. World Health Organization (WHO). Global strategy to reduce the harmful use of alcohol. 2015. Available at: <https://www.who.int/publications/i/item/9789241599931>
4. Hammond D. Tobacco Labelling and Packaging Toolkit: A guide to FCTC Article 11. 2009; Available at: https://www.who.int/fctc/guidelines/adopted/article_11/en/. Accessed September 2020.
5. Hammond D., McDonald P., Fong G., Stephen Brown K., Cameron R. The impact of cigarette warning labels and smoke-free bylaws on smoking cessation: evidence from former smokers. *Can J Public Health* 2003; **95**: 201–4.
6. Brewer N. T., Hall M. G., Noar S. M., Parada H., Stein-Seroussi A., Bach L. E., et al. Effect of pictorial cigarette pack warnings on changes in smoking behavior a randomised clinical trial. *JAMA Intern Med* 2016; **176**: 905–12.
7. Hammond D. Health warning messages on tobacco products: a review. *Tob Control* 2011; **20**: 327–37.
8. Noar S. M., Hall M. G., Francis D. B., Ribisl K. M., Pepper J. K., Brewer N. T. Pictorial cigarette pack warnings: a meta-analysis of experimental studies. *Tob Control* 2016; **25**: 341–54.
9. Thrasher J. E., Carpenter M. J., Andrews J. O., Gray K. M., Alberg A. J., Navarro A., et al. Cigarette warning label policy alternatives and smoking-related health disparities. *Am J Prev Med* 2012; **43**: 590–600.
10. Blackwell A. K. M., Drax K., Attwood A. S., Munafò M. R., Maynard O. M. Informing drinkers: can current UK alcohol labels be improved? *Drug Alcohol Depend* 2018; **192**: 163–70.
11. Martin-Moreno J. M., Harris M. E., Breda J., Møller L., Alfonso-Sanchez J. L., Gorgojo L. Enhanced labelling on alcoholic drinks: reviewing the evidence to guide alcohol policy. *Eur J Public Health* 2013; **23**: 1082–7.
12. Royal Society for Public Health (RSPH). Labelling the point: towards better alcohol health information. 2018. Available at: <https://www.rsph.org.uk/our-work/policy/drugs/labelling-the-point.html>. Accessed September 2020.
13. Zhao J., Stockwell T., Vallance K., Hobin E. The effects of alcohol warning labels on population alcohol consumption: an interrupted time series analysis of alcohol sales in Yukon, Canada. *J Stud Alcohol Drugs* 2020; **81**: 225–37.
14. Food Standards Agency. Food Standards Agency—Food labelling e-learning course. Available at: https://labellingtraining.food.gov.uk/module10/overview_1.html. Accessed September 2020.
15. Portman Group. Alcohol Marketing and Sponsorship. 2019. Available at: <https://www.portmangroup.org.uk/marketing-toolkit/>. Accessed September 2020.
16. Kersbergen I., Field M. Alcohol consumers' attention to warning labels and brand information on alcohol packaging: findings from cross-sectional and experimental studies. *BMC Public Health* 2017; **17**: 123.
17. Burton R., Henn C., Lavoie D., O'Connor R., Perkins C., Sweeney K., et al. A rapid evidence review of the effectiveness and cost-effectiveness of alcohol control policies: an English perspective. *Lancet* 2017; **389**: 1558–80.
18. Jané-Llopis E, Kokole D, Neufeld M, Hasan OSM, Rehm J. WHO Health Evidence Network Synthesis Report 68. 2020;112. Available at: <https://apps.who.int/iris/bitstream/handle/10665/332129/9789289054898-eng.pdf?ua=1>
19. Buykx P. An investigation of public knowledge of the link between alcohol and cancer. *BMC Public Health* 2015; **16**: 1194.
20. Clarke N., Pechey E., Mantzari E., Blackwell A. K. M., De-loyde K., Morris R. W., et al. Impact of health warning labels communicating the risk of cancer on alcohol selection: an online experimental study. *Addiction* 2021; **116**: 41–52.
21. Stafford L. D., Salmon J. Alcohol health warnings can influence the speed of consumption. *J Public Health* 2017; **25**: 147–54.
22. Wigg S., Stafford L. D. Health warnings on alcoholic beverages: perceptions of the health risks and intentions towards alcohol consumption. *PLOS ONE* 2016; **4**: e0153027.
23. Pechey E., Clarke N., Mantzari E., Blackwell A. K. M., De-Loyde K., Morris R. W., et al. Image-and-text health warning labels on alcohol and food: potential effectiveness and acceptability. *BMC Public Health* 2020; **20**: 376.
24. Weerasinghe A., Schoueri-Mychasiw N., Vallance K., Stockwell T., Hammond D., McGavock J., et al. Improving knowledge that alcohol can cause cancer is associated with consumer support for alcohol policies: findings from a real-world alcohol labelling study. *Int J Environ Res Public Health* 2020; **17**: –398.
25. Clarke N., Pechey E., Kosite D., König L. M., Mantzari E., Blackwell A. K. M., et al. Impact of health warning labels on selection and consumption of food and alcohol products: systematic review with meta-analysis. *Health Psychol Rev* 2020; 1–39.
26. Kersbergen I., Oldham M., Jones A., Field M., Angus C., Robinson E. Reducing the standard serving size of alcoholic beverages prompts reductions in alcohol consumption. *Addiction* 2018; **113**: 1598–608.
27. Setodji C. M., Martino S. C., Gong M., Dunbar M. S., Kusuke D., Sicker A., et al. How do tobacco power walls influence adolescents? A study of mediating mechanisms. *Health Psychol* 2018; **37**: 188–93.
28. Shadel W. G., Martino S. C., Setodji C. M., Dunbar M., Scharf D., Creswell K. G. Do graphic health warning labels on cigarette packages deter purchases at point-of-sale? An experiment with adult smokers. *Health Educ Res* 2019; **34**: 321–31.
29. Grummon A., Taillie L. S., Golden S., Hall M., Ranney L., Brewer N. Impact of sugar-sweetened beverage health warnings on beverage purchases: a randomised controlled trial (OR28–01-19). *Curr Dev Nutr* 2019; nzz042. OR28–01–19.
30. Hollands G. J., Bignardi G., Johnston M., Kelly M. P., Ogilvie D., Petticrew M., et al. The TIPPE intervention typology for changing environments to change behaviour. *Nat Hum Behav* 2017; **1**: 0140.
31. Kees J., Burton S., Craig Andrews J., Kozup J. Understanding how graphic pictorial warnings work on cigarette packaging. *J Public Policy Mark* 2010; **29**: 265–76.
32. Reynolds J. P., Archer S., Pilling M., Kenny M., Hollands G. J., Marteau T. M. Public acceptability of nudging and taxing to reduce consumption of alcohol, tobacco, and food: a population-based survey experiment. *Soc Sci Med* 2019; **236**: 112395.
33. Bush K., Kivlahan D. R., McDonell M. B., Fihn S. D., Bradley K. A. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory care quality improvement project (ACQUIP). Alcohol use disorders identification test. *Arch Intern Med* 1998; **158**: 1789–95.

34. Lindeløv JK. How to compute Bayes factors using *lm*, *lmer*, *BayesFactor*, *brms*, and *JAGS/stan/pymc3*. 2018. Available at: <https://vbn.aau.dk/da/publications/how-to-compute-bayes-factors-using-lm-lmer-bayesfactor-brms-and-j>
35. Braun V., Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006; **3**: 77–101.
36. Institute of Alcohol Studies. Licensing Laws. Available at: <http://www.ias.org.uk/Alcohol-knowledge-centre/Availability-and-licensing/Factsheets/Licensing-laws.aspx>. Accessed September 2020.
37. Mantzari E., Vasiljevic M., Turney I., Pilling M., Marteau T. Impact of warning labels on sugar-sweetened beverages on parental selection: an online experimental study. *Prev Med Rep* 2018; **12**: 259–67.
38. Donnelly G. E., Zatz L. Y., Svirsky D., John L. K. The effect of graphic warnings on sugary-drink purchasing. *Psychol Sci* 2018; **29**: 1321–33.
39. Hall M. G., Grummon A. H., Lazard A. J., Maynard O. M., Taillie L. S. Reactions to graphic and text health warnings for cigarettes, sugar-sweetened beverages, and alcohol: an online randomised experiment of US adults. *Prev Med* 2020; 106120.
40. Hammond D., Fong G. T., McDonald P. W., Cameron R., Brown K. S. Impact of the graphic Canadian warning labels on adult smoking behaviour. *Tob Control* 2003; **12**: 391–5.
41. Schüz N., Eid M., Schüz B., Ferguson S. G. Immediate effects of plain packaging health warnings on quitting intention and potential mediators: results from two ecological momentary assessment studies. *Psychol Addict Behav* 2016; **30**: 220–8.
42. Department of Health. UK Chief Medical Officers' Alcohol Guidelines Review. Summary of the proposed new guidelines. 2016. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/489795/summary.pdf. Accessed September 2020.
43. Sherk A., Thomas G., Churchill S., Stockwell T. Does drinking within low-risk guidelines prevent harm? Implications for high-income countries using the international model of alcohol harms and policies. *J Stud Alcohol Drugs* 2020; **81**: 352–61.
44. Cho Y. J., Thrasher J. F., Yong H.-H., Szklo A. S., O'Connor R. J., Bansal-Travers M., *et al.* Path analysis of warning label effects on negative emotions and quit attempts: a longitudinal study of smokers in Australia, Canada, Mexico, and the US. *Soc Sci Med* 2018; **197**: 226–34.
45. Vallance K., Stockwell T., Zhao J., Shokar S., Schoueri-Mychasiw N., Hammond D., *et al.* Baseline assessment of alcohol-related knowledge of and support for alcohol warning labels among alcohol consumers in northern Canada and associations with key sociodemographic characteristics. *J Stud Alcohol Drugs* 2020; **81**: 238–48.
46. Reynolds J. P., Stautz K., Pilling M., van der Linden S., Marteau T. M. Communicating the effectiveness and ineffectiveness of government policies and their impact on public support: a systematic review with meta-analysis. *R Soc Open Sci* 2020; **7**: 190522.
47. Hobin E., Weerasinghe A., Vallance K., Hammond D., McGavock J., Greenfield T. K., *et al.* Testing alcohol labels as a tool to communicate cancer risk to drinkers: a real-world quasi-experimental study. *J Stud Alcohol Drugs* 2020; **81**: 249–61.
48. Brewer N. T., Parada H., Hall M. G., Boynton M. H., Noar S. M., Ribisl K. M. Understanding why pictorial cigarette pack warnings increase quit attempts. *Ann Behav Med* 2019; **53**: 232–43.
49. Asbridge S. C. M., Pechey E., Marteau T. M., Hollands G. J. Effects of pairing health warning labels with energy-dense snack foods on food choice and attitudes: online experimental study. *Appetite* 2021; **160**: 105090.
50. Crockett R. A., King S. E., Marteau T. M., Prevost A. T., Bignardi G., Roberts N. W., *et al.* Nutritional labelling for healthier food or non-alcoholic drink purchasing and consumption. *Cochrane Database Syst Rev* 2018; **2**: CD009315.
51. Winstock A. R., Holmes J., Ferris J. A., Davies E. L. Perceptions of alcohol health warning labels in a large international cross-sectional survey of people who drink alcohol. *Alcohol Alcohol* 2020; **16**:55: 315–22.
52. Bareham B. K., Kaner E., Spencer L. P., Hanratty B. Drinking in later life: a systematic review and thematic synthesis of qualitative studies exploring older people's perceptions and experiences. *Age Ageing* 2019; **48**: 134–46.
53. Ally A. K., Lovatt M., Meier P. S., Brennan A., Holmes J. Developing a social practice-based typology of British drinking culture in 2009–2011: implications for alcohol policy analysis. *Addiction* 2016; **111**: 1568–79.
54. Stockwell T., Solomon R., O'Brien P., Vallance K., Hobin E. Cancer warning labels on alcohol containers: a Consumer's right to know, a Government's responsibility to inform, and an Industry's power to thwart. *J Stud Alcohol Drugs* 2020; **81**: 284–92.
55. Scarborough P., Adhikari V., Harrington R. A., Elhussein A., Briggs A., Rayner M., *et al.* Impact of the announcement and implementation of the UK soft drinks industry levy on sugar content, price, product size and number of available soft drinks in the UK, 2015–19: a controlled interrupted time series analysis. *PLOS Med* 2020; **17**: e1003025.
56. Shangquan S., Afshin A., Shulkin M., Ma W., Marsden D., Smith J., *et al.* A meta-analysis of food labeling effects on consumer diet behaviors and industry practices. *Am J Prev Med* 2019; **56**: 300–14.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Study flow diagram with randomisations.

Figure S2. Session allocation - randomisation procedure.

Figure S3. Shopper Lab layout.

Figure S4. Images of the shopping laboratory.

Figure S5. Example HWLs.

Table S7a. Sensitivity analysis 1. Manipulation check for primary outcome – proportion of drinks selected that were alcoholic ($n = 307$).

Table S7b. Sensitivity analysis 2. Participants identified by onsite research agency team for primary outcome - proportion of drinks selected that were alcoholic ($n = 386$).

Table S7c. Estimated means (robust SEs) for continuous secondary outcomes ($n = 396$).

Table S7d. Exploratory analysis. Updated model results with additional covariates included – primary outcome.

Supplementary material

S1. Randomisation details

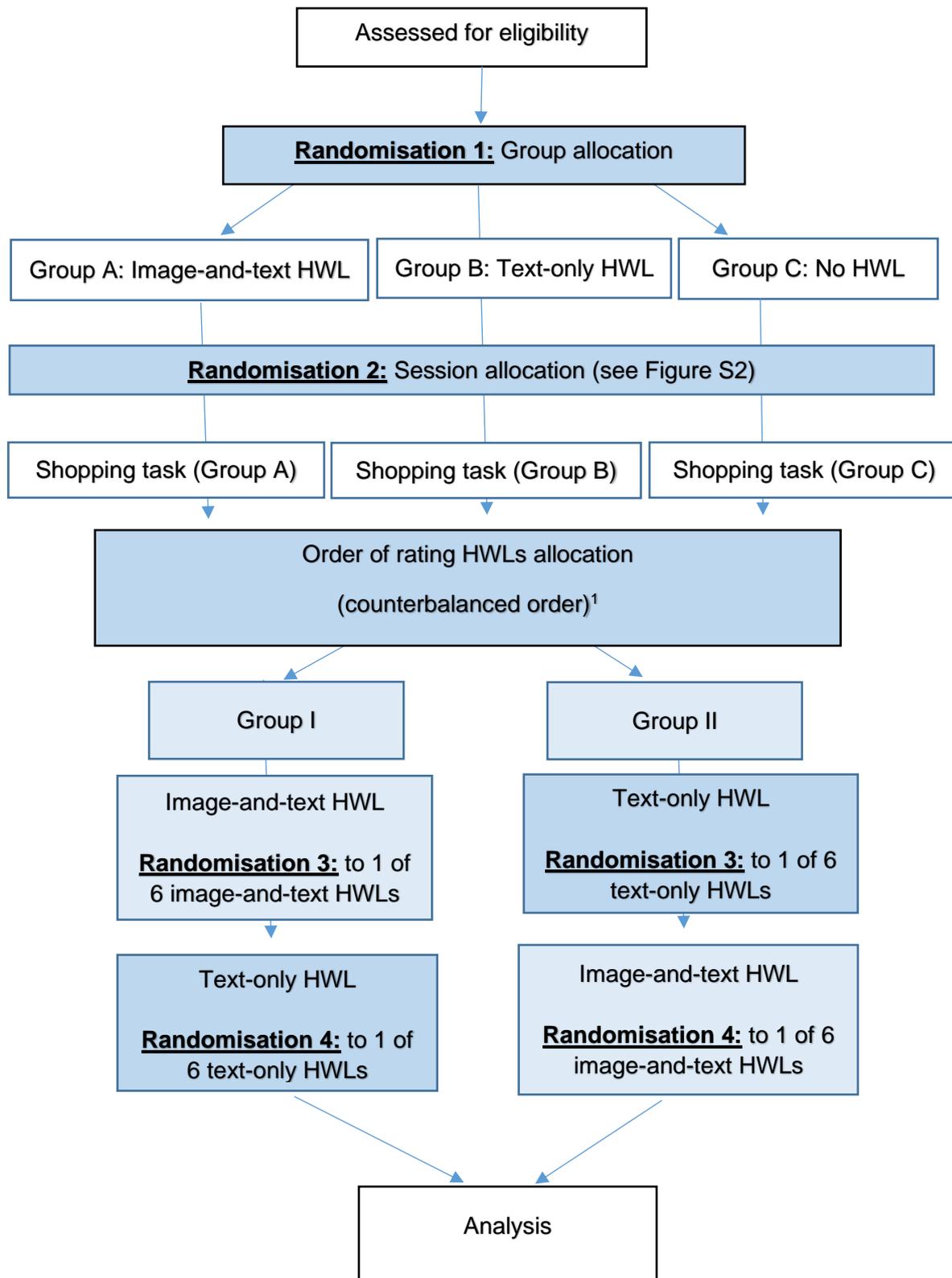
Session randomisation

The study took place across 18 part-day sessions, reflecting the capacity of the research agency to test the required number of participants. These comprised 12 sessions across six weekdays and six sessions across two weekend days. Each of the 18 sessions were allocated to display alcohol products in one of three ways according to the three study groups: displaying image-and-text HWLs, text-only HWLs or no HWLs. Stratified randomisation was used to allocate each of the 12 weekday sessions to one of six two-group combinations (AB,AC,BA,BC,CA,CB), avoiding combinations of the same groups occurring on the same day (i.e. AA,BB,CC). Each of the six weekend sessions were allocated to one of six three-group combinations that included all groups on any one day (ABC,ACB,BAC,BCA,CAB,CBA), omitting combinations that did not (e.g. AAA or AAB).

Participant allocation

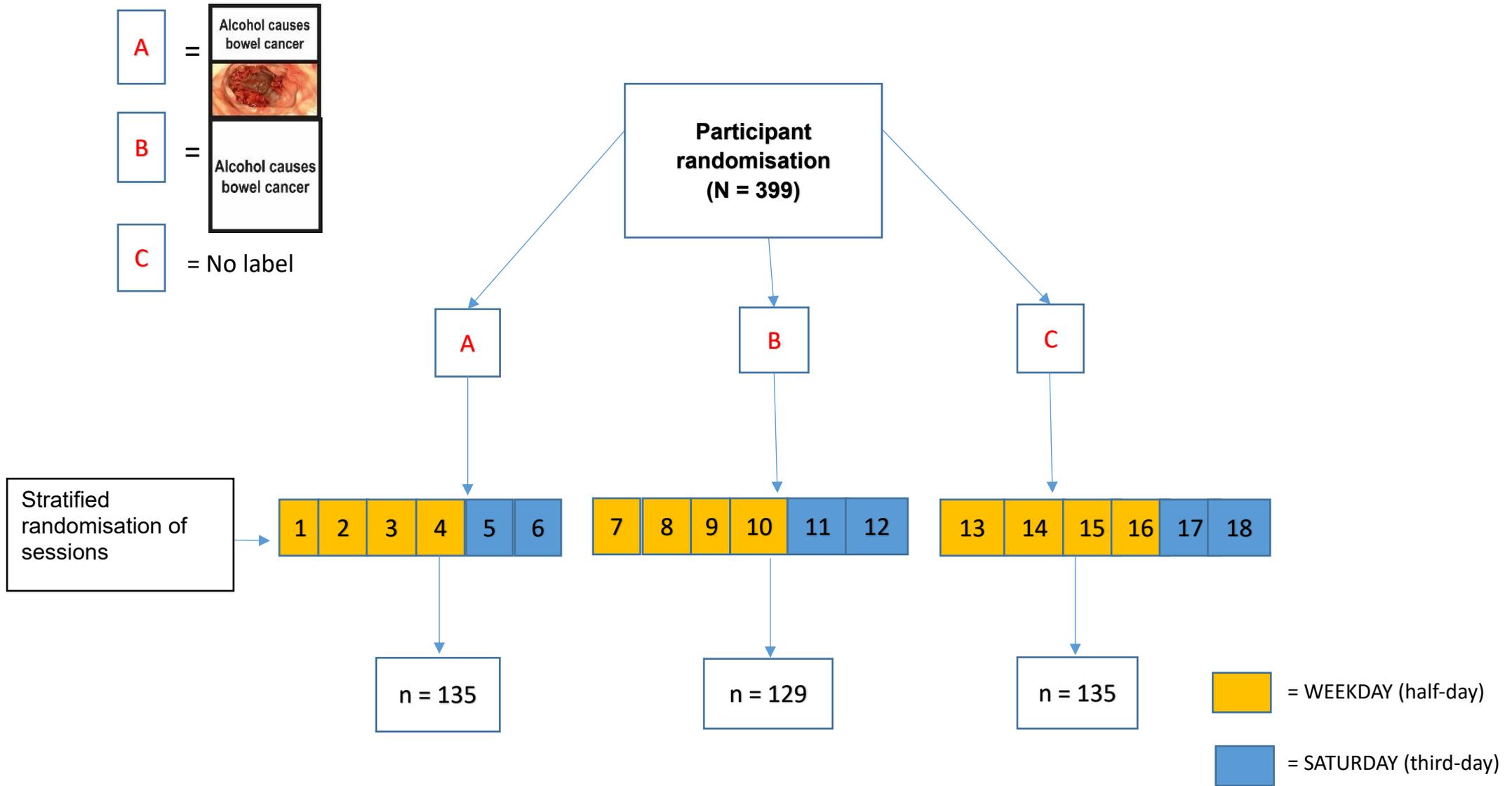
Eligible participants were randomly allocated using simple randomisation without stratification to one of the three groups. This occurred at the point of screening by the research agency. An eligibility criterion was being free to attend at least one session for each experimental group allocation. Once allocated to a group, participants' time slot for that group was confirmed.

Figure S1. Study flow diagram with randomisations



¹Difference from protocol: a pre-specified randomisation sequence was not used but the order in which the labels were presented was alternated and counterbalanced

Figure S2. Session allocation - randomisation procedure



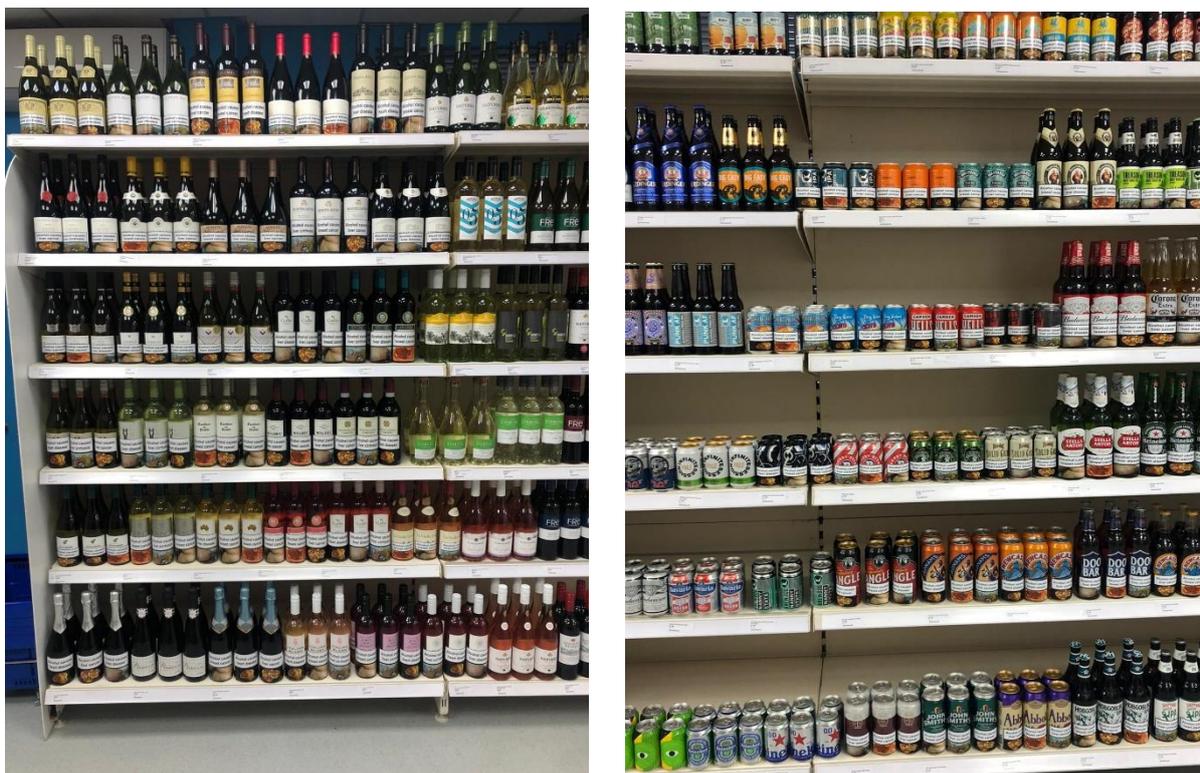
S3. Shopper Lab setup

The Shopper Lab displayed drinks (alcoholic and non-alcoholic) and snacks on two shelving units and snacks on three shelving units, similar to a convenience store set up, with products at least two deep on each shelf. The aisles displayed a range of premium and non-premium beer and wine in bottles and cans, soft drinks (branded and supermarket own labels) and alcohol-free beer and wine (144 different drink products in total: beer and wine (72 products), non-alcoholic beer and wine (36 products) and soft drinks (36 products)). There was an equal number of alcoholic and non-alcoholic options available. Snack products were displayed alongside alcohol products, in order to mimic real life stores more closely, and to mask the study aim, i.e. participants were asked to select both drinks and snacks in a shopping task. Snacks included nuts, popcorn, crackers, crisps, biscuits and confectionery.

Figure S3. Shopper Lab layout



Figure S4. Images of the shopping laboratory



S5. Details on HWL placement

There were six product units (e.g. six wine bottles) of each drink brand displayed on the shelves. In the HWL groups each product unit of the alcoholic drink brand displayed one of six image-and-text or text-only HWLs describing an adverse health consequence (bowel cancer, breast cancer, liver cancer, liver cirrhosis, liver disease, heart disease), with an equal number of each health consequence displayed (i.e. one of each health consequences on each of the six product units). The full range of labels can be found on the Open Science Framework (<https://osf.io/gye5n/>).

Figure S5. Example HWLs



S6. Statistical analysis

For all models, adjustment was made for: weekend or weekday (a binary variable); and morning or afternoon (as a binary variable); and clustering (18 level session variable) as a nested random variable. Evidence for residual clustering of outcomes within the 18 sessions was examined and the distributions of residuals were satisfactory. Adjusted means are reported in the text (unadjusted means in Table 2). Robust standard errors are reported. All models include the main effects only, as models were over-specified when interactions were included.

S7. Additional analyses

Sensitivity analysis

Two sensitivity analyses were conducted for the primary outcome: participants from the two HWL groups that answered ‘no’ to the manipulation check (‘Did you notice any labels on the products?’) (n = 85) and participants identified by the on-site research agency team (i.e. those that supplied nonsensical responses, did not adequately comply with the protocol, or did not engage with the task or follow instructions) (n = 6). Conclusions did not change.

Table S7a. Sensitivity analysis 1. Manipulation check for primary outcome – proportion of drinks selected that were alcoholic (n= 307)

Of the 392 participants contributing to the primary analysis, 85 were excluded for answering ‘no’ to the manipulation check (did you notice any labels on the products?).

	Estimated mean (robust SE) ¹	OR (95% CI) ¹	P value ¹
Image-and-text (n = 130)	0.46 (0.028)	1.18 (0.88, 1.57)	0.26
Text-only (n = 118)	0.40 (0.027)	1.02 (0.75, 1.39)	0.88
Control (n = 135)	0.44 (0.027)	-	-

¹ Main effects of study condition, weekend/weekday (a binary variable) and morning/afternoon (a binary variable), in relation to the timing of the shopping task, included in the model, with cluster variable.

Confidence interval (CI) Odds ratio (OR). Standard error (SE).

Table S7b. Sensitivity analysis 2. Participants identified by onsite research agency team for primary outcome - proportion of drinks selected that were alcoholic (n = 386)

Of the 392 participants contributing to the primary analysis, 6 participants were excluded as they were identified by the onsite research team as supplying nonsensical responses, not complying with the protocol, not engaging with the task or not following instructions.

	Estimated mean (robust SE) ¹	OR (95% CI) ¹	P value ¹
--	--	-----------------------------	----------------------

Image-and-text (n = 134)	0.45 (0.027)	1.07 (0.82, 1.40)	0.63
Text-only (n = 124)	0.41 (0.027)	0.90 (0.68, 1.17)	0.42
Control (n = 132)	0.44 (0.027)	-	-

¹ Main effects of study condition, weekend/weekday (a binary variable) and morning/afternoon (a binary variable), in relation to the timing of the shopping task, included in the model, with cluster variable.

Confidence interval (CI) Odds ratio (OR). Standard error (SE).

Ecological validity of setting

Overall, 57% (n = 228) of participants indicated that the shop felt at least somewhat typical (over 5 on the 10 point scale), with 28% (n = 116) giving a score of 8 or more. The mean score was 5.9 (SD = 2.4).

Based on free-text responses, approximately half of participants indicated that they thought the items were similar to those they would usually purchase. The reasons given for the shop not being similar were product or brand availability (most often related to specific snacks or soft drink products or brands being unavailable), shop layout (i.e. greater resemblance to a corner shop than a supermarket), pricing and references to the type of shop (i.e. more representative of a top-up weekend/evening than a weekly shop).

Thematic analysis

Three main themes were identified. The full descriptions of these themes and their subthemes are shown in Table 5.

Theme 1: Drinkers' position on the acceptability of HWLs, subthemes included general support for HWLs, general opposition to HWLs, unintended consequences of introducing HWLs and references to the drinking context and culture.

Theme 2: Potential impact of the HWLs, subthemes included how HWLs may impact how people behave, feel or think, message saliency and the potential ineffectiveness of labels.

Theme 3: Alcohol-related risk perception, subthemes included reflections on health communication and the personal relevance of health information.

Table S7c. Estimated means (robust SEs) for continuous secondary outcomes (n = 396)

	Group 1: Image-and-text HWL n = 135	Group 2: Text-only HWL n = 126	Group 3: Control n = 135
Secondary outcomes	Estimated mean (robust SE)		
Mean number of alcohol units selected ¹	3.18 (0.26)	2.83 (0.28)	2.96 (0.26)
Total number of alcohol units selected	11.01 (0.72)	10.78 (0.80)	11.10 (0.78)
Total spend on alcoholic drinks, in £	8.08 (0.66)	9.66 (1.04)	8.51 (0.85)
Total spend on alcoholic drinks as a proportion of total spend ¹	0.65 (0.03)	0.63 (0.03)	0.62 (0.03)

¹For these outcomes, 4 participants were excluded (3 in the image-and-text HWL group, 1 in the text-only HWL group) as they did not select any drinks, i.e. these outcomes would have been 0 / 0.

Table S7d. Exploratory analysis. Updated model results with additional covariates included – primary outcome

	OR (95% CI)¹	p value¹
Group 1: Image-and-text HWL (reference group = no HWL)	1.17 (0.89, 1.54)	0.26
Group 2: Text-only HWL (reference group = no HWL)	0.88 (0.67, 1.14)	0.33
Covariates		
Weekly alcohol consumed*	1.01 (1.01, 1.02)	< 0.001
Age (years)	1.00 (0.99, 1.01)	0.80
Gender (reference group = female)	1.13 (0.90, 1.43)	0.29
BMI category 1 (reference group = 6)	1.10 (0.20, 6.04)	0.91
BMI category 2 (reference group = 6)	0.51 (0.09, 2.95)	0.45
BMI category 3 (reference group = 6)	0.70 (0.13, 3.79)	0.67
BMI category 4 (reference group = 6)	0.67 (0.12, 0.36)	0.64

BMI category 5 (reference group = 6)	0.52 (0.10, 2.83)	0.45
Education category 1 (reference group = 6)	1.17 (0.61, 2.24)	0.63
Education category 2 (reference group = 6)	0.99 (0.66, 1.47)	0.95
Education category 3 (reference group = 6)	0.87 (0.59, 1.29)	0.49
Education category 4 (reference group = 6)	0.93 (0.62, 1.40)	0.73
Education category 5 (reference group = 6)	1.07 (0.71, 1.61)	0.74

BMI category (1=morbidly obese,...,6-underweight); Education category (1=None,...,6=Post-grad). Confidence interval (CI). Odds ratio (OR)

¹ This model also includes main effects for afternoon and weekend (as in original model in Table 3), with the additional covariates listed above. Model AIC (Akaike information criterion) = 1208.6. The original primary analysis model (from Table 3) AIC = 1196.4, indicating the model above was not an improvement on the original primary analysis.

* There was no evidence of an interaction between weekly alcohol consumed and HWL group.