
Article:
Kersbergen, I. orcid.org/0000-0002-8799-8963, Opazo Breton, M. orcid.org/0000-0002-1226-7541, Field, M. orcid.org/0000-0002-7790-5559 et al. (1 more author) (2022) The effect of alcohol packaging size and strength on U.K. alcohol consumers' classification of alcohol products as containing a single or multiple drinks. Psychology of Addictive Behaviors, 37 (7). pp. 969-976. ISSN 0893-164X

https://doi.org/10.1037/adb0000860

The effect of alcohol packaging size and strength on UK alcohol consumers' classification of alcohol products as containing a single or multiple drinks

Inge Kersbergen[1,*]

Magdalena Opazo Breton[1]

Matt Field[1]

Petra Meier[1,2]

[1] University of Sheffield, Sheffield, UK

[2] University of Glasgow, Glasgow, UK

* Corresponding author address: School of Health and Related Research; University of Sheffield; 30 Regent Street; Sheffield; S1 4DA; United Kingdom.

i.kersbergen@sheffield.ac.uk. +44 114 222 4293

Author note:

This research was funded by a Society for the Study of Addiction Griffith Edwards Academic Fellowship awarded to IK. This study's design and analysis plan were pre-registered; see https://osf.io/wyavx/. All anonymized data and analysis code are openly available at https://doi.org/10.15131/shef.data.17263604. The materials are available to researchers upon request at https://doi.org/10.15131/shef.data.17294252. The data in this manuscript were presented in June 2019 at the Kettil Bruun Society annual scientific meeting and in November 2019 at the Society for the Study of Addiction annual conference.

Abstract

Objective: Reductions to the size and strength of alcohol products prompt reductions in alcohol consumption, although these effects may be limited to single drinks rather than packages that contain multiple drinks. This study investigated what product characteristics predict whether a product is seen as a single drink and seeks to identify the thresholds beyond which products are considered to contain multiple drinks. **Methods:** Ninety-four UK drinkers from the Prolific participant panel categorized 250 alcohol products with varying packaging sizes and strengths into single or multiple drinks. We used multilevel logistic regression to investigate whether packaging size, strength, total alcohol content and container type predicted the likelihood that products were classified as a single drink across five drink types (beer, cider, ready-to-drink, spirits, wine). We used receiver operating characteristics curve analysis to identify the point at which products become too large or too strong to be considered a single drink by most drinkers. **Results:** Larger products, bottled drinks, products with higher ABV and higher alcohol content were more likely to be classified as containing multiple drinks. We report thresholds for packaging size, ABV and total alcohol content where products switch from being seen as a single drink to containing multiple drinks. The thresholds did not significantly differ between low risk and increased risk drinkers. **Conclusion:** The reported thresholds can help researchers and policy makers encourage more accurate self-monitoring of alcohol consumption. Future research should test whether single drink classifications moderate the effect of packaging size and strength reductions on alcohol consumption.

Keywords (up to 5): Alcohol, Product perceptions, Drink characteristics

Conflicts of interest: None

Public Significance Statement: This study showed that the size, strength and total alcohol content of alcohol products sold in UK stores affect whether these products are considered to

contain a single drink or multiple drinks. We report the different thresholds after which products are too large or too strong to be considered to contain a single drink. These thresholds can be used to encourage more accurate self-monitoring of alcohol consumption and to inform the development and evaluation of future interventions related to the size and strength of alcohol products.

The characteristics of alcoholic drinks such as serving size (total volume in a single drink), packaging size (total volume in a full alcohol container), and strength (alcohol by volume (ABV)) affect how much people drink. People consume less alcohol if they drink from smaller serving sizes (Kersbergen et al., 2018; Stevely et al., 2021), smaller wine bottles (Codling et al., 2020), and lower strength products (Perman-Howe et al., 2021). Therefore, changing these product characteristics could improve public health.

It is unclear which changes in product characteristics are likely to result in reduced consumption and to what extent those changes might lead to consumers (over)compensating for product changes. For example, using alcohol labels that highlight the low strength of the product led to increased consumption compared to the absence of such a label (Vasiljevic et al., 2018). Research on perceptions of food portion sizes demonstrates that people are likely to adjust their energy intake to compensate for portion size changes if portions are too large or too small to be considered to contain a single portion anymore (Haynes et al., 2019). Therefore, a reduction in the size of the food portion that is served is likely to yield reductions in total energy intake, but only if the smaller portion is still perceived as a whole portion by the consumer. There is insufficient research on the effect of alcohol packaging size reduction on consumption to assess whether this model can be generalized to alcohol consumption. If we generalized this model to the influence of packaging size on alcohol consumption, we would expect reductions in alcohol consumption by reducing the size of alcohol products that are considered to contain a single drink but a diminished effect of packaging size for alcohol products that are perceived as containing multiple drinks. It is important to understand what products are considered to contain a single or multiple drinks and what product characteristics predict this classification. This would allow future research to test whether classification as a single drink moderates the effect of packaging size or strength on alcohol consumption and inform future policy targeting these characteristics.

This information would also help us understand how people monitor their own alcohol consumption and compare them to drinking guidelines. Qualitative research showed that alcohol consumers preferred to count the number of drinks they consumed rather than the total number of standard drinks (Lovatt et al., 2015). Therefore, understanding how alcohol consumers judge whether alcohol containers contain a single or multiple drinks would show how alcohol consumers might understand drinking guidelines and regulate their drinking.

In this study, we investigated how product characteristics such as packaging size, strength, total alcohol content, and container type (e.g., cans, bottles) affected alcohol consumers' classification of packaged alcohol products as containing a single or multiple drink. First, we estimated the likelihood that alcohol products were classified as containing multiple drinks (versus single drinks) based on the products' characteristics. Then, we estimated the optimal threshold of the same product characteristics after which products are more likely to be classified as containing multiple drinks than a single drink. We also examined to what extent consumers' classifications and thresholds differed between low risk and increased risk drinkers.

Methods

Participants

We recruited 100 participants from the Prolific participant database (https://prolific.ac). Inclusion criteria were age ≥ 18, UK resident, and consumption of at least one UK unit of alcohol per week (8g of alcohol). We stratified recruitment by gender, annual household income and weekly alcohol consumption based on UK population statistics (Office for National Statistics, 2017; Osborne & Cooper, 2018). Our sample size was based on a power calculation (ipdpower package in Stata (Kontopantelis et al., 2016)) to detect a small effect of packaging size, hazardous drinking and their interaction on product classification

(Odds Ratio (OR) = 1.5 for main effects; OR = 1.2 for interactions) at > 95% power. The study received approval from the Research Ethics Committee at the University of Sheffield.

Stimuli

We extracted all alcohol products for sale on the websites of major UK supermarkets (Tesco, Sainsbury's, Asda). We extracted 3733 products with complete ABV and packaging size information. After removing alcohol-free products (<1% ABV; n = 125) and duplicates (n = 1190) we were left with 2418 unique products. We excluded packaging sizes that were more than 2.5 standard deviations (SD) above the drink type mean (i.e., 5L beer kegs, 2L cider bottles, 2.25L boxes of wine, 1.5L bottles of spirits, and 1.5L bottles of ready to drink products (RTDs; e.g., alcopops and pre-mixed cocktails) as we deemed it unlikely that anyone would consider these products to contain a single drink. Of the remaining 2290 products, we selected 100 products for five main drink types (beer, cider, RTDs, spirits, wine). We manually selected 80% of the products, ensuring that well-known brands and less common packaging sizes were represented in the sample, with the remaining 20% being selected at random (see Table 1 for product characteristics).

Procedure

After giving informed consent, participants were shown a random subset of 250 alcohol products in an online experiment in a randomized order (50 per drink type). Each product image was presented alongside a text description of the brand, packaging size and ABV (Figure 1). To ensure that participants could readily perceive size differences between products, the image size was proportional to the product's volume. For each product, participants reported whether the product contained 'one drink' or 'more than one drink'. We used attention checks to ensure that participants were paying attention to the products before responding (Oppenheimer et al., 2009): After 10 randomly selected products, participants

were asked to enter the brand name of the product on the previous trial. Then, participants completed a demographics questionnaire (age, gender, annual household income) and a measure of hazardous drinking (Alcohol Use Disorders Identification Test-consumption (AUDIT-c); Bush et al., 1998). The study took approximately 25 minutes. Participants received £4 reimbursement.

Measures

Our outcome variable was a binary variable based on participant's classification of alcohol product as either containing one (value "0") or multiple (value "1") drinks.

Our predictor variables were: packaging size (continuous variable in deciliters (100mls) ranging from 2.1dl to 7.5dl for beer; 2.5dl to 7.5dl for cider; 1.13dl to 10dl for RTDs; 0.2dl to 10dl for spirits; 1.87dl to 15dl for wine); ABV (continuous variable in % of alcohol by volume ranging from 1.2 to 9.2 for beer; 3.5 to 8.2 for cider; 3.5 to 11 for RTDs; 12 to 63 for spirits; 4 to 20 for wine); total alcohol content (continuous variable in UK units rounded to the nearest 0.5 unit ranging from 1 to 6 for beer; 1 to 5.5 for cider; 0.5 to 7.5 for RTDs; 0.5 to 43 for spirits; 1.5 to 18 for wine); container type (categorical variable based on the drink coming in a can, bottled, or other container type) and hazardous drinking (scale variable based on the AUDIT-c score ranging from 1 to 12). We also created interaction variables between hazardous drinking and packaging size and hazardous drinking and ABV to assess to what extent hazardous drinking moderates the effect of packaging size and ABV on drink classification. Based on Khadjesari et al (2017) we categorized participants as low-risk drinkers (AUDIT-c $<$ 8) and increased risk drinkers (AUDIT-c $\geq$ 8).

Statistical analysis

Drink classification

We estimated the likelihood of products being classified as containing multiple drinks, with packaging size, ABV, container type (can, bottled, or other container type), hazardous drinking and the interactions between hazardous drinking and packaging size and hazardous drinking and ABV as predictor variables using five multilevel logistic regressions with random intercepts and products nested within participants, one for each drink type category. We then repeated these analyses with total alcohol content as the predictor variable. The effect of total alcohol content on product classification was tested in separate models to the other predictor variables because it was highly correlated with packaging size (all r = .76 - .93) and ABV (all r = .23 - .85). All continuous variables were mean centered to aid interpretation of ORs and our confidence level was set to 99% ($\alpha = 0.01$), to adjust for multiple hypothesis testing.

Threshold determination

We used Receiver Operating Characteristics (ROC) curve analysis, accounting for clustering within participants (Liu & Wu, 2003), to determine thresholds for packaging size, ABV, and total alcohol content for classifying products as containing multiple drinks (vs a single drink). We used the area under the curve (AUC) to assess whether the predictor variables performed better than chance at classifying products as containing multiple drinks (AUC > 0.5). We used the maximized Youden Index (Fluss et al., 2005) to identify the highest level of packaging size, ABV, and total content at which products were considered to contain a single drink. The maximized Youden Index is the point at which the sum of the sensitivity and specificity is the greatest and therefore the threshold performs optimally at both avoiding false negatives (sensitivity) and avoiding false positives (specificity). We compared thresholds obtained with the maximized Youden Index with the highest level of packaging size, ABV, and total content at which products are considered to contain a single drink on >60% (Haynes et al., 2019) and >80% of observations. We also report sensitivity

and specificity of these thresholds for comparison. Sensitivity was calculated by dividing the number of true positive classifications (i.e., the number of products that was classified as containing multiple drinks by both the threshold and the participants) by the total number of products classified as containing multiple drinks by the participants. Specificity was calculated by dividing the number of true negative classifications (i.e., the number of products that was classified as a single drink by both the threshold and the participants) by the total number of products classified as containing a single drink by the participants.

Transparency and openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the current study. All data and analysis code are available at https://doi.org/10.15131/shef.data.17263604. Materials are available at https://doi.org/10.15131/shef.data.17294252. Data were analyzed using Stata 17 (StataCorp, 2021) and the cutpt module (Clayton, 2013). The study protocol and analysis plan were preregistered at https://osf.io/wyavx/ (Kersbergen, Field, & Meier, 2019).

Amendments to pre-registered study protocol

We made the following deviations from our pre-registered protocol:

1. We did not report results for the multilevel logistic regressions that only used packaging size as a predictor since these were consistent with the findings of the full regression models.
2. We included an interaction variable between hazardous drinking and ABV as a predictor of drink classification in our multilevel logistic regression models to assess how hazardous drinking moderated the effect of ABV on drink classification.
3. We conducted the multilevel logistic regressions analysis with total alcohol content predicting drink classification for each drink type separately (rather than across all

drinks), because total alcohol content was not linearly associated with drink classification and therefore the model across all drinks did not converge.

4. We reported additional thresholds (highest level at which products are considered to contain a single drink on >60% and >80% of observations) because inspection of descriptive plots of the proportion of products that were classified as containing a single drink showed that not all thresholds obtained with the maximized Youden Index met face validity (see Figures S1 to S4 in the supplementary materials). To allow for comparison with the thresholds obtained with the maximized Youden Index, we reported sensitivity and specificity of these additional thresholds.

Results

Product and participant characteristics

One hundred and seven participants consented, 102 of whom completed the study (two completed the study without collecting payment). Eight participants were excluded from analysis; seven because they failed at least two attention checks and one because they indicated that they misunderstood the question and rated products on whether they contained a single drink type or multiple drink types (e.g., gin and tonic, vodka and lemonade). Therefore, our final sample size was 94 (Table 2).

Drink classification (Table 3)

Packaging size

Across all drink types, packaging size significantly predicted the likelihood that a product was classified as containing multiple drinks. Participants were more likely to classify larger sized products as containing multiple drinks. We only found a statistically significant decrease in likelihood in our interaction variable between packaging size and hazardous drinking for spirits and wine, but not for beer, cider and RTDs. Increased risk drinkers were

more likely to classify larger spirits and wine products as containing a single drink than low risk drinkers. These interactions were statistically small and the thresholds for increased risk drinkers matched those for low risk drinkers (see Table S1 in the supplementary materials).

ABV

ABV was a statistically significant predictor of the likelihood that a container was classified as containing multiple drinks for all drink types except spirits, with participants being more likely to classify stronger products as containing multiple drinks. We found no evidence that hazardous drinking moderated the effect of ABV on drink classification.

Total alcohol content

Total alcohol content was statistically significantly associated with likelihood of classification as multiple drinks across all drink types. Participants were more likely to classify products with greater total alcohol content as containing multiple drinks.

Container type

Container type was a statistically significant predictor of the likelihood that a container was classified as containing multiple drinks for all drink types except spirits (which were all bottled). Participants were more likely to classify beer, cider and RTD products as a single drink if they were in a can compared to a bottle. Wine products were more likely to be classified as a single drink if they were in other container types (e.g., pouch) compared to a bottle, but there were no statistically significant differences between cans and bottles.

Thresholds (Table 4)

The AUC for packaging size and total alcohol content was significantly higher than .5 for all drink types (all packaging size AUCs > .62; all total alcohol content AUCs > .68), showing that packaging size and total alcohol content performed better than chance at

indicating whether a drink is considered to contain multiple drinks. The area under the curve for ABV was also significantly higher than .5 for beer (AUC = .56), cider (AUC = .58) and wine (AUC = .53), but not for spirits (AUC = .31) and RTDs (AUC = .30). Based on the maximized Youden Index, the largest packaging size that was considered to still contain a single drink is 568ml for beer, 440ml for cider, 440ml for RTDs, 200ml for spirits, and 375ml for wine. The highest ABV that was considered to still contain a single drink is 4.5% for beer, 6% for cider, 8.8% for RTDs, 50.2% for spirits, and 11.5% for wine. The largest total alcohol content that was considered to still contain a single drink is 2.5 UK units for beer, 3 UK units for cider, 2 UK units for RTDs, 4 UK units for spirits, and 5.5 UK units for wine.

As shown in Table 4, 13 out of 15 thresholds based on the Maximized Youden Index differed from the highest level of packaging size, ABV and total alcohol content at which >60% and >80% of products were classified as containing a single drink. In six instances, the combined sensitivity and specificity of the threshold obtained from single drink classifications on >60% and >80% of the trials were very close to the Maximized Youden Index (within 0.05 out of a possible 2 points difference). Therefore, these alternative thresholds led to similarly accurate classifications, whilst prioritizing either sensitivity or specificity over the Maximized Youden Index. The alternative packaging size threshold for cider was 500ml and had greater specificity. The alternative packaging size threshold for spirits was 50ml and had greater sensitivity. The alternative packaging size threshold for wine was 200ml and had greater sensitivity. The alternative total alcohol content threshold for cider was 3.5 UK units and had greater specificity. The alternative total alcohol content threshold for spirits was 2.5 UK units and had greater sensitivity. The alternative total alcohol content threshold for wine was 2.5 UK units and had greater sensitivity.

Discussion

We investigated to what extent packaging size, ABV, total alcohol content and container type influenced UK drinkers' assessment of whether the drink contained a single drink or multiple drinks. Our results showed that across drink types, larger products, products with higher ABV, products with higher alcohol content, and bottled drinks were more likely to be classified as containing multiple drinks. We report at what levels of packaging size, ABV and total alcohol content products switch from being perceived as containing a single drink to containing multiple drinks. The packaging size thresholds for beer and RTDs were greater and the thresholds for cider, spirits and wine were smaller than the median packaging size of eligible products sold on supermarket websites, suggesting that most beer and RTD products on the market are considered to contain a single drink, and most cider, spirits and wine products are considered to contain multiple drinks. We found no meaningful differences in the thresholds for packaging size, ABV and total alcohol content for low risk and increased risk drinkers, suggesting that these thresholds are likely to be universal among UK drinkers.

Our finding that packaging size predicted perceptions of whether products contained a single drink is in line with research on the effect of portion size on normality perceptions (Haynes et al., 2019). Therefore, reductions to the size of packaging for products that are considered to contain a single drink may lead to greater reductions in alcohol consumption compared to reductions in the size of products that are perceived as containing multiple drinks. However, reductions to the size of wine bottles (which contain multiple drinks) have led to reductions in alcohol consumption (Codling et al, 2020). This suggests that the perception that products contain a single drink is not the sole factor influencing the effect of alcohol size reductions on consumption and changes to products that contain multiple drinks are still worthwhile. Future research should investigate to what extent the perception that products contain a single drink moderates the effect of product size and strength on the amount of alcohol consumed during a drinking occasion.

Our results showed that packaging size was a stronger predictor than ABV of participants' judgements of whether a product contains a single or multiple drinks. In line with the findings by Lovatt et al. (2015), this suggests that drinkers who monitor their consumption by counting drinks rather than standard drinks or units are likely to use the product size to estimate the number of drinks they consumed. Therefore, it may be helpful to present drinking guidelines in terms of drink sizes to help drinkers compare their drinking to the recommendations. Additionally, researchers measuring self-reported alcohol consumption could use drink sizes to prompt more accurate estimates of alcohol consumption.

The key strengths of this study were that we assessed products that were on the UK market at the time of testing and are therefore representative of typical alcohol products, and we used quota sampling to recruit a sample that was representative of the UK population on gender, weekly alcohol consumption and annual household income. This study also had limitations. Firstly, we only tested UK products and UK drinkers. Given that exposure to different portion sizes affects which sizes are considered to be normal (Robinson & Kersbergen, 2018) these results may not generalize to markets where different sized alcohol products are sold. Secondly, we did not account for familiarity with the different drinks, which may have affected product classifications. However, as we showed participants a wide range of products, it is unlikely that they were unfamiliar with all presented drinks. Thirdly, we did not give participants the option to select "less than one drink", meaning that we are only able to assess the points at which single drinks become too large or strong to be seen as single drinks, but not the point at which they are so small or weak that they are seen as less than one drink. Fourthly, given that total alcohol content was strongly correlated with packaging size and ABV, we were unable to test these variables in the same model. Therefore, it is unclear whether packaging size or ABV predict drink classifications above and beyond total alcohol content. Finally, the analysis was restricted by the available products. Therefore, packaging size, ABV,

and total alcohol content were not equally distributed across drink types, causing large gaps between different levels of the variables. For example, there was only a 30ml gap between the smallest packaging sizes of spirits, but a 300ml gap between the largest sizes. Therefore, it is possible that the optimum threshold lies in between the possible sizes that are for sale, which we were unable to detect. This also likely resulted in the observed inconsistencies between the thresholds based on the ROC curve analysis and the indices based on the largest/strongest product that was still considered to contain a single drink on 60% or 80% of observations. The maximized Youden index indicates the threshold with the highest combined specificity and sensitivity and is therefore considered to be the optimal threshold. However, in some cases (the packaging size and the total alcohol content of cider, spirits and wine) the combined specificity and sensitivity of thresholds identified by the other indices were equal to or very close to the maximized Youden Index. In those cases, researchers and policy makers may wish to use the alternative threshold that prioritizes sensitivity or specificity according to their needs (Table 4).

To conclude, this study showed that the size, strength and total alcohol content of alcohol products sold in UK stores affect whether these products are considered to contain a single drink or multiple drinks. We report the different thresholds after which products are too large or too strong to be considered to contain a single drink. Researchers and policy makers could use these thresholds to encourage more accurate self-monitoring of alcohol consumption and inform the development and testing of future interventions related to product size, strength and total alcohol content.

References

Bush, Kristen., Kivlahan, D. R., McDonell, M. B., Fihn, S. D., & Bradley, K. A. (1998). The AUDIT alcohol consumption questions (AUDIT-C). *Archives of Internal Medicine*, *158*, 1789–1795. https://doi.org/10.1097/00000374-199811000-00034

Clayton, P. (2013). *CUTPT: Stata module for empirical estimation of cutpoint for a diagnostic test*. Statistical Software Components S457719, Boston College Department of Economics.

Codling, S., Mantzari, E., Sexton, O., Fuller, G., Pechey, R., Hollands, G. J., Pilling, M., & Marteau, T. M. (2020). Impact of bottle size on in-home consumption of wine: a randomized controlled cross-over trial. *Addiction*, *115*(12), 2280–2292. https://doi.org/10.1111/add.15042

Fluss, R., Faraggi, D., & Reiser, B. (2005). Estimation of the Youden Index and its associated cutoff point. *Biometrical Journal*, *47*(4), 458–472. https://doi.org/10.1002/bimj.200410135

Haynes, A., Hardman, C. A., Makin, A. D. J., Halford, J. C. G., Jebb, S. A., & Robinson, E. (2019). Visual perceptions of portion size normality and intended food consumption: A norm range model. *Food Quality and Preference*, *72*, 77–85.

Kersbergen, I., Field, M., & Meier, P. (2019, February 28). Classification of alcohol packaging sizes into single and multiple servings. https://doi.org/10.17605/OSF.IO/SQ7ZU

Kersbergen, I., Oldham, M., Jones, A., Field, M., Angus, C., & Robinson, E. (2018). Reducing the standard serving size of alcoholic beverages prompts reductions in alcohol consumption. *Addiction*, 1598–1608. https://doi.org/10.1111/add.14228

Khadjesari, Z., White, I. R., McCambridge, J., Marston, L., Wallace, P., Godfrey, C., & Murray, E. (2017). Validation of the AUDIT-C in adults seeking help with their drinking online. *Addiction Science & Clinical Practice*, *12*(1), 2. https://doi.org/10.1186/s13722-016-0066-5

Kontopantelis, E., Springate, D. A., Parisi, R., & Reeves, D. (2016). Simulation-Based Power Calculations for Mixed Effects Modeling: **ipdpower** in *Stata*. *Journal of Statistical Software*, *74*(12). https://doi.org/10.18637/jss.v074.i12

Liu, H., & Wu, T. (2003). Estimating the Area under a Receiver Operating Characteristic (ROC) Curve For Repeated Measures Design. *J Stat Softw*, *8*, 1–18. https://doi.org/10.18637/jss.v008.i12

Lovatt, M., Eadie, D., Meier, P. S., Li, J., Bauld, L., Hastings, G., & Holmes, J. (2015). Lay epidemiology and the interpretation of low-risk drinking guidelines by adults in the United Kingdom. *Addiction*, *110*(12), 1912–1919. https://doi.org/10.1111/add.13072

Office for National Statistics. (2017). *Gross household income by income decile group, UK, financial year ending 2016 - Office for National Statistics*. https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/expenditure/adhocs/006770grosshouseholdincomebyincomedecilegroupukfinancialyearending2016

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872. https://doi.org/10.1016/j.jesp.2009.03.009

Osborne, B., & Cooper, V. (2018). *Health Survey for England 2017: Adult health related behaviours* (Issue December).

Perman-Howe, P. R., Davies, E. L., & Foxcroft, D. R. (2021). The effect of alcohol strength on alcohol consumption: findings from a randomised controlled cross-over pilot trial. *Pilot and Feasibility Studies*, *7*(1), 1–13. https://doi.org/10.1186/s40814-021-00777-4

Robinson, E., & Kersbergen, I. (2018). Portion size and later food intake: evidence on the "normalizing" effect of reducing food portion sizes. *American Journal of Clinical Nutrition*, *107*(4), 640–646.

StataCorp. (2021). *Stata Statistical Software: Release 17*. StataCorp LLC.

Stevely, A. K., Holmes, J., & Meier, P. S. (2021). Combinations of Drinking Occasion Characteristics Associated with Units of Alcohol Consumed among British Adults: An Event-Level Decision Tree Modeling Study. *Alcoholism: Clinical and Experimental Research*, *45*(3), 630–637. https://doi.org/10.1111/acer.14560

Vasiljevic, M., Couturier, D., Frings, D., Moss, A. C., Albery, I. P., & Marteau, T. M. (2018). Impact of Lower Strength Alcohol Labeling on Consumption: A Randomized Controlled Trial. *Health Psychology*. https://doi.org/http://dx.doi.org/10.1037/hea0000622

Table 1. Product characteristics of 2290 unique products for sale in major UK supermarkets and the 500 selected stimuli for the present study.

Stimulus set	Drink type	ABV (%); median (range)	Packaging size of container (ml); median (range)	Number of discrete packaging sizes	Total alcohol content (UK units); median (range)
All eligible products (N = 2290)	Beer (n = 375)	4.8 (1.2 – 9.2)	375 (210 – 750)	24	1.9 (0.3 - 6.0)
	Cider (n = 101)	4.5 (3.5 – 8.2)	500 (250 – 750)	7	2.0 (1.0 - 5.6)
	Ready to drink (n = 109)	5.0 (3.5 – 11.0)	250 (113 – 1000)	9	1.3 (0.4 – 7.7)
	Spirits (n = 648)	40.0 (12.0 – 63.0)	700 (20 – 1000)	10	26.3 (0.7 – 44.1)
	Wine (n = 1057)	12.5 (4.0 – 15.0)	750 (187 – 1500)	7	9.4 (0.8 – 20.3)
Selected products (n = 500)	Beer (n = 100)	4.8 (3.5 – 8.0)	440 (210 – 750)	23	2.0 (0.9 – 6.0)
	Cider (n = 100)	4.5 (3.5 – 8.2)	500 (250 – 750)	5	2.0 (1 – 5.6)
	Ready to drink (n = 100)	5.0 (3.5 – 11.0)	275 (113 – 1000)	9	1.3 (0.4 – 7.7)
	Spirits (n = 100)	40.0 (15.0 – 50.2)	600 (20 – 1000)	7	17.3 (0.7 – 43.1)
	Wine (n = 100)	12.5 (5.5 – 14.5)	750 (187 – 1500)	6	8.8 (1.4 – 18.0)

Note: ABV = Alcohol by Volume.

Table 2. Participant characteristics.

	Total (N = 94)	Male (n = 45)	Female (n = 49)
Age; Mean (SD)	34.69 (11.74)	36.27 (12.17)	33.24 (5.49)
Hazardous drinking (AUDIT-c); Mean (SD)	5.96 (2.59)	6.47 (2.74)	5.49 (2.38)
Low risk drinkers (AUDIT-c < 8); n (%)	67 (71.3%)	28 (62.2%)	39 (79.6%)
Increased risk drinkers (AUDIT-c ≥ 8); n (%)	27 (28.7%)	17 (37.8%)	10 (20.4%)
Annual household income; median	£20,000 – £29,999	£20,000 – £29,999	£30,000 - £39,999

Note: AUDIT-c = Alcohol Use Disorders Identification Test-consumption

Table 3. Likelihood of products being classified as containing multiple drinks based on: packaging size, ABV, container type, hazardous drinking and the interactions between hazardous drinking and ABV and packaging size (Model 1); and total alcohol content (Model 2), split by drink type.

Model	Variable	Beer	Cider	RTD	Spirits	Wine
		OR [99% CI]	OR [99% CI]	OR [99% CI]	OR [99% CI]	OR [99% CI]
1	Constant	78.11 [13.31, 458.49]	8.02 [2.97, 21.67]	30.82 [12.60, 75.40]	276.19 [111.07, 686.78]	28.35 [16.87, 47.65]
	Packaging size (dl)	5.73 [4.57, 7.20]***	5.54 [4.41, 6.96]***	3.62 [3.15, 4.16]***	3.32 [2.95, 3.74]***	3.25 [2.93, 3.59]***
	ABV (%)	2.10 [1.66, 2.66]***	1.37 [1.20, 1.55]***	1.39 [1.27, 1.52]***	1.01 [0.99, 1.03]	1.13 [1.01, 1.26]**
	Container type: can (vs bottle)	0.30 [0.18, 0.49]***	0.46 [0.28, 0.76]***	0.30 [0.18, 0.49]***	-	0.46 [0.12, 1.74]
	Container type: other (vs bottle)	-	-	0.35 [0.08, 1.61]	-	0.00 [0.00, 0.01]***
	Hazardous drinking (AUDIT-c)	0.93 [0.48, 1.84]	0.74 [0.51, 1.07]	0.99 [0.71, 1.40]	0.72 [0.51, 1.02]	0.76 [0.63, 0.93]***
	Hazardous drinking x Packaging size	1.07 [0.99, 1.16]	0.94 [0.86, 1.02]	0.96 [0.92, 1.00]	0.92 [0.88, 0.96]***	0.95 [0.91, 0.98]***
	Hazardous drinking x ABV	1.01 [0.92, 1.10]	0.98 [0.94, 1.03]	1.02 [0.98, 1.05]	0.99 [0.99, 1.00]	0.99 [0.95, 1.03]
	Participant level variance (SE)	6.84 (1.25)	4.97 (0.86)	4.24 (0.75)	4.85 (0.87)	2.43 (0.44)
2	Constant	510.28 [163.94, 1588.32]	34.31 [16.11, 73.05]	727.12 [303.20, 1743.75]	1.51 [0.87, 2.62]	6.63 [4.16, 10.55]
	Total alcohol content (UK units)	12.61 [9.44, 16.84]***	5.69 [4.77, 6.77]***	7.80 [6.44, 9.45]***	1.33 [1.29, 1.37]***	2.38 [2.21, 2.56]***

Participant level variance (SE)	6.84 (1.32)	4.20 (0.73)	3.02 (0.57)	3.99 (0.72)	2.49 (0.44)

Note: * p < .01, ** p < .001, *** p < .0005. ABV = Alcohol by Volume. RTD = ready-to-drink. AUDIT-c = Alcohol Use Disorders Identification Test-consumption.

Table 4. Thresholds for packaging size (ml), ABV (%) and total alcohol content (UK units) after which products are considered to contain multiple drinks based on the Receiver Operating Characteristics curve analysis' Maximized Youden index and classification as a single drink >60% and >80% of observations, split by drink type.

Product characteristic	Drink type	AUC	Threshold Index								
			Maximized Youden index			Single drink on > 60% of observations			Single drink on > 80% of observations		
			Threshold	Sensitivity	Specificity	Threshold	Sensitivity	Specificity	Threshold	Sensitivity	Specificity
Packaging size (ml)	Beer	.78	**568**	.64	.87	660	.21	.99	**568**	.64	.87
	Cider	.62	**440**	.86	.51	500*	.34	.99	500*	.34	.99
	RTD	.79	**440**	.74	.93	**440**	.74	.93	**440**	.74	.93
	Spirits	.84	**200**	.89	.85	**50**	.91	.83	20	> .99	.04
	Wine	.81	**375**	.96	.78	200*	.97	.74	200*	.97	.74
ABV (%)	Beer	.56	**4.5**	.80	.36	7.5	.07	> .99	5.9	.17	.98
	Cider	.58	**6**	.49	.89	8.2	0	1	**6**	.49	.89
	RTD[1]	.30	**8.8**	.15	.98	**8.8**	.15	.98	**8.8**	.15	.98
	Spirits[1]	.31	**50.2**	0	1	**50.2**	0	1	N/A	N/A	N/A
	Wine	.53	**11.5**	.74	.47	7	.96	.05	7	.96	.05
Total alcohol content	Beer	.74	**2.5**	.68	.84	4.5	.14	> .99	**2.5**	.68	.84
(UK units)	Cider	.68	**3**	.49	.90	4	.34	.99	3.5*	.40	.97
	RTD	.78	**2**	.74	.91	3.5	.20	.99	3.5	.20	.99
	Spirits	.86	**4**	.90	.84	**2.5**	.91	.83	1.5	.99	.15
	Wine	.85	**5.5**	.92	.79	**2.5**	.97	.74	**2.5**	.97	.74

Note: AUC = Area under the curve. ABV = Alcohol by Volume. RTD = ready-to-drink. Low sensitivity is indicative of false negatives (incorrectly classified single drinks as multiple drinks). Low specificity is indicative of false positives (incorrectly classified multiple drinks as a single drink). Values in bold have the highest sum of sensitivity and specificity. * indicates thresholds whose sum of sensitivity and specificity is within .05 of the highest value. [1] The AUC indicated that ABV did not perform above chance level at classifying RTDs and spirits as containing a single or multiple drinks.

Figure 1. Example mock-up of a product shown to participants in the drink classification task.