# Shooting shots: Estimating alcoholic drink sizes in real life using event-level reports and annotations of close-up pictures

Short running title: Alcoholic drink sizes in real life

*Florian Labhart[1,2,3], Thanh-Trung Phan[1,4],*

*Daniel Gatica-Perez[1,4], & Emmanuel Kuntsche[2]*

[1]Idiap Research Institute, Social Computing Group, Martigny, Switzerland;
[2]Centre for Alcohol Policy Research, La Trobe University, Australia;
[3]Addiction Switzerland, Research Department, Lausanne, Switzerland;
[4]Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Correspondence should be addressed to: Florian Labhart, Idiap Research Institute, Rue Marconi 19, CH-1920 Martigny, Switzerland; Phone: +41277217731; Email: florian.labhart@idiap.ch

The authors do not have any conflicts of interest. The study was funded by the Swiss National Science Foundation (grant no. CRSII5_173696) and approved by the Lausanne and Zurich cantonal ethics committees for research on human beings (protocol 145/14).

**Abstract**

Introduction and Aims: Drinks consumed in real life are diverse, in terms of beverage type, container size and alcohol by volume (ABV). To date, most EMA studies have assessed drinking amounts with 'standard' drinks, although their event-level design allows for more advanced assessment schemes. The purpose of this empirical EMA study is to compare participants' estimates of alcoholic drink characteristics, assessed using drink-specific questions, with estimates generated by annotators based on pictures of the same drinks.

Design and Methods: On weekend nights, 186 young adults took 1,484 close-up pictures of their drinks using a custom-built smartphone application. Participants reported the beverage type, drink size, and ABV. Annotators described the beverage type, container size, and filling level. Correspondence between participants' and annotators' estimates were explored using descriptive statistics, difference tests and correlations.

Results: Annotators were unable to precisely identify the beverage types in most pictures of liqueurs, spirits and mixed drinks. Participants' drink size estimates converged with annotators' estimates of the container size for beer (41cl. corresponding to 16 grams of pure alcohol) and mixed drinks (28cl./35g), and of the content size for wine (10cl./9g). However, annotators estimated much larger sizes for liqueur/fortified wine (12cl./14g vs 7cl./9g) and spirits (8cl./26g vs 4cl./10g) than participants.

Discussion and Conclusions: Given the lack of visual cues on the beverage types, annotations of pictures should be considered as a complement to participants' reports rather than a substitute. Except for wine, drinks in real life vary largely and often exceed a 10-gram 'standard' drink.



**Keywords:** Drink size, Standard drink, Ecological momentary assessment, Smartphone application, Picture annotation

## Introduction

In alcohol research and policy, the amounts of alcohol consumed are principally measured in the number of 'standard drinks' that people consume within a given time frame [1,2]. The 'standard drink' concept builds on the premise that the amount of pure alcohol contained in any one drink is roughly comparable across beverage types, and that individuals can take account of standard drinks when recalling their own consumption. Standard drinks are therefore widely used to assess peoples' drinking behaviour in research [2] and clinical practices [3,4], as well as to communicate guidelines to encourage people to drink responsibly [5,6].

Yet, while the concept of 'standard drink' facilitates assessments and communications about alcohol use, its size is less standardised than what its name suggests. For instance, the amount of pure alcohol per standard drink varies across countries (e.g. United Kingdom: 8g; Australia, Switzerland: 10g; United States: 12–14g) [7–9]. Additionally, no single measurement can represent the diversity of drink sizes served or poured in real life [10], given that people often self-pour their drinks, bottles and cans are sold in varied sizes, and alcohol by volume (ABV) varies considerably within beverage types (i.e., 1-11% for beer, 7-14% for wine and 35-50% for spirits) [11]. Finally, free-pour experiments measuring the size of self-poured drinks demonstrated that drinkers tend to overestimate the size of standard drinks [see 12, for a review]. Over-pouring seems particularly frequent for wine, spirits and liqueur, in large and wide glasses, and can account for up to two standard drinks per glass [13]. The magnitude of over-pouring varies across beverage types, container sizes and individual habits [14], but is robust across countries [15].

Given their ability to record drinking behaviours in real time, ecological momentary assessment (EMA) studies [16,17] are particularly appropriate to capture the characteristics of real-life drinks. While most EMA studies assessed drinking amounts in number of standard drinks [18–21], several studies have developed alternative methods for capturing detailed characteristics of the drinks. A first method consists of separately assessing the beverage type (e.g., ‘beer’), size (e.g. ‘50 cl.’, ‘large beer’ or ‘pint’), and ABV [22,23]. This procedure allows for multiple combinations of size and ABV per beverage type. Evaluations of different smartphone apps using this assessment format have shown limited response burden [22,24]. A second approach collects close-up pictures of drinking and extracts their characteristics via manual annotations or computer vision algorithms [25]. This procedure has the advantage of limiting participants’ subjectivity by delegating the assessment of the drink characteristics to external coders. Although these methods cannot precisely measure the drink size in grams of pure alcohol, as done in experimental studies [26–28], they provide systematic estimates allowing for comparisons of the drink characteristics within and across beverage types.

As part of a larger EMA study on drinking and nightlife behaviours [29–31], study participants were requested to document their drinks using both a questionnaire on the beverage type, size, and ABV, and a close-up picture [24]. Visible drink characteristics on the pictures were then annotated by five independent research assistants (“annotators”). This study aims to compare the drink characteristics generated by participants and annotators for different beverage types, and to estimate the corresponding drink size. Alongside the development of sensors to unobtrusively collect data on people’s real-life behaviours [30,32], this study stands as an exploration of the feasibility of reducing participant burden in EMA studies by collecting pictures of drinks in place of questionnaires.

The first part of this study will assess the ability of annotators to identify the *beverage types* as reported by the participants. Because beer and wine have typical visual characteristics (e.g., beer colour, red wine colour, beer bottle shape, wine-glass shape), we hypothesise that annotators will accurately identify those two beverage types. However, we expect a low identification rate for other alcoholic beverage types (liqueur/fortified wine, spirits, cocktails) given the lack of typical visual characteristics (e.g., vodka might look like water and cocktails like fruit juices).

Secondly, we will compare estimates of the *drink sizes* in centilitres from participants and annotators, separately for the most common beverage and container types (e.g. for beer: aluminium can, bottle, glass and plastic cup). Because both assessment methods used beverage-specific questions accounting for variations in drink sizes, we hypothesise that both methods will provide similar estimates.

Thirdly, we will estimate, in an exploratory way, the corresponding *alcohol content* in grams of pure alcohol per beverage and container types. Because, previous evidence has shown that over-pouring was likely for all beverage types [12], we expect that average drink sizes will exceed 10grams of pure alcohol for all or almost all beverage types.

## Materials and Methods

### Study design

Participants were recruited from two major nightlife hubs in Switzerland [33]. Eligible participants were aged between 16 to 25, owned an Android smartphone, had been out in the city at least twice in the past month, and had consumed alcohol at least once in the past month. After providing consent, participants installed a custom-developed app (Youth@Night) on their smartphone to document 10 weekend (Friday and Saturday) nights over seven consecutive weekends [24]. The app recorded various aspects of the participants' drinking and nightlife behaviours using questionnaires, pictures, videos and sensors [24,32]. Specifically, participants were requested to take a picture of each drink consumed before the first sip and to describe it with three drink-specific questions (see Measures below). Guidelines (i.e., portrait mode, drink in the centre, no artistic filter, etc.) were shown before taking each picture to help participants produce a high-quality representation of the drink. The study was approved by the Lausanne and Zurich Cantonal Ethics Committees for Research on Human Beings (protocol 145/14).

## Sample

In total, the 241 participants (mean age=19.1, SD=2.4; 46.5% women) documented 2,540 drinks on 1,291 participant-nights. Of those, 130 pictures were unusable due to missing questionnaires or corrupted files. From the 2,410 annotated pictures, 143 were excluded for the following reasons: 49 showed no drink (34.3%), 64 were too dark (44.8%), 39 were blurry (27.3%) and 64 showed several drinks but none in the foreground (44.8%). Altogether, 1,501 pictures represented alcoholic drinks and 766 non-alcoholic drinks. Because the analysis aimed to estimate single drink characteristics, we excluded 17 pictures of alcoholic drinks (1.1%) that participants or annotators described as containing more than 50 centiliters, or for which annotators could not estimate the size. The final set comprised 1,484 close-up pictures of alcoholic drinks reported by 186 participants.

## Measures

### *Participants*

After taking the picture of the drinks, the app provided participants with a predefined list of six alcoholic (‘Beer/cider’, ‘Wine/champagne’, ‘Liqueur/fortified wine’, ‘Shots’, ‘Straight spirit’ and ‘Cocktail/alcopops’) and six non-alcoholic beverages [24]. Subsequently, depending on the beverage type, lists of the most common sizes and alcohol by volume (ABV) were provided (Table 1). The midpoint of ABV ranges was taken (e.g., 5-10%=7.5%).

--Table 1--

### *Annotators*

Five research assistants inspected the entire set of pictures in a random order and annotated features of the each drink appearing in the foreground. All were familiar with the beverage types and brands consumed by young adults in Switzerland. A random selection of 5% of the pictures were annotated twice to measure intra-annotator consistency.

We used the same 12 *beverage type* categories as in the app (see above) plus ‘impossible to say’. Annotators were instructed to pick a response only if they were almost certain of the answer, or otherwise tick ‘impossible to say’.

To estimate the *drink size*, annotators first described the drink *container size* and then its *filling level*. Response options for the container size included an indication of the volume in centilitres and corresponding typical containers: ‘5cl (e.g., shot)’, ‘10cl (e.g., white wine glass)’, ‘15cl (e.g., small plastic cup)’, ‘20cl (e.g., small table glass)’, ‘25cl (e.g., small beer bottle/can, small wine glass)’, ‘33cl (e.g., medium beer bottle/can, standard table glass)’, ‘40cl (e.g., large wine glass)’, ‘50cl (e.g., large beer bottle/can)’, and ‘impossible to say’. Response options for the filling level included 10-percent increments from 0% to 100%, plus ‘impossible to say’.

The *container material* was assessed with the following response options: ‘Glass’, ‘Ceramic’, ‘Plastic’, ‘Aluminium’, ‘Paper’, and ‘Other’. Response options for the *container shape* were: ‘Long/tall glass’, ‘Medium/short glass’, ‘Small/shot glass’, ‘Cup’, ‘Mug’, ‘Wine glass’, ‘Bottle’, ‘Can’, and ‘Other’. For the analyses, the two categories (container shape and material) were combined to identify the most prevalent *container type* combinations per beverage type.

Intra- and inter-annotator agreements were excellent for all measures (Table 2) [34]. Annotators’ answers were aggreated as follows: (a) if three or more annotators agreed on one option, this value was selected (96.9% of the cases); (b) ‘impossible to say’ answers were excluded and, if two annotators agreeded on one option while the others ticked different options, this value was selected (1.6%); (c) for the remaining pictures, the mean was taken for continuous variables and, for categorial variables, pictures were manually checked by the research team (1.5%).

--Table 2--

## Analytical strategy

Descriptive and bivariate statistics were used to compare the drinks size (in cl.) estimated by the participants and annotators. Differences between sources were measured with *t*-tests and Pearson correlations, and variations between container types with one-way Analysis of Variance (ANOVA). Effect sizes were adjusted for the effect of drink reports being nested within participants using Stata 14 [35]. It should be noted that, in contrast to participants who only estimated the content size, annotators estimated separately the container size and its filling level. This procedure was used to account for the fact that, unlike beer glasses that are usually poured up to the top, wine and spirits are usually poured only partially. Participants' estimates were therefore compared to annotators' estimates of both the container size and the content size (=container size*filling level).

Participants' and annotators' drinks size estimates (in cl.) were converted into grams of pure alcohol using the formula: drink size (in ml.)*ABV (in %)*0.793g/ml. (specific alcohol density at 20°C). Because ABV was not measured by annotators, the ABV reported by the participants was used for both sources in an exploratory way. Differences between sources were measured with *t*-tests and Pearson correlations, also adjusted for the nested structure of the data.

# Results

## Identification of beverage types

Annotators were able to correctly identify a large majority of beer and wine pictures (Table 3). They identified most pictures of liqueur/fortified wine as containing alcohol, but often attributed them to other beverages categories (spirits and mixed drinks). Annotators were unable to provide a classification for most pictures of spirits or mixed drinks and ticked the option ‘Impossible to say’.

--Table 3--

**Drink sizes (in centilitres)**

Participants estimated the highest average drink size for beer/cider (40.6 cl.), followed by mixed drinks, wine/champagne, liqueur/fortified wine, and straight spirits, in that order (Table 4). Annotators’ container size estimates were similar to participants’ drink size estimates for beer/cider and cocktails, i.e., drinks that usually fill the entire glass, as shown by non-significant *t*-test and fair correlation coefficients. Regarding drinks that usually fill half or less of the glass, like wine, fortified wine and spirits, annotators’ estimates of the content size (i.e., container size multiplied by filling level) closely aligned with participants’ estimates only for wine. For straight spirits and liqueur/fortified wine, significant differences were found between participants and annotators for both the container and content size estimates.

Significant variations across container types were reported by participants and annotators for beer/cider and cocktails. In particular, beer cans and cocktails served in long plastic cups were estimated as larger than in other containers. Interestingly, annotators systematically estimated the content of wine glasses to be approximately 10cl. regardless of variations in container size.

--Table 4--

**Alcohol content (in grams)**

Participants reported the highest ABV for spirits (37.1%), followed by liqueur/fortified wine, mixed drinks, wine and beer/cider, in that order (Table 5). Building on the results in Table 4, drink size (in grams) for annotators were estimated using the container size for beer/cider and cocktails and the content size for wine, liqueur and spirits. Similar drink size estimates were found between participants and annotators for beer and wine. Beers were equivalent to 1.6 standard drinks on average and more than 80% of those exceeded 10grams. Beer cans contained almost two standard drinks each and were larger than the other beer containers. Conversely, wine glasses were slightly smaller than one standard drink on average, without variation across container types, and almost every second wine glass was under 10grams. While participants estimated liqueur and spirits glasses to contain about one standard drink on average, annotators' estimates were 1.5 times larger for liqueurs and 2.5 times larger for spirits. Finally, both sources estimated that mixed drinks/cocktails contained at least three standard drinks on average, with little variation across container type.

--Table 5--

## Discussion

This study aimed to compare the characteristics and sizes of alcoholic drinks in real life generated by participants (documenting their drinks through drink-specific questions in a smartphone app) and by research assistants (annotating close-up pictures of the same drinks). These two assessment methods were designed to collect ecologically valid information on the characteristics of real-life drinks and to reflect the diversity of drink containers and sizes within and across several alcoholic beverage types. In addition, this study explored the possibility of reducing burden in EMA studies by requesting participants only take a picture of their drink and letting annotators extract information on the characteristics.

The study focused first on the ability of annotators to assess the drink characteristics. Results showed that this is challenging because several characteristics were not distinctly shown on the pictures. While annotators correctly identified beer and wine glasses in most cases, most likely the results of the typical drink container shapes and liquid colours of those beverage types, such visual cues were apparently missing for the other drinks. Findings showing that annotators often confused liqueur/fortified wine with spirits and chose the ‘impossible to say’ response option for most spirits and mixed drinks confirmed our first hypothesis that annotators could only correctly identify beer and wine pictures. Consequently, participants’ reports appear more insightful to determine the beverage category.

ABV cannot generally be determined in a picture either. In some cases, this information might be recognizable from can or bottle brands but remains unavailable for a glass of beer, wine, spirits, or mixed beverages. In fact, ABV might also be unknown to the participants, for example when the drink is prepared or poured by another person. To partly address this issue, future research might collect pictures of the ABV indication on the bottle alongside the glass, although this will likely increase participant burden and might be hardly feasible in restaurants or pubs.

Altogether, the findings show that the provision of a close-up picture, as done in this study, cannot replace participants’ reports. However, the study proves the utility of such pictures to augment participants’ self-reports with information on the containers’ size, percentage of filling, shape, and material. Although not used in this study, annotations might provide additional information on the drinking occasion, such as the type of location, the social context or on-going activities [25]. Regardless, given that participants reported that taking pictures of drinks can disrupt the on-going social dynamics [24,31], future studies should provide an option for participants to take pictures to supplement their self-reports whenever they can, rather than force it for each drink.

The second part of this study compared the drink size estimates, in centilitres, between participants and annotators. While we hypothesised that both sources would estimate similar amounts, their estimates corresponded, after accounting for partial filling, only for wine. Regarding beer and mixed drinks, participants’ and annotators’ estimates of the container size, rather than the content size, matched. This result highlights a flaw in the annotation tool, namely that we forgot to consider that drink containers are never filled more than 80 or 90%. Beer glasses are typically larger than the expected amount to leave room for the foam. This explains why the annotators' content size estimates were underestimated for beer and mixed drinks. Additionally, participants' estimates were smaller than annotators' estimates for liqueur/fortified wine and spirits, even when accounting for the partial filling level. While we cannot determine which source is closer the reality, these results underline the difficulty in assessing the size of drinks with small volume but high alcohol content.

Finally we explored the corresponding alcohol content, in grams of pure alcohol, per beverage type. This analysis was done with the exploratory purpose of providing more sensible figures of the alcohol content, using the ABV estimated by participants. Results partly confirmed our third hypothesis by showing that the estimated average drink size was larger than 10grams for almost all beverages types, but wine. Unlike evidence showing that over-pouring wine is the norm in Australia [36], the UK [13,28] or the US [37], results suggest that people in Switzerland pour on average one standard drink independently of the container size. Consistent with evidence from Spain [38], a likely explanation is that pouring small amounts per wine glass is a cultural habit of wine-producing countries. The findings that beers generally exceeded one standard drink is not surprising given that most standard bottles and cans exceed 25cl. The fact that data were collected on weekend nights, where one large beer is easier to carry (e.g., pre-drink in parks [39]) and cheaper than two smaller beers, might also play a role in this result. Regarding spirits and liqueurs, while participants' estimates corresponded to about one standard drink, annotators estimated those 1.5 times larger for liqueur/fortified wine and 2.5 times larger for spirits. A possible explanation is that the presence of ice cubes or thick glass bottoms biased annotators' estimations. However, given the vast literature documenting people's tendency to over-pour spirits and liqueurs [12], it is very likely that participants underestimated the actual size of these two beverage types. Finally, both participants and annotators reported that self-mixed drinks and cocktails were on average equivalent to three standard drinks. The fact that many cocktail recipes contain at least two types of spirits may explain this result.

**Limitations and future directions**

A number of limitations should be acknowledged. First, the exact drink size and alcohol content have not been experimentally measured. Results are estimates from six independent people (1 participant and 5 annotators) but none can be considered as ground truth. Future studies should include validation pictures, prepared and measured by the research team, in the annotation set to measure annotators' accuracy. Training sessions might also be organised to increase their accuracy. Second, the questions used to record the drink characteristics from participants and annotators were not strictly identical, since annotators estimated the container size and the filling level separately. This might have decreased the comparability of the two methods and increased the risk of errors in the annotators' estimations. The fact that annotators and participants provided similar estimates of the size of wine glasses highlights the conditionally correspondence of both methods, which nevertheless needs further validation. Third, some participants explained that the first sip is often part of the social ritual of saying 'cheers' and that they sometimes took the picture after one or more sips [24]. Consequently, the filling level estimated by annotators might be smaller than the initial beverage poured. Future research should include the possibility for participants to tag pictures that were not taken before the first sip. Fourth, the study did not account for potential influences of the context (e.g., number of drinks already consumed, type of setting) or of the participants (e.g., gender, age) on their drink reports. Although annotators' estimations were not affected by such influences, future studies might investigate whether participants' reports vary within and across drinking occasions. Finally, data were collected among young adults on weekend nights. The findings might not be applicable to other age groups, weekdays or to other times of the day.

## Conclusion

This study used in-situ reports and annotations of close-up pictures of 1,484 alcoholic drinks provided by young adults on weekend nights to estimate the characteristics and size of drinks in real-life situations. The results concur with previous evidence [12] that (1) drink sizes in real life vary largely and consumed amounts cannot be precisely assessed by counting ‘standard’ drinks, and (2) that single drinks often exceed 10grams of pure alcohol. These results have implications for alcohol research and prevention. Regarding research, the drink-specific questions (beverage type, drink size and ABV) completed by the participants provided detailed data on the drink characteristics and are recommended for any future EMA study on alcohol use. Annotators, however, had difficulties identifying beverage types but provided detailed data on other characteristics, such as the container shape and material. Annotations of pictures should be considered to complement participants’ in-situ reports rather than as a substitute. Regarding prevention and health communication, people are likely to drink excessively when considering their usual drinks as being ‘standard’, despite adhering to safe drinking guidelines. It is therefore crucial to recommend people count their drinks in terms of ‘small’ drink sizes, rather than ‘standard’, because some usual containers can contain up to three units per drink.

## Acknowledgements

The study was funded by the Swiss National Science Foundation though the Dusk2Dawn project (grant no. CRSII5_173696) We would like to thank all participants for the provision of the pictures and Lucille Chapuis, Laura Délèze, Marie-Noëlle McGarrity, Eva Schneider and Noémie Vuilleumier for the patient completion of the annotation task.

## References


1. Devos-Comby L, Lange J. “My Drink is Larger than Yours”? A Literature Review of Self-Defined Drink Sizes and Standard Drinks. Curr Drug Abuse Rev. 2008 Jun 1;1(2):162–76.
2. Gmel G, Rehm J. Measuring Alcohol Consumption. Contemp Drug Probl. 2004 Sep;31(3):467–540.
3. Babor TF, de la Fuente JR, Saunders J, Grant M. The Alcohol Use Disorders Identification Test: Guidelines for use in Primary Care. Geneva: World Health Organization; 2001.
4. Mayfield D, McLeod G, Hall P. The CAGE questionnaire: validation of a new alcoholism screening instrument. Am J Psychiatry. 1974;131(10):1121–3.
5. Stockwell T, Butt P, Beirness D, Gliksman L, Paradis C. The basis for Canada’s new low-risk drinking guidelines: A relative risk approach to estimating hazardous levels and patterns of alcohol use: A relative risk approach to drinking guidelines. Drug Alcohol Rev. 2012 Mar;31(2):126–34.
6. Rehm J, Patra J. Different guidelines for different countries? On the scientific basis of low-risk drinking guidelines and their implications: Scientific basis of low-risk drinking guidelines. Drug Alcohol Rev. 2012 Mar;31(2):156–61.
7. Miller WR, Heather N, Hall W. Calculating standard drink units: international comparisons. Addiction. 1991 Jan;86(1):43–7.
8. World Health Organization. International guide for monitoring alcohol consumption and related harm. 2000 [cited 2018 Mar 20]; Available from: http://whqlibdoc.who.int/hq/2000/WHO_MSD_MSB_00.4.pdf
9. Turner C. How much alcohol is in a ‘standard drink’? An analysis of 125 studies. Addiction. 1990 Sep;85(9):1171–5.

10. Carruthers SJ, Binns CW. The standard drink and alcohol consumption. Drug Alcohol Rev. 1992 Oct;11(4):363–70.
11. Kerr WC, Stockwell T. Understanding standard drinks and drinking guidelines: Understanding standard drinks and drinking guidelines. Drug Alcohol Rev. 2012 Mar;31(2):200–5.
12. Schultz NR, Kohn CS, Schmerbauch M, Correia CJ. A systematic review of the free-pour assessment: Implications for research, assessment and intervention. Exp Clin Psychopharmacol. 2017;25(3):125–40.
13. Gill J, O'May F. Practical demonstration of personal daily consumption limits: a useful intervention tool to promote responsible drinking among UK adults? Alcohol Alcohol. 2007 Feb 16;42(5):436–41.
14. White AM, Kraus CL, Flom JD, Kestenbaum LA, Mitchell JR, Shah K, et al. College Students Lack Knowledge of Standard Drink Volumes: Implications for Definitions of Risky Drinking Based on Survey Data: Alcohol Clin Exp Res. 2005 Apr;29(4):631–8.
15. Zandy SL, Pang JS, Ho M-HR, Matthews DB. Singaporean College Students Overpour Drinks Similar to Western Populations: Influence of Peer Presence in a Simulated Alcohol-Pouring Task. Alcohol Clin Exp Res. 2013 Nov;37(11):1963–70.
16. Shiffman S. Ecological momentary assessment (EMA) in studies of substance use. Psychol Assess. 2009 Dec;21(4):486–97.
17. Bolger N, Davis A, Rafaeli E. Diary Methods: Capturing Life as it is Lived. Annu Rev Psychol. 2003 Feb;54(1):579–616.
18. Merrill JE, Wardell JD, Read JP. Drinking Motives in the Prospective Prediction of Unique Alcohol-Related Consequences in College Students. J Stud Alcohol Drugs. 2014 Jan;75(1):93–102.

19. O'Donnell R, Richardson B, Fuller-Tyszkiewicz M, Liknaitzky P, Arulkadacham L, Dvorak R, et al. Ecological momentary assessment of drinking in young adults: An investigation into social context, affect and motives. Addict Behav. 2019 Nov;98:106019.

20. Dulin PL, Alvarado CE, Fitterling JM, Gonzalez VM. Comparisons of alcohol consumption by timeline follow back vs.smartphone-based daily interviews. Addict Res Theory. 2017 May 4;25(3):195–200.

21. Labhart F, Kuntsche E. When yesterday's consumption strikes back: Deviation from usual consumption inversely predicts amounts consumed the next weekend evening: Prediction of next day's consumption. Drug Alcohol Rev. 2014 Jul;33(4):385–92.

22. Poulton A, Pan J, Bruns Jr LR, Sinnott RO, Hester R. A Smartphone App to Assess Alcohol Consumption Behavior: Development, Compliance, and Reactivity. JMIR MHealth UHealth. 2019 Mar 25;7(3):e11157.

23. Monk RL, Heim D, Qureshi A, Price A. "I Have No Clue What I Drunk Last Night" Using Smartphone Technology to Compare In-Vivo and Retrospective Self-Reports of Alcohol Consumption. Ryabinin AE, editor. PLOS ONE. 2015 May 19;10(5):e0126209.

24. Labhart F, Tarsetti F, Bornet O, Santani D, Truong J, Landolt S, et al. Capturing drinking and nightlife behaviours and their social and physical context with a smartphone application – investigation of users' experience and reactivity. Addict Res Theory. 2020;28(1):62–75.

25. Phan T-T, Labhart F, Gatica-Perez D. My Own Private Nightlife: Understanding Youth PersonalSpaces from Crowdsourced Video. Proc ACM Hum-Comput Interact. 2019;

26. Kuendig H, Kuntsche E. Solitary Versus Social Drinking: An Experimental Study on Effects of Social Exposures on In Situ Alcohol Consumption. Alcohol Clin Exp Res. 2012 Apr;36(4):732–8.

27. Barnett NP, Wei J, Czachowski C. Measured alcohol content in college party mixed drinks. Psychol Addict Behav. 2009;23(1):152–6.

28. Boniface S, Kneale J, Shelton N. Actual and Perceived Units of Alcohol in a Self-Defined “Usual Glass” of Alcoholic Drinks in England. Alcohol Clin Exp Res. 2013 Jun;37(6):978–83.

29. Labhart F, Anderson KG, Kuntsche E. The Spirit Is Willing, But the Flesh is Weak: Why Young People Drink More Than Intended on Weekend Nights-An Event-Level Study. Alcohol Clin Exp Res. 2017 Nov;41(11):1961–9.

30. Santani D, Biel J-I, Labhart F, Truong J, Landolt S, Kuntsche E, et al. The night is young: urban crowdsourcing of nightlife patterns. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing [Internet]. ACM Press; 2016 [cited 2018 Aug 6]. p. 427–38. Available from: http://dl.acm.org/citation.cfm?doid=2971648.2971713

31. Truong J, Labhart F, Santani D, Gatica-Perez D, Kuntsche E, Landolt S. The emotional entanglements of smartphones “in the field”: on social positioning, power relations, and research ethics. Area. 2019;52(1):81–8.

32. Santani D, Do TMT, Labhart F, Landolt S, Kuntsche E, Gatica-Perez D. DrinkSense: Characterizing Youth Drinking Behavior using Smartphones. IEEE Trans Mob Comput. 2018;17(10):1536–1233.

33. Labhart F, Santani D, Truong J, Tarsetti F, Bornet O, Landolt S, et al. Development of the Geographical Proportional-to-size Street-Intercept Sampling (GPSIS) method for recruiting urban nightlife-goers in an entire city. Int J Soc Res Methodol. 2017 Nov 2;20(6):721–36.

34. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess. 1994;6(4):284.

35. StataCorp. Stata Statistical Software: Release 14. Coll Stn TX StataCorp LP. 2015;

36. Banwell C. How many standard drinks are there in a glass of wine? Drug Alcohol Rev. 1999 Mar 1;18(1):99–101.

37. Kaskutas LA, Graves K. Pre-pregnancy drinking: how drink size affects risk assessment. Addiction. 2001 Aug;96(8):1199–209.

38. Gual A, Martos AR, Lligona A, Llopis JJ. Does the concept of a standard drink apply to viticultural societies? Alcohol Alcohol. 1999 Mar 1;34(2):153–60.

39. Labhart F, Kuntsche E. Development and validation of the predrinking motives questionnaire: Predrinking motives questionnaire. J Appl Soc Psychol. 2017 Mar;47(3):136–47.

Table 1: Drink size and alcohol content categories presented in the drink-specific drop-down lists in the Youth@Night app

| | Drink size | Alcohol by volume (ABV) |
|---|---|---|
| Beer, cider | ‘Small (25 cl.)’, ‘Medium (33 cl.)’, ‘Large’ (50 cl.) | ‘Light (2-4%)’, ‘Medium (4-6%)’, ‘Strong (6% or more)’ |
| Wine, champagne | ‘Small (5 cl.)’, ‘Medium (10 cl.)’, ‘Large (15 cl.)’ | ‘Light (8-10%)’, ‘Medium (10-12%)’, ‘Strong (12-14%)’ |
| Liqueur, fortified wine | ‘Small (4 cl.)’, ‘Medium (8 cl.)’, ‘Large (12 cl.)’ | ‘Light (10-15%)’, ‘Medium (15-20%)’, ‘Strong (20% or more)’ |
| Straight spirits | ‘Small (2 cl.)’, ‘Medium (4 cl.)’, ‘Large (6 cl.)’ | ‘Light (25-35%)’, ‘Medium (35-45%)’, ‘Strong (45% or more)’ |
| Alcoholic mixed drinks (e.g., cocktail, alcopops, premix, self-mix) | ‘5 cl.’, ‘10 cl.’, ‘20 cl.’, ‘30 cl.’, ‘50 cl.’, ‘1 litre’, ‘1.5 litre’ | ‘0-5%’, ‘5-10%’, ‘10-15%’, ‘15-20%’, ‘20-25%’, ‘25-30%’, ‘30% or more’ |

Table 2: Intra- and inter-annotator levels of agreement

| | Double coding within annotators | | Agreement across annotators |
|---|---|---|---|
| | All annotators (combined measure) | Per annotator (range) | |
| | Kappa / Correlations[1] | Kappa / Correlations[1] | ICC(2,k) |
| Beverage type | 0.893 | 0.797-0.941 | 0.952 |
| Drink container size | 0.913 | 0.897-0.938 | 0.928 |
| Percentage of filling | 0.951 | 0.906-0.975 | 0.950 |
| Drink container material | 0.943 | 0.895-0.987 | 0.993 |
| Drink container shape | 0.906 | 0.841-0.961 | 0.981 |

Note: 1) Kappa used for categorical variables (beverage type, drink material and shape); Pearson correlation used for ordinal variables (drink size and percentage of filling)

Table 3: Comparison of annotators' beverage types classification to participants' report

| Participants | | Annotators | | | |
|---|---|---|---|---|---|
| | | Same drink | Another alcoholic drink | Non-alcoholic drink | Impossible to say |
| | Total | N (%) | N (%) | N (%) | N (%) |
| Beer, cider | 759 | 737 (97.1%) | 2 (0.3%) | 3 (0.4%) | 17 (2.2%) |
| Wine, champagne | 172 | 119 (69.2%) | 5 (2.9%) | 6 (3.5%) | 42 (24.4%) |
| Liqueur, fortified wine | 45 | 2 (4.4%) | 21 (46.7%) | 6 (13.3%) | 16 (35.6%) |
| Straight spirits | 84 | 19 (22.6%) | 8 (9.5%) | 9 (10.7%) | 48 (57.1%) |
| Mixed drinks, cocktails | 314 | 104 (33.1%) | 23 (7.3%) | 25 (8.0%) | 162 (51.6%) |

Table 4: Estimations of drink sizes by participants and annotators, per beverage and container type

| | | Participants | Annotators | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Drink size | Container size | | | Filling level | Content size | | |
| | | | | Difference with participants' estimates | | | | Difference with participants' estimates | |
| | | cl. | cl. | Diff. test | Correlation | % | cl. | Diff. test | Correlation |
| | N | Mean (SD) | Mean (SD) | *t* [*p*] | *r* [*p*] | Mean (SD) | Mean (SD) | *t* [*p*] | *r* [*p*] |
| Beer, cider | 759 | 40.6 (9.8) | 40.8 (8.8) | -0.45 [.657] | 0.65 [<.001] | 64.1 (29.5) | 24.0 (12.6) | 11.02 [<.001] | 0.19 [.008] |
| Aluminium can | 248 | 48.9 (4.5) | 49.7 (2.1) | -2.69 [.009] | 0.62 [.013] | 88.8 (31.6) | 44.4 (15.8) | 0.80 [.443] | -0.09 [.371] |
| Bottle | 245 | 33.7 (6.7) | 34.7 (4.9) | -1.93 [.057] | 0.53 [<.001] | 66.8 (28.7) | 22.6 (9.9) | 8.94 [<.001] | 0.13 [.188] |
| Glass | 176 | 39.4 (10.6) | 41.0 (9.2) | -1.29 [.203] | 0.41 [<.001] | 59.1 (28.6) | 24.7 (13.7) | 7.03 [<.001] | 0.13 [.161] |
| Plastic cup | 90 | 39.0 (10.0) | 32.8 (4.8) | 5.10 [<.001] | 0.07 [.339] | 64.0 (29.6) | 21.3 (10.9) | 10.61 [<.001] | 0.09 [.553] |
| ANOVA (*F* [*p*]) | | 128.48 [<.001] | 576.43 [<.001] | | | | 6.98 [<.001] | | |
| Wine, champagne | 172 | 10.1 (3.3) | 24.4 (4.5) | -28.38 [<.001] | 0.21 [.023] | 40.7 (22.0) | 9.8 (5.6) | 0.55 [.587] | 0.22 [.014] |
| Small Wine Glass | 59 | 10.7 (3.0) | 24.2 (2.1) | -22.82 [<.001] | 0.09 [.606] | 40.9 (16.5) | 9.9 (3.9) | 1.21 [.233] | 0.10 [.593] |
| Big glass wine | 65 | 10.0 (3.4) | 26.8 (3.8) | -24.14 [<.001] | 0.14 [.322] | 35.5 (21.2) | 9.6 (6.2) | 0.40 [.688] | 0.16 [.309] |
| Straight glass | 20 | 9.8 (3.4) | 21.8 (5.0) | -12.53 [<.001] | 0.58 [.036] | 53.2 (22.9) | 11.4 (6.0) | -1.10 [.305] | 0.55 [.037] |
| Cup / goblet | 28 | 9.5 (3.7) | 21.2 (5.9) | -8.63 [<.001] | 0.23 [.112] | 43.9 (29.6) | 9.2 (6.9) | 0.16 [.873] | 0.30 [.219] |
| ANOVA (*F* [*p*]) | | 0.81 [.495] | 9.88 [<.001] | | | | 0.39 [.759] | | |
| Liqueur, fortified wine | 45 | 6.8 (3.1) | 27.2 (8.3) | -14.96 [<.001] | 0.15 [.390] | 44.2 (32.1) | 11.6 (10.4) | -2.58 [.016] | 0.23 [.206] |
| Glass | 28 | 6.4 (3.1) | 25.1 (7.8) | -12.08 [<.001] | 0.13 [.642] | 37.7 (28.0) | 8.7 (7.2) | -1.60 [.127] | 0.27 [.248] |
| Other | 17 | 7.3 (2.9) | 30.6 (8.1) | -10.26 [<.001] | 0.08 [.707] | 58.3 (36.9) | 18.0 (13.6) | -2.38 [.041] | 0.03 [.934] |
| ANOVA (*F* [*p*]) | | 0.83 [.370] | 4.96 [.033] | | | | 4.84 [.037] | | |
| Straight spirits | 84 | 3.5 (2.2) | 18.1 (9.5) | -11.06 [<.001] | 0.01 [.945] | 56.1 (33.9) | 8.5 (6.7) | -4.76 [<.001] | 0.13 [.285] |
| Glass | 51 | 3.8 (2.6) | 17.3 (9.6) | -7.36 [<.001] | -0.03 [.878] | 61.2 (34.1) | 9.3 (7.5) | -3.47 [.002] | 0.06 [.653] |
| Plastic Cup | 33 | 3.0 (1.1) | 19.5 (9.2) | -8.03 [<.001] | 0.26 [.134] | 47.0 (32.0) | 7.2 (4.7) | -3.03 [.008] | 0.40 [.029] |
| ANOVA (*F* [*p*]) | | 3.72 [.061] | 0.57 [.454] | | | | 0.85 [.362] | | |
| Mixed drinks, cocktails | 314 | 27.7 (11.6) | 29.0 (7.2) | -1.56 [.123] | 0.34 [<.001] | 66.3 (24.1) | 19.1 (8.8) | 8.61 [<.001] | 0.34 [<.001] |
| Medium glass | 108 | 27.6 (11.7) | 29.2 (4.9) | -1.19 [.239] | 0.22 [.045] | 69.3 (22.9) | 20.4 (8.0) | 4.99 [<.001] | 0.26 [.010] |
| Medium plastic cup | 113 | 26.1 (10.6) | 25.2 (6.1) | 0.79 [.431] | 0.31 [.009] | 61.7 (24.7) | 14.9 (6.8) | 8.14 [<.001] | 0.34 [.021] |
| Long glass | 62 | 27.8 (11.9) | 31.3 (6.3) | -2.20 [.035] | 0.44 [.002] | 68.5 (23.4) | 21.8 (9.4) | 2.76 [.010] | 0.26 [.053] |
| Long plastic cup | 31 | 33.7 (12.9) | 37.5 (10.1) | -1.14 [.269] | 0.25 [.293] | 67.0 (26.9) | 23.6 (11.8) | 3.26 [.006] | 0.56 [.015] |
| ANOVA (*F* [*p*]) | | 2.26 [.087] | 14.46 [<.001] | | | | 9.15 [<.001] | | |

Notes: cl. = centilitres; Beverage types as reported by the participants; Container types as classified by the annotators; Effect sizes of mean difference tests (Diff. test) and Pearson correlations (Correlation) adjusted for the nested structure of the data.

Table 5: Participants' and annotators' estimates of ABV, drink sizes in grams and percentage of drinks larger than 10 grams, per beverage and container type

| | | ABV (in %) | Grams of pure alcohol | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Source: | | Participants | Participants | | Annotators | | | | Difference test [1] |
| | | | | | Container size | | Content size | | |
| | N | Mean (SD) | Mean (SD) | >10g. | Mean (SD) | >10g. | Mean (SD) | >10g. | *t* [*p*] |
| Beer, cider | 759 | 4.9 (0.7) | 15.8 (4.8) | 82.7% | 15.8 (4.4) | 89.7% | | | -0.12 [.908] |
| Aluminium can | 248 | 5.0 (0.6) | 19.4 (3.1) | 97.6% | 19.7 (2.6) | 99.6% | | | -2.67 [.010] |
| Bottle | 245 | 4.8 (0.7) | 12.9 (3.3) | 75.9% | 13.2 (2.8) | 88.2% | | | -1.86 [.066] |
| Glass | 176 | 4.9 (0.9) | 15.3 (5.3) | 74.4% | 15.8 (4.5) | 86.4% | | | -0.92 [.363] |
| Plastic cup | 90 | 4.7 (0.8) | 14.7 (4.8) | 76.7% | 12.2 (2.8) | 73.3% | | | 5.36 [<.001] |
| ANOVA (*F* [*p*]) | | 2.49 [.063] | 78.96 [<.001] | | 127.02 [<.001] | | | | |
| Wine, champagne | 172 | 11.3 (1.5) | 9.1 (3.4) | 47.7% | | | 8.6 (4.7) | 39.5% | 1.09 [.277] |
| Small Wine Glass | 59 | 11.3 (1.4) | 9.7 (3.1) | 50.8% | | | 8.8 (3.7) | 42.4% | 1.44 [.158] |
| Big glass wine | 65 | 11.3 (1.5) | 9.0 (3.5) | 41.5% | | | 8.3 (4.7) | 32.3% | 0.88 [.387] |
| Straight glass | 20 | 11.9 (1.7) | 9.3 (3.5) | 70.0% | | | 10.6 (5.4) | 60.0% | -0.93 [.378] |
| Cup / goblet | 28 | 10.6 (1.5) | 8.1 (3.5) | 39.3% | | | 7.5 (5.5) | 35.7% | 0.45 [.661] |
| ANOVA (*F* [*p*]) | | 1.58 [.202] | 1.11 [.352] | | | | 1.29 [0.286] | | |
| Liqueur, fortified wine | 45 | 16.8 (3.8) | 9.1 (4.8) | 42.2% | | | 14.1 (11.7) | 46.7% | -2.26 [.032] |
| Glass | 28 | 18.2 (3.3) | 9.4 (5.1) | 42.9% | | | 12.0 (10.8) | 42.9% | -1.30 [.211] |
| Other | 17 | 14.6 (3.6) | 8.6 (4.3) | 41.2% | | | 18.7 (12.6) | 52.9% | -2.08 [.068] |
| ANOVA (*F* [*p*]) | | 13.79 [.001] | 0.25 [.622] | | | | 2.48 [.127] | | |
| Straight spirits | 84 | 37.1 (7.0) | 10.3 (6.6) | 47.5% | | | 25.6 (21.5) | 75.0% | -4.23 [<.001] |
| Glass | 51 | 37.8 (6.7) | 11.5 (7.7) | 70.0% | | | 27.2 (22.7) | 85.0% | -3.29 [.003] |
| Plastic Cup | 33 | 36.1 (7.5) | 8.6 (4.0) | 25.0% | | | 22.7 (19.1) | 65.0% | -2.40 [.030] |
| ANOVA (*F* [*p*]) | | 0.65 [.424] | 4.45 [.041] | | | | 0.30 [.585] | | |
| Mixed drinks, cocktails | 314 | 15.4 (6.3) | 33.7 (20.5) | 74.8% | 36.0 (17.1) | 80.6% | | | -2.06 [.042] |
| Medium glass | 108 | 16.3 (6.1) | 35.7 (21.3) | 75.0% | 38.0 (14.8) | 84.3% | | | -1.44 [.155] |
| Medium plastic cup | 113 | 14.9 (6.4) | 30.6 (19.5) | 70.8% | 30.3 (15.1) | 72.6% | | | 0.17 [.868] |
| Long glass | 62 | 14.9 (6.1) | 34.1 (20.8) | 80.6% | 37.3 (18.0) | 90.3% | | | -1.73 [.094] |
| Long plastic cup | 31 | 15.0 (6.8) | 35.8 (20.0) | 77.4% | 43.9 (22.9) | 77.4% | | | -1.66 [.115] |
| ANOVA (*F* [*p*]) | | 0.73 [.534] | 0.81 [.490] | 82.7% | 3.02 [.034] | | | | |

Notes: 1) Difference between participants' and annotators' estimations; Beverage types as defined by the participants; Container types as classified by the annotators; Effect sizes of difference tests and ANOVAs adjusted for the nested structure of the data.