#Drink Or #Drunk: Multimodal Signals and Drinking Practices on Instagram

Thanh-Trung Phan
Idiap Research Institute & EPFL
Switzerland
tphan@idiap.ch

Skanda Muralidhar
Idiap Research Institute & EPFL
Switzerland
smuralidhar@idiap.ch

Daniel Gatica-Perez
Idiap Research Institute & EPFL
Switzerland
gatica@idiap.ch

ABSTRACT

The understanding of alcohol consumption patterns, especially those indicating negative drinking behavior, is an important issue to researchers and health policymakers. On social media, people share daily activities, including alcohol consumption, representing these moments through images and text. This work, using a five-year dataset from Instagram, analyzes what machine-extracted textual and visual cues reveal about trends of casual drinking (concepts gathered around #drink) and possible negative drinking (concepts gathered around #drunk). Our analysis reveals that #drunk posts occur more frequently in party occasions and nightlife locations, with a higher presence of people, while #drink posts occur at food locations, with a higher presence of drink containers. Manual coding further shows that #drunk posts have a higher chance of being perceived as potentially objectionable. A random forest classifier shows that #drink and #drunk posts can be discriminated with accuracy up to 82.3%. These results have important implications for alcohol research among youth.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing**; *Ubiquitous and mobile computing design and evaluation methods*;

KEYWORDS

Alcohol consumption, social media, Instagram, deep learning

PervasiveHealth'19, May 20–23, 2019, Trento, Italy

ACM ISBN 978-1-4503-6126-2/19/05…$15.00
https://doi.org/10.1145/3329189.3329193

ACM Reference Format:
Thanh-Trung Phan, Skanda Muralidhar, and Daniel Gatica-Perez. 2019. #Drink Or #Drunk: Multimodal Signals and Drinking Practices on Instagram. In *The 13th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth'19), May 20–23, 2019, Trento, Italy*. ACM, New York, NY, USA, Article 39, 10 pages. https://doi.org/10.1145/3329189.3329193

1 INTRODUCTION

Alcohol plays an important role in many cultures including social interaction and health [59]. However, people also abuse alcohol, leading to negative social and health outcomes including injuries, violence, accidents, and fatalities [27, 55, 61]. In this work, we examine the drinking behavior of youth using self-representation on social media (Instagram), with the aim to uncover patterns of positive drinking and potential negative drinking.

Drinking culture "refers to the customs and practices associated with the consumption of alcoholic beverages" [3] and is prevalent in many societies. As part of this, social drinking or responsible drinking, practices correspond to "casual drinking of alcoholic beverages in a social setting without an intent to become intoxicated" [3]. This drinking pattern is in contrast with negative drinking, which involves alcohol intake "far beyond that which is socially acceptable" [3]. In practice, alcohol consumption often begins as a casual, social activity [36] until excessive consumption leads to negative effects [27]. Due to the above mentioned negative consequences, it is important to limit excessive drinking by setting up prevention efforts [30], and to understand the possible transitions between casual drinking and negative drinking.

Researchers in psychology and alcohol consumption have studied drinking behavior from the perspective of drinking motives [18, 19, 42] and consequences of drinking [27, 55, 61]. Most of these works collected alcohol consumption data using face-to-face interviews or paper-and-pencil questionnaires [18, 19] , which are known to have limitations due to incorrect reporting, e.g. limited recall [24, 43]. The advent of ubiquitous sensors and smartphones aided researchers to collect larger amounts of data including in-situ responses via SMS on features phone [44], wearable sensor data [9, 11, 38], and hybrid data (including sensor data, and human-generated data like photos, captions, and location) [16, 56, 57, 62]. However, these methods of data collection have the disadvantage of being intrusive because participants are asked to intentionally report their alcohol intake.

As a complement to such data collection methods, social media from Twitter and Instagram also provide in-situ data related to drinking events, including photos, captions, locations, and other semantic information. By taking advantage of the large scale of users and media posts, researchers can study aggregate alcohol patterns in countries or cities [33, 39, 51, 53]. It must be noted that social media data also has some limitations, in terms of population bias and the lack of continuous temporal data for individual users.

In this work, using data from Instagram collected over five years, we study two aspects of drinking culture (social drinking and negative drinking) inside Switzerland (as a European case study) using users' posts (both textual and visual). We hypothesize that #drunk (or its variations) might be indicative of **possibly negative drinking**; vs. #drink (or its variations) as **casual drinking** (i.e., with a more positive connotation). Specially, we investigate two research questions:

RQ1: Are there significant differences depending on the specific representation used to characterize the act of drinking? More specifically, what are the textual content, visual content, and perceived differences between possibly negative drinking (e.g. posts containing #drunk or related hashtags) and casual drinking (e.g. posts containing #drink)?

RQ2: If such differences exist, how can textual and visual features of Instagram posts be used to automatically classify #drink and #drunk drinking episodes?

The specific contributions of the paper are the following: (1) We curate a 1.7M Instagram dataset collected over 5 years inside Switzerland to obtain a corpus for #drink and #drunk. This is done using a dictionary of hashtags defined manually for #drink, #drunk, drinking occasion, location, social context, and alcohol categories. (2) We automatically extract textual and visual features. Textual features include drinking occasion, location social contexts, and alcohol categories. Visual features were extracted using Microsoft Azure and consist of visual autotags, descriptions, and categories. In addition, we obtain crowdsourced perceptions of drinking motivations from #drink and #drunk posts using a popular online platform. (3) An analysis of textual features indicates that #drunk posts occur more often for nightlife and events locations, and for party occasions, while #drink occurs often in food locations like restaurants. (4) Visual analysis shows that photos in #drink posts contain higher presence of drink-related artifacts (like beverages, glasses, or tables), while, #drunk posts have a higher presence of people, specifically males. (5) The analysis of manually annotated perceived drinking attributes in posts shows that #drunk posts have higher scores for drinking motives (social, coping, enhancement, and conformity, as defined by a validated short questionnaire) compared to #drink. Furthermore, 19% of #drunk posts and 6% of #drink posts were labeled as potentially problematic. (6) We show that using machine learning algorithms, textual and visual cues are able to discriminate #drink and #drunk. Specifically, textual cues achieve a classification accuracy of 82.3%, while visual cues achieve an accuracy of 75.0%.

This work has implications not just for alcohol consumption research but also for automatic classification of potential negative drinking social media posts in health tracking applications. Our work could also be useful for understanding consequences related to mental and physical health through self-representation in social media.

2 RELATED WORK

Self-Presentation and Social Media

Self-presentation refers to how people try to manage impressions of themselves with respect to other people or entities [58]. Goffman explained it as strategic activities of the individual to "convey an impression to others which it is in his interests to convey" [28]. In detail, self-presentation relates to how people try to make up attitudes and reactions of audiences through the presentation of self-relevant information. In the age of the Internet, people have developed strategies for self-presentation in personal web spaces [35] or online dating environments [25]. The boom of social networks in the last decade, e.g. Instagram, has opened more channels for people to self-present.

The epitome of self-presentation on Instagram is the selfie [60] which is an efficient medium to occupy audiences' attention. In previous work, [12] shows that photos with faces are 38% more likely to receive likes and 32% more likely to receive comments. Likes and comments are means of responding to other Instagram users' posts. Based on a study of 27K teens and adults in Instagram [34], teens tend to post fewer photos but more hashtags and to get more likes/comments. In addition, teens show a higher sense of self-presentation than adults through posting more selfies and self-expression photos. Interestingly, self-presentation on Instagram can reveal aspects of user's mental health. For instance, [54] uses color analysis, metadata components (i.e. image filters), and face detection to diagnose rate of depression. In addition, the relationship between self-presentation behavior in Instagram and narcissism is investigated in [50], finding that the higher a user scores in narcissism, the more often they post selfies or update their profile pictures. Hence, self-presentation on Instagram has potential to be informative of other user behavioural patterns. In our work, we consider alcohol-tagged posts as one of kind of self-presentation that could reveal drinking practices of Instagram users.

Alcohol Consumption and Social Media

Alcohol consumption is a topic of great interest to researchers and policymakers. The literature has shown that drinking motivations can be broadly categorized into four types: social, enhancement, coping, and conformity [18, 20, 21]. Specifically in the context of young adults and adolescents, it has been shown that alcohol is often used as a stimulant for initiating conversations (social) and/or due to peer-pressure (conformity) [52]. Previous works also indicate that some young adults indulge in heavy drinking with the motivation of enhancement [41], leading to alcohol abuse [37], [47].

Traditionally, alcohol research has relied on paper-and-pencil questionnaires and/or face-to-face interviews for data collection [18, 20, 21, 63]. With the advances in ubiquitous computing and the widespread availability of smartphones, the use of mobile technology for data collection in alcohol research has increased. Specifically, literature has shown the validity of data collected using mobile phone applications or wearable devices [9, 11, 13, 38, 49]. Participants are asked to use smartphones to report their drinking events [43, 56]. Researchers collect data on drinking through reported drink images, social context, ambiance context, etc. In addition, participants are asked to answer questionnaires or drinking habits, which are used as validation data. For example, [57] uses sensor and log data to classify drinking nights with 76.6% accuracy.

As mentioned above, alcohol-tagged posts could reveal drinking practices of Instagram users. [14] stated that Instagram was the most likely destination for posts glamorizing college drinking. By asking volunteers to answer surveys about themselves and give access right of their social network accounts, researchers can collect participants' ground truth and their social network data. [32] asked 192 participants (mean age 20.64, 132 women and 54 men) to give access to their Facebook and/or Instagram profiles and their timelines which are analyzed to discover users' behaviour w.r.t drinking. Similarly, through group interviews with 37 young women aged 16-21, [10] explored experiences of drinking and intoxication, the use of social networks in their drinking cultures and the display of drinking practices on social media. All this work has involved manual work, e.g. no machine learning has been used.

In another direction, researchers collected available posts from social networks in a larger scale. In [53], we investigated the food and drink patterns in Switzerland by using a Instagram corpus. They created a vocabulary of food and drinks by manually annotating the top hashtags on Instagram and classified six daily meals at 61.7% and healthy/unhealthy posts at 85.8%. Similarly, [7] and [48] use hashtags to analyze food/drink patterns including alcohol consumption. These works did not investigate particular drinking patterns like casual drinking or possibly negative drinking, as we do in our paper. The work in [51] investigated posts containing alcohol-related hashtags (textual cues) and Instagram user demographics (visual cues). This work obtained users' demographics from photos via Face++ (an online face processing platform). This work found patterns of alcohol consumption w.r.t time, age, location, and exposure to specific alcohol types, e.g. beer or vodka. We also use hashtags, captions, meta data, and picture contents on Instagram posts as users' self-presentation features, to examine the novel angle of distinguishing between casual and possibly negative drinking behaviours through two categories: *#drink and #drunk.*

3 #DRINK AND #DRUNK DATASETS

We use a dataset collected in our previous work [53]. The dataset contains 2.8M picture posts between November 2010 and March 2016 within Switzerland. This was then filtered to have at least one hashtag per post, leading to 1.7M posts. In this paper, by using the alcohol-related vocabulary from the food & drink dictionary defined in [53], we harvest 34K alcohol drink posts from the 1.7M dataset. We call this dataset the *34K dataset.* We hypothesize that posts containing #drink and similar hashtags, co-occurring with explicit alcohol hashtags (e.g. #wine, #beer, etc.), could be evidence for casual drinking, while those posts containing #drunk and similar hashtags could represent possibly negative drinking. In addition, we hypothesize that users posting about these two types of drinking practices also use hashtags related to social relationships (#friend, #family), occasions (#party, #festival), and locations (#bar, #restaurant). From the 34K dataset, we extracted and manually annotated the top 2000 hashtags ranked by their frequency. Based on the meaning of the 2000 hashtags, we manually annotated, grouped, and defined a #drink/#drunk hashtag dictionary, a location hashtag dictionary with ten venue categories, an occasion hashtag dictionary with six categories (*travel_vacation, holiday, sport, party, festival, other*), and a social hashtag dictionary with five categories (*friend, alone, partner_spouse, family, other*). In detail, the #drink/#drunk hashtag dictionary has 20 hashtags for casual drinking (*drinkup, alcoholinfused, alcoholdrink, alcoholicdrink, alcoholicdrinks, alcoholsucks, drink, drinks, drinking, instadrink, drinkingcraft, drinklocal, boire, drank, drinkin, instadrinks, saufen, slurp, trinken, drinkporn*) and 9 hashtags for possibly negative drinking (*drunk, wasted, getdrunk, ivresse, boozing, instadrunk, tipsy, drunken, getdrunk*, пьянствубой) for a total of 29.

The hashtags are in multiple languages (German, French, English, and Russian) reflecting that Switzerland is a multilingual country and a tourist destination. With the location dictionary, we borrowed the definition from Foursquare venue categories, namely, Arts & Entertainment, College & University, Events, Food, Nightlife Spots, Outdoors & Recreation, Professional & Other Places, Residence, Shop & Services, and Travel & Transport, and Other or None [5]. Then, for the #drunk construct, we used the 1.7M dataset to filter out all posts which mentioned at least one of the hashtags in the #drunk hashtag dictionary. For the #drink construct, in order to avoid posts with non-alcoholic drinking, we selected posts with both an explicit alcohol hashtag (e.g. #wine) and at least one of the hashtags in the #drink dictionary. This resulted in a corpus of 2046 #drink and 1323 #drunk posts. We call this corpus *the #Drink/#Drunk corpus*. The corpus is composed of 1451 and 952, users respectively. Similarly to [53], we also link our original corpus to Foursquare venue categories. In the end, the corpus filtered by occasion, social, context location, alcohol hashtag dictionaries, and 4sq categories is shown in Table 1.

In order to validate the above 29 drink/drunk hashtags as being related to positive and negative drinking, we asked 10 trusted annotators to think about the meaning of each hashtag and search it on Instagram first and Google Images later. Then, we asked three questions. Specifically, we asked "Please think about this hashtag and the pictures you have searched and rate it according to its positive connotation". The second question was worded similarly asking for negative connotation. The annotator had to answer this on a 5-point Likert scale (1=strongly disagree and 5=strongly agree). In the last question, we asked "Do you think that if users posted those pictures they could have bad consequences?" with two answers: 1 (Yes) or 0 (No). To assess agreement of annotators, we used Intraclass Correlation Coefficients (ICC(2,k)) as recommended in [40]. ICC(2,k) of #drink and #drunk on positive/negative connotations is good (0.60-0.80) as shown in Table 1. We observe a difference between the mean for positive (4.06 vs. 2.52) and negative (1.66 vs. 3.22) connotation for #drink and #drunk hashtags. In addition, the mean of bad consequence of #drink hashtags is 0.16 while the value of #drunk hashtags is 0.54. These results suggest that our hypothesis of #drink hashtags as signaling casual, positive drinking, and of #drunk hashtags as signalling more negative drinking is reasonable.

4 #DRINK AND #DRUNK ANALYSIS (RQ1)

In this section,we analyze patterns of drinking (expressed by #drink and #drunk) using textual content (hashtags) and visual content (image) from the Instagram posts, as well as other attributes obtained from human perception.

Textual Content

As a first step, we manually extract all hashtags from the posts in our corpus (1,323 #drunk and 2,046 #drink) and cluster them according to (a) alcohol type (wine, beer, spirit & cocktails, and others), (b) occasion (holidays, events, party, travel, sports and festival), (c) social context (friend, family, partner_spouse, alone and other), and (d) location categories (10 4sq venues) as described in Section 3.

Table 1: Instagram datasets used in this work.

Corpus and its derives	#drink	#drunk
#Drink/#Drunk corpus	2,046	1323
#posts containing alcohol hashtags	2,046	453
#posts containing occasion hashtags	881	682
#posts containing social hashtags	594	585
#posts containing location hashtags	608	266
#posts linked to 4sq venues	1351	859
ICC(2,k) for negative connotation	0.80	0.60
Mean of negative connotation	1.66	3.22
ICC(2,k) for positive connotation	0.72	0.66
Mean of positive connotation	4.06	2.52
ICC(2,k) for bad consequence	0.68	0.56
Mean of bad consequence	0.16	0.54

The descriptive statistics for each group of hashtags are presented in Table 2. We use unpaired T-tests to compare the two groups for each variable. As p-values are known to be not sufficiently informative [45, 64], we additionally use effect size, namely Cohen's d (CD), and 95% confidence interval (95% CI) to expand our understanding of statistical significance [45, 64]. In Table 2, the cases that are statistically significant (by combining the effect size and the CI not including zero) are the number of alcohol hashtags (medium effect size), social hashtags (small effect size) and location hashtags (small effect size). In other words, #drunk posts use fewer alcohol-related hashtags (e.g. wine, beer) and location hashtags but slightly more hashtags related to social interaction. As a basic point of reference to these numbers, we compute descriptive statistics on a disjoint, random sample of 2046 posts from the general Instagram 1.7M-post dataset containing at least one hashtag. Compared to our #drink/#drunk dataset, this random sample of general Instagram posts differs for all variables.

As the next step, we study the distribution of #drink and #drunk posts for the various categories (alcohol type, occasion, and location), see Figure 1. We also study the distribution of the social context but it is not shown for space reason. Unsurprisingly, friends are the dominant social context, with most posts that contain a social hashtag refer to friends ($0.84 - 0.91$). Figure 1a shows the distribution of *alcohol category* hashtags. We observe that #drunk posts have a similar frequency for wine, beer, and spirit ($0.25 - 0.30$), while #drink posts have a substantially higher frequency for *spirit & cocktails* (0.45), indicating that cocktail & spirit are popular representations of the #drink concept. Figure 1b shows the distribution of *occasion* hashtags. We observe that #drunk posts are often mentioned at parties (0.53), while #drink posts spread over other occasions. Figures 1c and 1d show the distribution of location categories for #drink and

Table 2: Descriptive statistics of hashtags for random posts (N=2046), #drink posts (N=2046) and #drunk posts (N=1323). (CD denotes Cohen's d, MD denotes mean difference, and CI denotes confidence interval)

# hashtag per post	Random Posts				#Drink				#Drunk				#Drink vs. #Drunk			
	mean	med	sd	max	mean	med	sd	max	mean	med	sd	max	MD	CD[95% CI]	T-Test	p-value
# hashtags	8.40	5	7.98	35	15.45	15	7.51	30	14.06	12	7.96	30	1.39	0.18 [0.11, 0.25]	5.1	1.0e-6
# words	5.42	2	12.06	222	4.17	1	8.76	186	3.35	1	5.81	71	0.82	0.11 [0.04, 0.18]	3.27	0.001
#comments	2.29	1	6.30	113	1.43	1	2.35	33	1.62	1	2.35	25	-0.19	-0.08[-0.15,-0.01]	-2.3	0.021
# Alcohol hashtags	0.02	0	0.23	7	2.08	1	2.32	18	0.67	0	1.61	16	1.41	0.68 [0.61, 0.75]	20.88	1.0e-90
# Social hashtags	0.08	0	0.36	5	0.43	0	0.86	8	0.64	0	0.89	6	-0.22	-0.25 [-0.32, -0.18]	-7.0	1.0e-11
# Occasion hashtags	0.40	0	1.11	15	0.70	0	1.05	11	0.73	1	0.88	5	-0.03	-0.03 [-0.10, 0.04]	-0.98	0.327
# Location hashtags	0.17	0	0.45	5	0.37	0	0.65	5	0.25	0	0.55	4	0.12	0.20 [0.13, 0.27]	5.9	1.0e-8

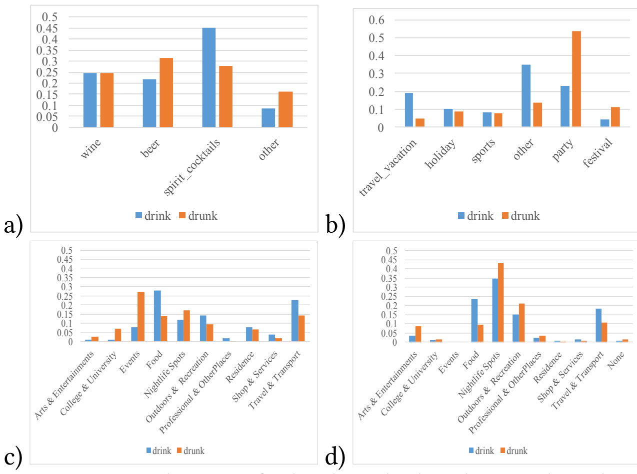

Figure 1: Distribution of #drink and #drunk posts based on: (a) Alcohol categories (b) Occasion (c) Location-related Hashtags (d) Foursquare venues

#drunk posts based on 4sq-linked venues and location hashtags, respectively. While the two methods that associate the location with posts produce different distributions for #drink and #drunk, we observe three consistent trends. Specifically, the distribution of #drink is higher than #drunk at food locations and travel & transport, while #drunk is higher at nightlife spots. On the other hand, private places (residence) are poorly represented, even though it is known that alcohol drinking at home is common among young people [56]. This quantitative analysis complements previous results in the literature reporting about the drinking location of young people using qualitative methods [15, 23].

In summary, #drink and #drunk posts include references to friends, parties/events, and non-private venues, with #drunk posts being more often associated with parties and nightlife, and #drink posts with food and travel & transport. We also showed some small to medium effects regarding the frequency of use of alcohol, social, and location hashtags.

Visual Content

In this section, we examine the photos of #drink and #drunk posts through visual cues (*categories, autotags, and descriptions*) and facial features (*face presence and expressions*) extracted using the Computer Vision API [1] and Face API [4] from Microsoft Azure [26, 29, 31].

Visual Categories. Each image has at least one category name from the top 15 parent/child hereditary hierarchies [1]. Figure 2a shows the distribution of #drink and #drunk posts for these 15 visual categories. We observe that considerably more photos from #drunk posts than from #drink posts have people in them (0.55 vs. 0.21), while photos from #drink posts include drinks, indoor, and food. This differentiated trend suggests that #drunk photos are more often about people, while #drink photos have a wider variety of content, adding drinks and food content to human presence.

Visual Autotags and Descriptions. Autotags correspond to over 2,000 recognizable objects, living beings, scene hierarchy, and actions. These autotags are returned along with a confidence value. We use these autotags for our dataset, with 153 autotags having confidence value higher than 90%. Figure 2a shows the distribution of the top vision autotags with confidence above 90%. We again observe that #drunk posts have a higher frequency than #drink posts for autotags related to people, while #drink has a higher frequency for indoor scenes and objects including tables, cup, and bottle. Additionally, visual descriptions correspond to full semantic descriptions like "glass of beer on the table" for each image. Figure 2c shows that #drunk photos contain photos with people posing, while #drink photos contain objects related to eating/drink: (*"a glass of beer on a table, a glass of wine,"*, etc.). This confirms that #drunk photos are quite often about people, while #drink photos also depict content related to the drinks themselves.

Face Features. The Face API detects faces and recognizes facial expressions and face exposure of people in pictures. We use a smile cue (the percentage of smiling faces in a photo, ranging from 0 to 1), and an exposure level cue (the average of the exposure of all detected faces in a photo). Photographic exposure corresponds to the amount of light on the face, ranging from zero (underexposed faces) to one (overexposed faces). Table 3 shows the descriptive statistics of face features. We observe that 55% (727/1,323) of #drunk posts contain faces compared to 21% (429/2,046) of #drink. Furthermore, the mean number of faces in #drunk posts is higher than in #drink posts. This corroborates our previous observations that #drunk posts tend to contain more people. We then compute differences for the two groups w.r.t. the number of male and female faces, the percentage of smiling

Table 3: Descriptive statistics and significance test for #drink posts (N=429) and #drunk posts (N=727) containing at least one face (except (*) estimated from N = 2046 #drink posts and N=1323 #drink post). (CD denotes Cohen's d, MD denotes mean difference, and CI denotes confidence interval)

Vision Feature	Feature	#Drink		#Drunk		#Drink vs. #Drunk			
		mean	sd	mean	sd	MD	CD [95% CI]	T-test	p-value
Face	#face (*)	0.44	0	1.33	1	-0.89	-0.89 [-1.00,-0.78]	-14.68	1.0e-45
	#face	2.13	1.55	2.47	2.20	-0.35	-0.17 [-0.30,-0.05]	-3.08	0.002
	#male	0.99	1.06	1.32	1.48	-0.33	-0.25 [-0.37,-0.13]	-4.38	1.3e-5
	#female	1.14	1.30	1.15	1.52	-0.01	-0.01 [-0.13,0.11]	-0.15	0.88
Expression	smile	0.71	0.40	0.60	0.43	0.11	0.25 [0.17,0.33]	6.21	1.0e-9
	exposure	0.58	0.18	0.62	0.19	-0.03	-0.18 [-0.26,-0.10]	-4.47	1.0e-5

faces in a photo, and the average exposure of faces in a photo (Table 3). An unpaired T-test followed by the estimation of effect size and 95% CI indicate that the differences between the two groups w.r.t. the number of faces (in the full sample, with large effect size), and the number of male faces and the smiling cue (in the subsample of images containing at least one face, both with small effect size) are statistically significant, as the corresponding CIs do not include zero. In contrast, the difference in the number of female faces is not significant (p=0.88, negligible effect size and CI including zero in the subsample of images containing at least one face).

In summary, by using state-of-art computer vision algorithms and several ways of characterizing the visual content of images, we found that that #drunk posts significantly depict more people (and males in particular) in the corresponding photos, while #drink posts relate to both people and eating/drinking activities.

Human Perception of Drinking Posts

In this section, we study a final issue: are #drink and #drunk posts perceived differently by human observers? For this, we analyze how #drink and #drunk posts are perceived based on three issues: drinking motives, context, and problematic issues. We randomly chose 200 #drunk posts and 200 #drink posts from 400 distinct Instagram users, including photos, captions, and hashtags. These 400 posts were then manually coded on three dimensions: (1) 12 questions of an adapted questionnaire on four drinking motives (social, coping, enhancement, conformity) using a 1-to-5 Likert scale [42]; (2) two questions about the context of the posts, namely social relationship (*Who does the person appear to be with?*) and place (*Where does the person appear to be?*); and (3) problematic issues with the posts, with yes/no choices (*Do you think this post could have negative consequences on the user?*). Annotations were conducted on Crowdflower [2], with each post annotated by 5 raters.

Perceived Drinking Motives. For the four dimensions of drinking motives, the descriptive statistics are shown in Table 4. All means are higher for the #drunk group. For #drunk posts, social and conformity motives are the two motives and have a mean equal or higher than 4.0 (i.e., one point above the middle of the Likert scale). Social and enhancement motives are the top two motives, for the #drink group. Similarly to the previous analysis in this section, we calculate effect size and 95% CI for the differences between #drink and #drunk for each drinking motive, as shown in Table 4. The differences for the four drinking motives appear to be significant, with one medium effect size (enhancement), three large effect size (social, coping, conformity) and all CIs not including zero.

Perceived Problematic Posts in Context. Among the 400 annotated posts, there are 11 #drink and 37 #drunk posts perceived as having problematic consequences (for this difference, Fisher's Exact Test score = 3.9, p=8.5e-5). In the questionnaire, we also asked observers *"Why do you say so?"*. We summarize the content of the answers into three groups: (1) people in pictures seem to behave inappropriately, appear naked, or are seen as too young; (2) alcoholic drinks appear to be combined with energy drinks (e.g. Red Bull) or with other hard alcohol, which makes observers believe this might lead to negative consequences; and (3) observers stated that the people in photos appeared to have drunk in excess and might drive a car afterwards. Next, we examine these problematic posts in the context of place and social relations which are shown in Figure 3. First, Figure 3a shows that most of the cases of problematic #drink and #drunk posts correspond to bars/pubs, nightclubs, and personal places. Problematic #drunk posts are relatively more frequent than those for #drink at bar/pub, in accordance with the results discussed earlier in this section in Figure 1c,d. Interestingly, problematic #drunk posts are perceived as happening in personal places with 13%. This result differs from our earlier observations in this section, and opens a relevant question about drinking and private places. Second, the question *"Who does the person appear to be with?"* shows that half of the problematic posts have an *Impossible to say* answer, which indicates that annotators cannot infer any social relationships. A manual inspection of these posts showed that there are several reasons for this ambiguity e.g. photos depict two or more alcoholic drinks but not visible people, etc. In Figure 3b, we see that for the rest of the cases people are perceived to be in the company of same-sex friends, with higher values for #drunk posts than #drink posts in problematic context. These results also match the results using visual/textual content discussed earlier in this section. In contrast, #drink posts are perceived to be more problematic than #drunk when in mixed company.

In summary, using crowdsourced annotations for 400 #drink and #drunk posts, we obtain three results related to

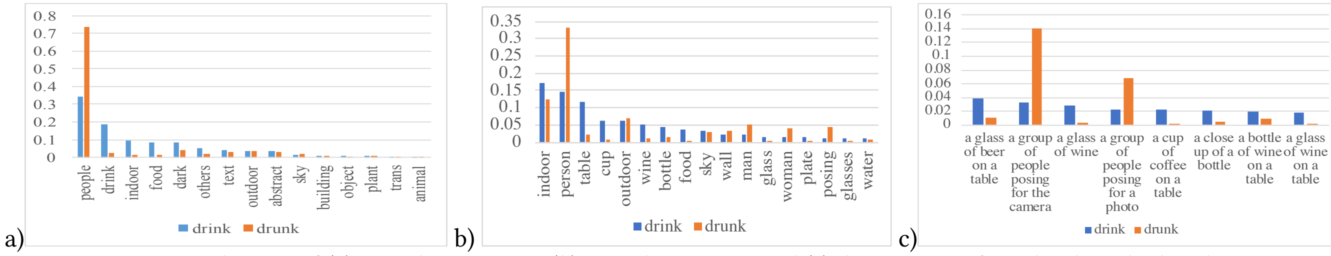

Figure 2: Distribution of (a) visual categories, (b) visual autotags, and (c) descriptions for #drink and #drunk posts.

Table 4: Drinking motives descriptive stats and significance for #drink posts (N=200) and #drunk posts (N=200). (CD denotes Cohen's d, MD denotes mean difference, and CI denotes confidence interval)

Motives	#Drink		#Drunk		#Drink vs. #Drunk			
	mean	SD	mean	SD	MD	CD[95% CI]	T-Test	p-value
Social	3.77	0.56	4.29	0.43	-0.52	-1.04[-1.25,-0.83]	-10.38	2.2e-16
Coping	3.0	0.56	3.8	0.70	-0.80	-1.28[-1.49,-1.06]	-12.79	2.2e-16
Enhancement	3.63	0.32	3.82	0.35	-0.18	-0.55[-0.75,-0.35]	-5.46	8.3e-08
Conformity	3.4	0.51	4.0	0.56	-0.65	-1.23[-1.44,-1.01]	-12.25	2.2e-16

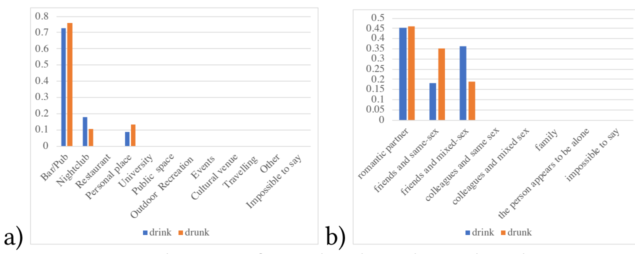

Figure 3: Distribution of 11 #drink and 37 #drunk posts possibly having problematic consequences of (a) Place (b) Social

perception of Instagram drinking practices: perceived drinking motives are scored higher for #drunk posts, for which social and conformity motives are the top ones; the perceived drinking context generally agrees with the trends obtained with visual and textual content results in the previous two subsections; and significantly more #drunk posts were perceived as problematic (19%) compared to #drink posts (6%).

5 CLASSIFYING #DRINK AND #DRUNK(RQ2)

We now investigate how the textual and visual cues analyzed in the previous section could be used to automatically discriminate between #drink and #drunk posts, defining a binary classification task.

Feature Extraction

In the first step, we extract a number of features from the visual and textual modalities (shown in Table 5 with name, description, type and group). Note that all hashtags used to define the #drink and #drunk clusters (see Section 3) have been removed so they are not part of the features. Picture caption (PC) contains the numerical count of hashtags in the caption (general hashtags, words, alcohol hashtags, occasion hashtags, social hashtags, alcohol category hashtags), and the categories of social context and occasions. Time (T) corresponds to the hour and weekday when photos are posted.

Table 5: Features for classification of #drink and #drunk posts with the following group features Picture Caption (PC), Vision Autotag (VA), Fine Vision Autotags (VCA), Vision Colors (VCO), Face (F), Attention (A), Time (T).

Feature	Description	Type	Group
tagCount	Total number of hashtags	numeric (1)	PC
wordCount	Total number of words in caption	numeric (1)	PC
alcoholTag	Total number of alcohol hashtags	numeric (1)	PC
alcoholCategory	Total number of wine/ beer/ spirit&cocktails hashtags	numeric (3)	PC
socialCount	Total number of social-related hashtags	numeric (1)	PC
occasionCount	Total number of occasion-related hashtags	numeric (1)	PC
socialCategory	5 social hashtags	categories (5)	PC
occasionCategory	6 occasion hashtags	categories (6)	PC
Visual Autotags	Generated autotags for auto description from Azure Vision	categories (520)	VA
Fine Visual Autotags	VA with confidence values higher 90%	categories (153)	VAC
Visual Categories	Generated categories of image from Azure Vision	categories (67)	VCA
faceCount	Number of total faces and male faces in picture	numeric (2)	F
age	Min, Max, Mean of age	numeric (3)	F
exposure	Min, Max, Mean of distribution of exposure	numeric (3)	F
commentCount	Total number of comments of picture	numeric (1)	A
likeCount	Total number of likes of picture	numeric (1)	A
hours	Hour when picture is posted (in minutes)	numeric (1)	T
day	Day when picture is posted (weekday)	numeric (1)	T

Attention (A) includes the count of comments and likes on posts.

Classification Task

To classify #drink vs. #drunk posts, we use a Random forest (RF) algorithm. In parameter setting, we set ntree=500, mtry as recommended by [46] and GridSearch supported by [6]. Then, we use 10-fold cross validation over 10 times for accuracy evaluation. For the classification experiments, we use 1,242 data points for the #drunk class, and randomly sample the same number of posts from the #drink class to have a balanced dataset (total: 2,484 data points) with a random baseline of 50%.

Classification Results and Comparison

Classification results are shown in Table 6. With visual features, we see that the visual autotags produce the highest accuracy (75%), followed by fine visual autotags (68.9%), visual categories (67.4%), and face features (66.6%). With textual features, we see that the picture captions (PC) produce the highest accuracy (82.3%), followed by time (60.8%) and attention (52.5%).

Table 6: Classification accuracy of #drink/#drunk classifier for visual and textual.

Modality	Feature	Accuracy
	Baseline	**50.0**
Visual	Visual Autotags (VA)	**75.0**
	Fine Visual Autotags (VAC)	68.9
	Visual Categories (VCA)	67.4
	Face (F)	66.6
Textual	Picture Caption (PC)	**82.3**
	Attention (A)	52.5
	Time (T)	60.8

In summary, basic visual and textual features are useful to classify #drink and #drunk. We see that the textual content not only plays a better role in classification (82.3%) but also take less effort as they are directly available in the posts.

6 DISCUSSION AND IMPLICATIONS

In this section, we discuss the results presented in the last two sections and their implications.

Social signals and Instagram drinking posts. The higher use of social-related hashtags (Table 1), combined with the machine recognition results (Figure 2 and Table 3) and the social motives for drinking (Table 4), suggest that #drunk posts might carry a stronger social connotation. In contrast, the higher visual presence of drink-related artifacts (glasses, bottles, etc.) in #drink posts suggests that the conveyed signal, while clearly social, could also serve other objectives related to documenting specific drink items, or moments when people are alone. The finding adds to the literature investigating social media practices related to personal tracking of food and drink [17, 22]. Future qualitative research interviewing Instagram users could complement and deepen these results from the perspective of health-related issues.

Places and Instagram drinking posts. Our analysis also showed some differences with respect to the use of places (Fig. 1c,d). We observe more cases of #drunk posts in bars/pubs, nightclubs, and personal places (Fig. 3a). This result, though not surprising, has implications for health and security in and around these places. These also have implications for police authorities and policymakers. In this dataset, private places are underrepresented (see category "residence" in Fig. 1c and Fig. 3a). This is important as research has shown that alcohol is often consumed by youth at home [8, 56]. Further research could specifically investigate the issue of alcohol drinking and social media practices in the home context.

Human perception of Instagram drinking posts. Social motives, corresponding to external orientation and positive reward [18], were perceived as the strongest motive for both #drink and #drunk posts. On the other hand, #drunk posts have higher scores for all motives compared to #drink posts, with large effect sizes (Table 4). To our knowledge, this issue has not been previously studied. Finally, 19% of #drunk posts and 6% of #drink posts were labeled as potentially problematic (Fig 3c). The comments by the external observers highlight a variety of issues. While previous survey-based research has found that youth are generally aware of the potential negative consequences of excessive drinking (33.7% of young Europeans believe that overdrinking might make them do something they would regret, and 10% believe they might get in trouble with the police) [8], our analysis shows that these kind of photos are clearly in circulation within Instagram. Furthermore, the risks of sharing posts about alcohol drinking take on a new dimension with the current advances in machine recognition.

Machine recognition of Instagram drinking posts and health-drinking tracking applications. Our results showed that #drink and #drunk posts can be discriminated up to 82%. Given the current trends in deep learning, we anticipate that performance could increase if larger datasets were available and more advanced models were applied. Our results suggest that the recognition of different forms of drinking could be automated to some degree in the future. Possible health applications include privacy-sensitive tools for self-reflection and self-management of personal habits; and anonymized contributions towards public health studies. There are, however, important risks associated with the inference of such kind of personal information, especially if used for other purposes that are not directly meant to support or benefit users.

7 CONCLUSION

In this work, we investigated the patterns in drinking behavior (both social drinking and potential negative drinking) inside Switzerland using Instagram posts from a five-year Instagram dataset. We conclude this paper by summarizing the findings of the two research questions we posed, and by discussing limitations and future directions.

Regarding RQ1, the following patterns were observed: (a) textual features indicated that a majority of the drink and drunk posts include references to friends (over 84% of posts), parties/events, and non-private venues (nightlife spots, outdoors, food, and travel & transport). In particular, #drink posts occurred more often with food and travel transport, while #drunk posts occurred more often at parties and nightlife. (b) Visual features indicated a social tendency among #drunk posts with a higher presence of people, specifically males, while #drink posts contain the higher presence of drink-related artifacts (like beverages, glasses, tables). (c) manual coding indicated #drunk was rated higher for all drinking motivations, with social and conformity being the top ones. The perceived drinking context is in line with trends obtained using the visual and textual content. Furthermore, significantly greater drunk posts were perceived as being problematic (19%) compared to drink posts (6%).

Regarding RQ2, we observed that textual and visual cues in posts are able to discriminate #drink and #drunk, with textual cues showing improved classification accuracy (82.3%) but representing less computational effort due to their direct availability in captions as compared to visual cues. We believe that our work has implications not just for alcohol consumption research but also for automatic classification of potential negative drinking of health-drinking tracking applications.

Our dataset, curated in Switzerland, could have limitations of generality w.r.t. other locations. In addition, over 1M out of the 2.8M posts contain no hashtags, which leaves many Instagram users (who follow this practice) out of our analysis. In the future, we plan to investigate Instagram alcohol consumption in private vs. public spaces and understand possible differences.

ACKNOWLEDGMENTS

This work has been funded by the Swiss National Science Foundation through the Dusk2Dawn Sinergia project, and a Swiss Government Excellence Scholarship.

REFERENCES

[1] 2018. Computer Vision API on Microsoft Azure. (2018). Retrieved April 12, 2018 from https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/home

[2] 2018. Crowdflower Website. (2018). Retrieved September 05, 2018 from https://make.figure-eight.com/choose_login

[3] 2018. Drinking culture. (2018). Retrieved April 16, 2018 from https://en.wikipedia.org/wiki/Drinking_culture

[4] 2018. Face API on Microsoft Azure. (2018). Retrieved April 12, 2018 from https://docs.microsoft.com/en-us/azure/cognitive-services/face/overview

[5] 2018. Foursquare Venue Categories. (2018). Retrieved April 12, 2018 from https://developer.foursquare.com/categorytree

[6] 2018. Sklearn with GridSearchCV. (2018). Retrieved September 05, 2018 from http://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html

[7] Sofiane Abbar, Yelena Mejova, and Ingmar Weber. 2015. You tweet what you eat: Studying food consumption through twitter. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3197–3206.

[8] Barbro Andersson, Björn Hibell, François Beck, Marie Choquet, Anna Kokkevi, Anastasios Fotiou, Sabrina Molinaro, Alojz Nociar, Januzs Sieroslawski, and Marcis Trapencieris. 2007. Alcohol and drug use among European 17–18 year old students. *Data from the ESPAD Project. The Swedish Council for Information on Alcohol and Other Drugs (CAN) and the Pompidou Group at the Council of Europe* (2007).

[9] Zachary Arnold, Danielle Larose, and Emmanuel Agu. 2015. Smartphone inference of alcohol consumption levels from gait. In *Healthcare Informatics (ICHI), 2015 International Conference on*. IEEE, 417–426.

[10] Amanda Marie Atkinson and Harry Robin Sumnall. 2016. 'If I don't look good, it just doesn't go up': A qualitative study of young women's drinking cultures and practices on Social Network Sites. *International Journal of Drug Policy* 38 (2016), 50–62.

[11] Sangwon Bae, Denzil Ferreira, Brian Suffoletto, Juan C Puyana, Ryan Kurtz, Tammy Chung, and Anind K Dey. 2017. Detecting Drinking Episodes in Young Adults Using Smartphone-based Sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 2 (2017), 5.

[12] Saeideh Bakhshi, David A Shamma, and Eric Gilbert. 2014. Faces engage us: Photos with faces attract more likes and comments on instagram. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 965–974.

[13] Joan-Isaac Biel, Nathalie Martin, David Labbe, and Daniel Gatica-Perez. 2017. BitesnBits: Inferring Eating Behavior from Contextual Mobile Data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2017), 125.

[14] Sarah C Boyle, Andrew M Earle, Joseph W LaBrie, and Kayla Ballou. 2017. Facebook dethroned: Revealing the more likely social media destinations for college students' depictions of underage drinking. *Addictive behaviors* 65 (2017), 63–67.

[15] Paul Chatterton and Robert Hollands. 2003. *Urban nightscapes: Youth cultures, pleasure spaces and corporate power*. Vol. 18. Psychology Press.

[16] Yohan Chon, Nicholas D Lane, Fan Li, Hojung Cha, and Feng Zhao. 2012. Automatically characterizing places with opportunistic crowdsensing using smartphones. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 481–490.

[17] Chia-Fang Chung, Elena Agapie, Jessica Schroeder, Sonali Mishra, James Fogarty, and Sean A. Munson. 2017. When Personal Tracking Becomes Social: Examining the Use of Instagram for Healthy Eating. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1674–1687. https://doi.org/10.1145/3025453.3025747

[18] M Lynne Cooper. 1994. Motivations for alcohol use among adolescents: Development and validation of a four-factor model. *Psychological assessment* 6, 2 (1994), 117.

[19] M Lynne Cooper, Marcia Russell, Jeremy B Skinner, and Michael Windle. 1992. Development and validation of a three-dimensional measure of drinking motives. *Psychological assessment* 4, 2 (1992), 123.

[20] W Miles Cox. 1990. *Why people drink: Parameters of alcohol as a reinforcer*. Gardner Press.

[21] W Miles Cox and Eric Klinger. 1988. A motivational model of alcohol use. *Journal of abnormal psychology* 97, 2 (1988), 168.

[22] Munmun De Choudhury, Mrinal Kumar, and Ingmar Weber. 2017. Computational Approaches Toward Integrating Quantified Self Sensing and Social Media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 1334–1349. https://doi.org/10.1145/2998181.2998219

[23] Jakob Demant and Sara Landolt. 2014. Youth drinking in public places: The production of drinking spaces in and outside nightlife areas. *Urban Studies* 51, 1 (2014), 170–184.

[24] Ola Ekholm. 2004. Influence of the recall period on self-reported alcohol intake. *European Journal of Clinical Nutrition* 58, 1 (2004), 60.

[25] Nicole Ellison, Rebecca Heino, and Jennifer Gibbs. 2006. Managing impressions online: Self-presentation processes in the online dating environment. *Journal of computer-mediated communication* 11, 2 (2006), 415–441.

[26] Hao Fang, Saurabh Gupta, Forrest Iandola, Rupesh Srivastava, Li Deng, Piotr Dollár, Jianfeng Gao, Xiaodong He, Margaret Mitchell, John Platt, et al. 2015. From captions to visual concepts and back. (2015).

[27] Gerhard Gmel, Jürgen Rehm, Emmanuel Kuntsche, et al. 2003. Binge drinking in Europe: definitions, epidemiology, and consequences. *Sucht* 49, 2 (2003), 105–116.

[28] Erving Goffman. 1959. The Presentation of Self in. *Butler, Bodies that Matter* (1959).

[29] Yandong Guo, Lei Zhang, Yuxiao Hu, Xiaodong He, and Jianfeng Gao. 2016. Ms-celeb-1m: Challenge of recognizing one million celebrities

in the real world. *Electronic Imaging* 2016, 11 (2016), 1–6.

[30] David Hawks, Katie Scott, Nyanda McBride, Paul Jones, Tim Stockwell, World Health Organization, et al. 2002. Prevention of psychoactive substance use: A selected review of what works in the area of prevention. (2002).

[31] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[32] Hanneke Hendriks, Bas Van den Putte, Winifred A Gebhardt, and Megan A Moreno. 2018. Social Drinking on Social Media: Content Analysis of the Social Aspects of Alcohol-Related Posts on Facebook and Instagram. *Journal of Medical Internet Research* 20, 6 (2018), e226.

[33] Nabil Hossain, Tianran Hu, Roghayeh Feizi, Ann Marie White, Jiebo Luo, and Henry Kautz. 2016. Inferring fine-grained details on user activities and home location from social media: Detecting drinking-while-tweeting patterns in communities. *arXiv preprint arXiv:1603.03181* (2016).

[34] Jin Yea Jang, Kyungsik Han, Patrick C Shih, and Dongwon Lee. 2015. Generation like: comparative characteristics in Instagram. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 4039–4042.

[35] Hope Jensen Schau and Mary C Gilly. 2003. We are what we post? Self-presentation in personal web space. *Journal of consumer research* 30, 3 (2003), 385–404.

[36] Lloyd Johnston. 2010. *Monitoring the future: National results on adolescent drug use: Overview of key findings.* Diane Publishing.

[37] Sylvia Kairouz, Louis Gliksman, Andrée Demers, and Edward M Adlaf. 2002. For all these reasons, I do... drink: a multilevel analysis of contextual reasons for drinking among Canadian undergraduates. *Journal of studies on alcohol* 63, 5 (2002), 600–608.

[38] Hsin-Liu Cindy Kao, Bo-Jhang Ho, Allan C Lin, and Hao-Hua Chu. 2012. Phone-based gait analysis to detect alcohol usage. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 661–662.

[39] Daniel Kershaw, Matthew Rowe, and Patrick Stacey. 2014. Towards Tracking and Analysing Regional Alcohol Consumption Patterns in the UK Through the Use of Social Media. In *Proceedings of the 2014 ACM Conference on Web Science (WebSci '14)*. ACM, New York, NY, USA, 220–228. https://doi.org/10.1145/2615569.2615678

[40] Terry K Koo and Mae Y Li. 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine* 15, 2 (2016), 155–163.

[41] Emmanuel Kuntsche, Ronald Knibbe, Gerhard Gmel, and Rutger Engels. 2005. Why do young people drink? A review of drinking motives. *Clinical psychology review* 25, 7 (2005), 841–861.

[42] Emmanuel Kuntsche and Sandra Kuntsche. 2009. Development and validation of the drinking motive questionnaire revised short form (DMQ–R SF). *Journal of Clinical Child & Adolescent Psychology* 38, 6 (2009), 899–908.

[43] Emmanuel Kuntsche and Florian Labhart. 2012. Investigating the drinking patterns of young people over the course of the evening at weekends. *Drug and alcohol dependence* 124, 3 (2012), 319–324.

[44] Emmanuel Kuntsche and Florian Labhart. 2013. ICAT: development of an internet-based data collection method for ecological momentary assessment using personal cell phones. *European Journal of Psychological Assessment* (2013).

[45] Dong Kyu Lee. 2016. Alternatives to P value: confidence interval and effect size. *Korean journal of anesthesiology* 69, 6 (2016), 555–562.

[46] Andy Liaw, Matthew Wiener, et al. 2002. Classification and regression by randomForest. *R news* 2, 3 (2002), 18–22.

[47] Sean Esteban McCabe. 2002. Gender differences in collegiate risk factors for heavy episodic drinking. *Journal of Studies on Alcohol* 63, 1 (2002), 49–56.

[48] Yelena Mejova, Hamed Haddadi, Anastasios Noulas, and Ingmar Weber. 2015. # foodporn: Obesity patterns in culinary interactions. In *Proceedings of the 5th International Conference on Digital Health 2015*. ACM, 51–58.

[49] Aehong Min, Daehyoung Lee, and Patrick C Shih. 2018. Potentials of Smart Breathalyzer: Interventions for Excessive Drinking Among College Students. In *International Conference on Information*. Springer, 195–206.

[50] Jang Ho Moon, Eunji Lee, Jung-Ah Lee, Tae Rang Choi, and Yongjun Sung. 2016. The role of narcissism in self-promotion on Instagram. *Personality and Individual Differences* 101 (2016), 22–25.

[51] Ran Pang, Agustin Baretto, Henry Kautz, and Jiebo Luo. 2015. Monitoring adolescent alcohol use via multimodal analysis in social multimedia. In *Big Data (Big Data), 2015 IEEE International Conference on*. IEEE, 1509–1518.

[52] John Petraitis, Brian R Flay, and Todd Q Miller. 1995. Reviewing theories of adolescent substance use: organizing pieces in the puzzle. *Psychological bulletin* 117, 1 (1995), 67.

[53] Thanh-Trung Phan and Daniel Gatica-Perez. 2017. Healthy# fondue# dinner: analysis and inference of food and drink consumption patterns on instagram. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*. ACM, 327–338.

[54] Andrew G Reece and Christopher M Danforth. 2017. Instagram photos reveal predictive markers of depression. *EPJ Data Science* 6, 1 (2017), 15.

[55] Jürgen Rehm, Robin Room, Kathryn Graham, Maristela Monteiro, Gerhard Gmel, and Christopher T Sempos. 2003. The relationship of average volume of alcohol consumption and patterns of drinking to burden of disease: an overview. *Addiction* 98, 9 (2003), 1209–1228.

[56] Darshan Santani, Joan-Isaac Biel, Florian Labhart, Jasmine Truong, Sara Landolt, Emmanuel Kuntsche, and Daniel Gatica-Perez. 2016. The night is young: urban crowdsourcing of nightlife patterns. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 427–438.

[57] Darshan Santani, Trinh-Minh-Tri Do, Florian Labhart, Sara Landolt, Emmanuel Kuntsche, and Daniel Gatica-Perez. 2017. DrinkSense: Characterizing Youth Drinking Behavior using Smartphones. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM.

[58] Barry R Schlenker. 2012. Self-presentation. (2012).

[59] Anil J Shrikhande. 2000. Wine by-products with health benefits. *Food Research International* 33, 6 (2000), 469–474.

[60] Flávio Souza, Diego de Las Casas, Vinícius Flores, SunBum Youn, Meeyoung Cha, Daniele Quercia, and Virgílio Almeida. 2015. Dawn of the selfie era: The whos, wheres, and hows of selfies on Instagram. In *Proceedings of the 2015 ACM on conference on online social networks*. ACM, 221–231.

[61] Bruce Taylor, HM Irving, Fotis Kanteres, Robin Room, Guilherme Borges, Cheryl Cherpitel, Thomas Greenfield, and Jürgen Rehm. 2010. The more you drink, the harder you fall: a systematic review and meta-analysis of how acute alcohol consumption and injury or collision risk increase together. *Drug & Alcohol Dependence* 110, 1 (2010), 108–116.

[62] Tingxin Yan, Matt Marzilli, Ryan Holmes, Deepak Ganesan, and Mark Corner. 2009. mCrowd: a platform for mobile crowdsourcing. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*. ACM, 347–348.

[63] Svetlana Yarosh. 2013. Shifting dynamics or breaking sacred traditions?: the role of technology in twelve-step fellowships. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3413–3422.

[64] Koji Yatani. 2016. Effect Sizes and Power Analysis in HCI. In *Modern Statistical Methods for HCI*. Springer, 87–110.