Risky drinking social worlds in Victoria: A social media analysis

Authors
Dr Ashir Ahmed
Professor Jennifer Martin
Mr David Towl
Mr Zac Haussegger

This research was supported by Access Health and Community
Acknowledgements
We would like to thank Access Health and Community for sponsoring this research, especially Mr David Towl for his helpful advice, contribution and support. We also thank the survey respondents who provided valuable information and helped shape this report.

Disclaimers
This report does not constitute Victorian Government policy. The content of this publication is solely the responsibility of the authors and does not necessarily represent the views of the Victorian Government.

Authors’ affiliation
Dr Ashir Ahmed  
Professor Jennifer Martin  
Mr David Towl  
Mr Zac Haussegger  

Swinburne University of Technology  
Federation University  
Access Health and Community  
Swinburne University of Technology

Publication design by Diane Yuri

Copyright © The authors 2021.  
Published by Swinburne University of Technology and Access Health & Community.

Except where otherwise noted, this work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

ISBN 978-1-925761-08-5 (PDF)  
ISBN 978-1-925761-34-4 (Print)  
DOI https://doi.org/10.25916/y690-8y89
Contents

Executive summary 5
Risky drinking 6
Social Worlds 7
Talkwalker 9
Study design 11
Key findings 14
  Gender 14
  Age 15
  Family status 16
  Occupation 16
  Social interests 17
  Location 18
  Devices 19
  Conversations 20
  Sentiment 21
Conclusion 22
Recommendations 23
References 25
List of Tables 27
List of Figures 27
The aim of the project is to answer the question of: “What are the social worlds of people engaging in risky drinking behaviours in Victoria?”
Executive summary

Alcohol use is at the forefront of public health policy in Australia. Not only does chronic use increase the risk of cardio-metabolic disease, cancer, liver disease and mental health disorders, intoxication in young people can lead to dangerous behaviour and increase the risk of unintentional injury and death (Griswald et al., 2018). On a global scale, alcohol use was the seventh leading risk factor for disability-affected-life-years (DALY) and deaths in 2016, with Australia in the top bracket – over 80% – of current drinking prevalence in its population (Griswald et al., 2018). In 2015, alcohol use contributed to 4.5% of the total burden of disease in Australia (AIHW, 2019). In an effort to address risky alcohol use, the Victorian Health Promotion Foundation (VicHealth) has identified “preventing harm from alcohol” as one of five strategic priorities in its Strategic Plan 2013-2023 (VicHealth, 2013 p. 23).

Developing an effective intervention requires a better and improved understanding of the social worlds of risky drinking behaviour. Hence, this project makes use of social media data to examine the social worlds of drinking behaviour in Victoria extending the current understanding of people involved in online conversation regarding risky drinking. The aim of the project is to answer the question of: “What are the social worlds of people engaging in risky drinking behaviours in Victoria?”

The current study plans to gain more detailed and nuanced data from social media to gain an increased depth of understanding of the social worlds where this risky drinking conversation and behaviour occurs. In addition, the findings of this report provide a catalyst for further in-depth and longitudinal analysis of qualitative and quantitative social media data to inform policy makers develop more effective interventions to deal with the problem of risky drinking.
Risky drinking

Alcohol use is a widespread issue affecting people of different gender, age, social and professional status and geographical location. High levels of alcohol consumption can lead to serious health and social issues. In Australia, 1 in 6 people (17%) consume alcohol at levels that place them at an increased lifetime risk of alcohol-related disease or injury (AIHW, 2019).

The National Health and Medical Research Council of Australia (NHMRC) recognises the broad role that alcohol has on the national burden of disease, as it considers its immediate effects (namely intoxication and the risk of injury), the cumulative effect on mental health, the risk of cardiovascular and metabolic disease, and the social effects of alcohol use such as family disruption, violence, crime, road accidents and work-related harm.

The current study relies upon the following NHMRC definition of ‘risky drinking’. Risky drinking is

‘...drinking more than 10 standard drinks per week, or more than 4 standard drinks on any one day. Drinking in this way can lead to both these immediate and long-term adverse effects of alcohol use.’ (NHMRC, 2019, p. 1)

The characteristics of people engaged in online group conversations about drinking alcohol, referred to as ‘social worlds’, was examined using social media data analysis.
Social Worlds

Social worlds is a term used in social practice theory to describe settings and cultures with shared understandings, values and behaviours.

For the purpose of this study an alcohol culture or ‘social world’ is:

“The way a group of people drinks, including the shared understanding of formal rules, social norms, practices, values and beliefs around what is and what is not socially acceptable when members of the group get together.” (VicHealth and La Trobe University’s Centre for Alcohol Policy Research (CAPR), 2019, p. 6)

Social media data analysis has demonstrated ability to provide information on the characteristics of users, such as who they are, where they are, personal preferences and what they discuss online. Generally, much of the statistical inferences within social media use is utilised commercially and this is evident in the types of sources that display this data. Freely available public statistics are provided by digital analytics websites such as datareportal.com with a reported 3 billion plus active users of social media worldwide (Datareportal, 2020). Social media analysis is also conducted for targeted marketing of products and services with Talkwalker a main platform. An innovation of this study is the extension of the application of Talkwalker to social worlds. Further details of Talkwalker are provided below.

Social media research has focused on the various ways people interact and behave online. In particular, there is a known intersection between health and social media that is currently under-utilised and requiring further research (Velasco et al., 2014; Zhao & Zhang, 2017). User-generated alcohol-related posts by youth, across a variety of social media platforms, has been linked to self-reported problems with alcohol (Curtis et al., 2018). In 2016, a study by VicHealth and La Trobe University identified at risk populations within the construction industry (VicHealth & CAPR, 2017a), people in rural and regional areas (VicHealth & CAPR, 2017b), sports bar settings (VicHealth & CAPR, 2017c); and more recently, middle aged same sex attracted female as being groups of particular interest for risky drinking cultures (MacLean et al., 2019). There is an element of interconnectedness that is shared between the nature of social networking and the communal aspects of drinking cultures that is further investigated in the current study.
The shift in policy, transitioning away from the individual and towards the social world, comes from within social practice theory, as governments seek new ways of achieving their public health goals.

This is evidenced by Meier, Warde and Holmes’s (2018) study being used as the basis for VicHealth and CAPR’s Alcohol Cultures Framework (2019), which argued that previously held individualist explanations of drinking behaviour have clouded the academic understanding of behaviours that cannot be accounted for without considering the diverse, context specific practices that go along with them. The literature highlighted this missing link and stresses that more research should be conducted into investigating this causal relationship and exploring the possibility of locating risky drinkers via these user-generated posts about alcohol. The current study was designed to address the gap by exploring the ability of social data analytics to locate the social worlds of risky drinkers.

“

There is an element of interconnectedness that is shared between the nature of social networking and the communal aspects of drinking cultures that is further investigated in the current study.

“
Talkwalker

Talkwalker is a social media analytics platform that uses artificial intelligence to monitor and analyse online conversations in real-time across social networks, news websites, blogs and forums in number of different languages.

The role that Talkwalker has in this study is to take the search terms provided by the surveys (as well as a literature review conducted by the research team) and identify the social worlds of users who use those terms when engaging in interactions online. In addition to the qualitative social media data such as posts, tweets and comments, Talkwalker examines various dimensions of users’ demographics involved in online communication.

Talkwalker monitoring works by allowing the execution of a search of all publicly available data on various social media platforms, and returns all items that match its search criteria. These criteria can be set not only with a specific search term query but also additional preferences such as timeframe, region, language and media and use of a particular device type. Talkwalker can return the additional information of the user’s demographics, sentiment of the post, and themes alongside each data point it returns.

The collection of data by Talkwalker is governed by the privacy policy of the social media platform the data exists within. Talkwalker has the ability to analyse a number of parameters pertaining to a user’s profile when a post is found that matches the keywords. Details such as location, age, occupation, gender and interests can either be found explicitly on the user’s public profile or be implied from their “bio” or pages that they follow. These demographic dimensions provide insights on gender, age, social and family status, interests and location of those who constitute the social worlds of risky drinking culture.

Importantly, when analysing Talkwalker’s global summaries of the data, each parameter may have a different sample size. This is because on some users’ profiles the gender of the user may be available (either explicitly or implicitly), and on others, it won’t. At the same time, the occupation of the person might be available for the user whose gender could not be obtained, but could not for the user whose gender was recorded.
We were working within the parameters of the functionality that was offered by Talkwalker and note that a number of the functions, such as attribution of gender, are controversial and this will be discussed further in a separate publication. The following findings must be considered in light of these functionality features.

Talkwalker monitoring works by allowing the execution of a search of all publicly available data on various social media platforms, and returns all items that match its search criteria.
Study design

The study aim is to provide useful insights into how people engage in conversations regarding drinking behaviour online. Twitter, Facebook and Reddit were selected to search for relevant data as their privacy policies allowed access to relevant types and amounts of historical data. The licensing agreements of data monitoring on different social media and university ethics guidelines allowed for the collection of data on public social media platforms only (Ethics approval: Ref: 20202944-4889). As a result, data from closed-groups and private forums is not included.

The study question is “What are the social worlds of people engaging in risky drinking behaviours in Victoria?”

The study location is the state of Victoria, recognising that people in different geographic locations have varying relationships with alcohol and ‘social worlds of drinking culture’. Social worlds that have an aspect of risky drinking behaviour are the target of this study, as increased understanding of drinking cultures will better inform successful interventions in the future (MacLean et al., 2019).

We use age, gender, occupation, interests and location to uncover different risky drinking cultures within Victoria using social media data analytics.

The initial task was to identify key search terms and hashtags used in risky drinking conversations on social media that people use when referring to alcohol on social media. This initial task involved two components. First a scoping review of the literature was conducted. This was followed by two online surveys conducted between August and September 2020 with: (1) the general population and (2) health and human service clinicians working with risky drinkers. Details of the literature review and survey will be published elsewhere. Both surveys were distributed using Qualtrics software.

The role of scoping literature review is to inform policy decision makers when circumstances dictate a need for information in a timely manner (Langlois et al., 2017). Furthermore, scoping reviews have been shown to
play a role when organisational and commissioning bodies see that a more succinct and timely process suits their own needs (Polisena et al., 2015). For these reasons, a scoping literature review of peer reviewed articles was chosen as the most appropriate method. Most of the literature on this topic was confined to the last 10 years with over 50% of the included studies were conducted with adolescents and young people at university.

The general population online survey was decided upon as it casts a wide net to uncover individuals engaging in risky drinking that a targeted population group might have missed. This could be due to a misunderstanding of what risky drinking actually means. It is indeed possible that people engaging in behaviours that match the definition in this study might not consider themselves ‘risky drinkers’. It is well known that assigning labels to people in this way can lead to denial and, in these circumstances, declining participation. Online survey invitations were distributed to local community groups within Victoria such as sporting clubs, early childhood centres and schools with 7 respondents.

The inclusion of the worker population worker survey via networks of alcohol and other drug services, social workers, general practitioners and counsellors aimed to also collect key search terms as well as providing another perspective on risky drinkers. Key questions were asked that allowed the investigators to formulate an idea of the types of drinking cultures practitioners thought were most at risk. Four workers responded.

The results from the online surveys about possible search terms and hashtags were compiled for analysis in Talkwalker. A total of 41 search terms were generated for Talkwalker analysis. Each of these terms was mentioned on multiple occasions in the literature review and survey data. In addition to the search terms, there are Talkwalker specific filters that direct the search to only take in results from Victoria, Australia from Facebook, Twitter, or Reddit and are in English.

Following the collection of the data, measures were taken to ensure that posts were relevant to the project aim. Two random samples of 200 posts were taken from the entire dataset that was generated by the initial topic query from the 2-month data collection period from October 2020 to December 2020. Each sample was assessed by a member of the investigative team and their recommendations for the adjustment of the query were tested and then implemented across the entire dataset.
Recommendations were based on the particular post’s relevance to the study question on risky drinking. Words or phrases that were commonly used with particular search terms outside of a drinking context were excluded during the data cleaning process, resulting in a final set of 29 search terms.

During data cleaning, analysis focused on whether or not the post was referring to drinking when it used the search term. This process was repeated a second time on the resulting, now smaller, dataset, with the refinements implemented from the first iteration. In this instance, members of the investigative team double handled the same sample of 200. This second data cleaning enabled the team to reach a saturation point whereby no new information was being found towards the end of this sample. In this sample, greater than 80% of posts were deemed to be relating to drinking behaviours by each investigator and the data cleaning process ended.

“The inclusion of the worker population worker survey via networks of alcohol and other drug services, social workers, general practitioners and counsellors aimed to also collect key search terms as well as providing another perspective on risky drinkers.”
Key findings

Talkwalker collected 4,714 results from social media over the 2-month period from Oct 15 to Dec 15 2020. After the data cleaning process, the refined dataset was reduced to 1,021 results. Of these results, greater than 99% were found on Twitter and the rest on Reddit. The final search terms are detailed on Table 1.

<table>
<thead>
<tr>
<th>#afterworkdrinks</th>
<th>blotto</th>
<th>hair of the dog</th>
<th>tipsy</th>
</tr>
</thead>
<tbody>
<tr>
<td>#cheers</td>
<td>#liftyourspirit</td>
<td>zoom drink</td>
<td>virtual drink</td>
</tr>
<tr>
<td>#cheerstotheweekend</td>
<td>#pinotgrigio</td>
<td>happy hour</td>
<td>it’s 5 o’clock somewhere</td>
</tr>
<tr>
<td>#cocktailhour</td>
<td>#pubcrawl</td>
<td>inebriated</td>
<td>legless</td>
</tr>
<tr>
<td>#drinkdrankdrunk</td>
<td>#quarantini</td>
<td>intoxicated</td>
<td>wine not</td>
</tr>
<tr>
<td>foursies</td>
<td>#saturdaynight</td>
<td>pissed</td>
<td>wine o’clock</td>
</tr>
<tr>
<td>#drunk</td>
<td>#whynot</td>
<td>shitfaced</td>
<td>wine time / winetime</td>
</tr>
</tbody>
</table>

Table 1: Talkwalker search terms

Gender

According to the National Alcohol and Drug Knowledgebase (NADK), Australian men usually drink more alcohol than women on a single occasion of drinking (NADK, 2021). Women are more likely than men to drink 1-2 standard drinks per occasion, and are also more likely to be abstainers. On a single occasion of drinking, 21% of Australian men abstain from alcohol consumption, compared to 26% of women.
The study findings indicate that the gender representations are in line with the findings from the NADK, as shown in Figure 1. It is interesting to note that Talkwalker applies ‘male’ and ‘female’ as the only two classifications for gender. People who identify as a non-binary gender or who do not want to disclose their gender identity have not been represented in the social media data.

Age

As seen in Figure 2, the most common age-group was 25-34 years, with almost half of users who posted in this age group. The second most common age-group was 18 and 24 years. Since the legal age to drink in Australia is 18 years, this age group was less likely to have representation. According to the social media data analysis, senior people (age 65+) had least representation. It is important to note that the social media use of this age group is the lowest (47%) of all age groups in Australia (Sensis, 2017). Therefore, social media uptake by age group needs to be considered when interpreting this kind of data.
Family status

‘Parents’ emerged as the most common group of family status and represents 66.4% of the collected data. Based on the family status, the next in line is ‘married’ (29.3%), whereas ‘senior’ is the least commonly mentioned status. These findings on family status are aligned with what has been discussed in the age groups in the previous section and presented in Figure 3.

![Figure 3: Risky drinking and family status](image1)

Occupation

The occupation of the members of social worlds is an important piece of information that reflects on what these people do for a living and helps to understand their social status. The social media analysis suggests that the social worlds of risky drinking behaviours constitutes a range of occupations with ‘authors/writers’ at the top of the list, whereas people who classified themselves as ‘scientist’ are at the end with 3.2% (see Figure 4). However, a very large number of people participating in online conversations have not mentioned their occupations specifically, leaving some ambiguity in this field.

![Figure 4: Risky drinking and occupation](image2)
Social interests

In line with the top occupation as ‘author/writer’, literature/books emerged as the most popular interest among social media users included in the project as seen in Figure 5. Although these finding are surprising, they are mutually inclusive. Similar to the data findings on occupations, a very large number of people (44.9%) have not mentioned their interests specifically, which leaves some ambiguity about the interest of this large unclassified cohort.

![Figure 5: Risky drinking and social interests](image-url)
The geographic boundary for data collection for this project was set as ‘Victoria, Australia’. Therefore, all the posts matching the search terms that were either published from within Victoria or where the author has set their location within Victoria were included. In terms of a rural and urban divide, the urban areas of Victoria significantly outnumbered rural localities. As shown in Figure 6, social media data analysis found that 89% of the social media posts matching the selection criteria were generated from the Melbourne area with the remaining 11% including a mix of rural, regional and urban areas beyond the capital city.
Devices

The national trend of using social media by various handheld devices was also reflected in the study findings (Sensis, 2017). Accessing social media using mobile phones was the key preference of the study population. Data analysis of the devices used for social media suggests that the vast majority of people use various types of mobile devices including iPhone, androids, iOS tablets and MS Windows (see Figure 7). The large number of people using handheld devices implies there is more spontaneity and speed in sharing content as compared to people using desktop computers. Moreover, handheld devices are generally more responsive to GPS location services and provide more specific information to the users based on their changing position as compared to stationary devices such as a desktop computer.

Figure 7: Devices used by risky drinkers
Conversations

In excess of 99% of the social media posts were on Twitter. All social media posts included in the analysis match the search terms and other criteria such as location in Victoria and English language. In-depth analysis of these posts was conducted to identify patterns of words, terms and phrases that are commonly used in online conversations.

Figure 8 shows the frequency of the most popular words and emojis that were used in the posts. The bigger the size of the word, the higher the frequency of use in the online conversations.

In addition to the popularity of various words, we also examined the string of words to identify the most popular two-word strings used online (See Figure 10).
Sentiment

The analysis of sentiment expressed in the posts revealed less than half of the posts indicated positive sentiment in relation to drinking alcohol (see Figure 10). Further exploration of sentiment is promising, particularly as datasets get larger, to reveal risky drinking behaviours.

Figure 10: Risky drinking and sentiment
Summary

The study findings provide a global perspective on the data as a whole, and is the closest we can get to generating archetypes of particular people talking about a topic such as risky drinking.

According to the study findings, the most likely social world engaging in risky drinking behaviour is 25-34-year-old males, who are authors or writers, are parents, and enjoy literature and books.

Of course, as discussed above, there was no one group as well defined as this. It is considered the most likely as these were the most popular categories. Reading the results in this way points to a number of confounding variables within the study. These included the fact that more than half of the included articles in the scoping review targeted adolescent and undergraduate university students, and that there are larger numbers of younger people than older people who are active on social media. There seems to be a disproportionately higher rate of authors and writers within our occupation findings in comparison to the general population. This may be an artefact of Twitter rather than Victoria’s risky drinking population. As more than 99% of the posts collected are from Twitter, it can be said that any of the demographic data collected is representative of Twitter users in Victoria who engage in risky drinking rather than, more broadly, Victorians who engage in risky drinking.

Using social media data to find people engaging in risky drinking behaviours is promising. In saying this, there is limited confidence in providing the social world of 25-34-year-old males, who are authors or writers, are parents, and enjoy literature and books, as the target of Victoria’s next alcohol intervention strategy. The project does indicate, however, that in a world where social media is at the front of corporate consumer marketing strategy, it too can be at the front of public health intervention and strategy. The project has demonstrated that discussions around drinking do take place online, and there is capacity within sophisticated social media data analytics to find them.
Recommendations

**Recommendation 1:**
Conduct a qualitative data analysis on the raw data for a more in-depth understanding of the drinking cultures and risky drinking behaviours in Victoria. If possible, this would be the gold standard for this sort of project.

**Recommendation 2:**
Conduct a larger scale search within Victoria using the methods adopted in this study incorporating the technical learnings gained in the process.

**Recommendation 3:**
Be clear about which social worlds are being targeted, and aim to utilise a more effective strategy for monitoring Facebook and other sites as deemed appropriate.

**Recommendation 4:**
Adopt machine learning to fine tune the sentiment algorithm within Talkwalker to identify risky drinking behaviours.
The project has demonstrated that discussions around drinking do take place online, and there is capacity within sophisticated social media data analytics to find them.
References


List of Tables

Table 1: Talkwalker search terms  Page 14

List of Figures

Figure 1: Risky drinking and gender  Page 15
Figure 2: Risky drinking age groups  Page 15
Figure 3: Risky drinking and family status  Page 16
Figure 4: Risky drinking and occupation  Page 16
Figure 5: Risky drinking and social interests  Page 17
Figure 6: Location of risky drinking cultures  Page 18
Figure 7: Devices used by risky drinkers  Page 19
Figure 8: Most popular words arising from posts  Page 20
Figure 9: Most popular themes arising from posts  Page 20
Figure 10: Risky drinking and sentiment  Page 21