A Smartphone App to Assess Alcohol Consumption Behaviours: Development, Compliance, and Reactivity

Antoinette Poulton¹, MEd; Jason Pan¹; Loren Richard Bruns Jr², PhD; Richard O Sinnott², PhD; Robert Hester¹, PhD

¹Melbourne School of Psychological Sciences, University of Melbourne, Parkville 3010, VIC, Australia
²Computing and Information Systems, University of Melbourne, Parkville 3010, VIC, Australia

Corresponding author:
Antoinette Poulton
Melbourne School of Psychological Sciences,
University of Melbourne,
Parkville 3010, VIC, Australia
Tel.: +61 3 8344 6377
Fax: +61 3 9347 6618
antoinette.poulton@unimelb.edu.au
poultonantoinette@gmail.com
Abstract

Background: There are disadvantages – largely related to cost, participant burden, and missing data – associated with traditional electronic methods of assessing drinking behaviour in real-time. Smartphone apps participants download to their own phones might minimise some of these disadvantages. To date, however, few researchers have detailed the process involved in developing custom-built apps for use in the experimental arena or explored methodological concerns regarding compliance and reactivity.

Objective: This paper describes the process used to guide development of a custom-built smartphone app designed to capture alcohol intake behaviour in the healthy population. Methodological issues related to compliance with, and reactivity to, study protocols, as a function of hazard/non-hazard drinker status are also explored.

Methods: An iterative development process that included elements typical of agile software design guided the creation of the CNLab-A app. Healthy individuals used the app to record alcohol consumption behaviour each day for 21 days. Submissions were either event- or notification-contingent. We considered the size and diversity of the sample plus assessed the data for evidence of app protocol compliance and reactivity as a function of hazard/non-hazard drinker status.

Results: CNLab-A yielded a large and diverse sample ($N = 671, M_{age} = 23.12$). On average, participants submitted data on 20.27 ($SD = 1.88$) days. Both hazard and non-hazard drinkers were highly compliant with app protocols. Linear growth analyses revealed hazardous drinkers decreased their alcohol intake by 0.80 standard drinks over the 21-day experimental period. There was no change to the drinking of non-hazard individuals.
Conclusions: Smartphone apps participants download to their own phones are an effective and methodologically sound means of obtaining alcohol consumption information for research purposes. While further investigation is required, such apps might, in future, allow for a more thorough examination of the antecedents and consequences of drinking behaviour.

Keywords Alcohol; Smartphone apps; Research app development; Compliance; Reactivity
Introduction

Advantages associated with real-time (or near real-time) methods of assessing alcohol consumption behaviour for research purposes have been widely documented in recent years [1–3]. Such methods – which are increasingly electronic in nature – are advocated on the basis they allow data to be captured repeatedly, in the natural environment, and in the absence of the researcher [2,3]. This facilitates the collection of actual intake information rather than the summary data commonly elicited from more traditional retrospective methods of assessing drinking. As such, recall and response biases are thought minimised and the ecological validity of the data enhanced [3]. Real-time methods additionally reduce the quantity of missing data and can yield a diverse and potentially very large sample [4,5]. Crucially, real-time data enables variations in behaviour to be examined across time and in concert with cognitive, affective, environmental, and/or physiological factors; in this way, the antecedents and consequences of drinking behaviour can also be investigated [2,3].

Real-time electronic methods of collecting alcohol consumption information for research purposes have evolved rapidly over the last two decades. Early studies typically featured hand-held electronic devices [6,7] or interactive voice response systems [8,9]. In the case of the former, participants used the device as an electronic diary; in the later, the interactive voice response system was programmed to call participants on researcher-supplied cellular phones multiple times per day so an automated questionnaire could be administered. Text messaging protocols – using either researcher-supplied devices or participants’ mobiles – have also been employed. In such studies, text messages direct participants to follow a link in order to complete an online survey via their mobile phone browser [10], or ask them to respond to
simple questions about their alcohol consumption via text [11]. With the advent of smartphones, researchers have, increasingly, been programming study-specific applications (apps) for use in such studies. In most cases, this involves either providing participants with a phone preloaded with the app [12,13], or loading the app onto participants’ own phones and downloading the data at end of the experimental period [14].

There are a number of disadvantages associated with utilising the aforementioned electronic protocols. There are costs, for instance, associated with programming, supplying, and training participants to use electronic devices that are not their own [15–18]. Participants must also be provided with some means whereby they can recharge these devices [16]. If they have their own mobile, they will be carrying two devices during the experimental period, which might prove unduly burdensome and reduce compliance. Questionnaires completed via cell phone browsers do not always scale well to all mobile devices and require Internet connectivity [19]. Surveys conducted via text message are necessarily limited in scope and, in the absence of Internet connectivity, potentially costly to participants [20]. In all cases, participants are required to visit the lab at least twice during the experimental period: they must collect the device or have the program loaded onto their own phone plus undergo training and they must return the device or have the data downloaded off their phone [18,21]. These disadvantages potentially diminish some of the benefits of using electronic devices to collect real-time alcohol intake behaviour; specifically, sample size and diversity may be a function of the cost of supplying the device to those willing and able to attend the lab, and data may go unrecorded where Internet connectivity or cost to the participant is an issue.
Using apps participants download to their own smartphones – without ever visiting a lab – might enable researchers to more fully realize the benefits of collecting alcohol consumption information electronically and in real time. Smartphone penetration currently stands at upward of 70% across many developed nations and is growing rapidly across developing ones [22]. Across the United States, United Kingdom and Australia, more than 93% of 18 to 34 year olds own a smartphone [23]. Taking advantage of high smartphone ownership rates by using participants’ own devices represents a substantial cost reduction both in terms of equipment and training. Asking individuals to download apps via marketplace vendors and having data stored on the phone until it can be automatically uploaded to a server via a Wi-Fi connection, alleviates participants of the need to visit the lab and reduces the likelihood data will be lost or go unrecorded. Any necessary protocol training can be conducted via digital means (e.g., video, embedded in the app). Advantages of real-time assessment related to completeness of data, sample size and diversity might therefore be preserved in studies employing apps downloaded to participants’ smartphones.

There are nonetheless a number of development and methodological concerns pertaining to assessing alcohol consumption behaviour via smartphone apps that warrant further investigation. Given any app must be programmed for two different frequently updating operating systems with distinct deployment protocols, researchers are likely to require the assistance of one or more app programmers in the development phase. The literature offers little guidance regarding the software development process as it relates to the behavioural sciences (although see [24]). Examples of how to effectively manage and integrate the requirements and
SMARTPHONE APP DEVELOPMENT, COMPLIANCE AND REACTIVITY

effectations of multiple stakeholders when developing apps for behavioural studies are consequently required. Although several recent studies have reported promising results with regard to the validity of app-based methods of capturing alcohol intake information [25–27], issues related to app protocol compliance and reactivity have received less attention. Compliance refers to the extent to which participants adhere to study requirements/protocols throughout the experimental period, while reactivity describes a process whereby the monitoring of a behaviour results in a change in that behaviour over time [1–3]. Several commentators have suggested reactivity in the context of real-time research requires further investigation [21,28]. It is a phenomenon that can emerge for several reasons. Participants might become more aware of the behaviour and are consequently motivated to implement change or the demands of the protocol might provoke a tendency to satisfice [28]. In any real-time study of alcohol intake, a decrease in drinking could be evidence of the former, while a decrease in responding “yes” to the behaviour (and therefore in having to respond to further questions), might suggest the later.

In this paper, we describe the process used to guide development of a custom-built smartphone app designed to capture alcohol intake behaviour in the healthy population for research purposes. We also evaluate methodological issues related to compliance with, and reactivity to, study protocols, as a function of hazard/non-hazard drinker status.

**Materials and Methods**
App Development

The CNLab-A app represents the outcome of an iterative development process that included elements of agile software design; namely, requirements analysis, feature and interface design, and app implementation [24,29].

Requirements analysis

In the requirements analysis phase, the research team determined key variables of interest. To this end, empirical definitions of excessive and binge drinking and variables derived from validated retrospective measures of drinking were examined [30–34]. We aimed to elicit data that would assist in establishing percentage of drinking/non-drinking days, daily/average total standard drinks, daily/average drinking rate, highest drink count in two hours, and blood alcohol content (BAC).

After soliciting advice from the programmer and giving due consideration to the ethical implications, we decided that while the date/time of app submissions would be automatically logged, geo-location would not be recorded due to concerns such data may undermine efforts to preserve user anonymity [35,36]. In early iterations of the app, responses related to date of birth, sex, height, and weight were also recorded. Date of birth and sex details were utilised to link data from the app with information collected via other means (e.g., online surveys or in the lab). Height and weight were used to determine BAC. In this early iteration, all data were uploaded to a dedicated commercial web-server company account. Given concerns articulated in the literature regarding privacy and the secure storage of behavioural data [35,36], later iterations of the app did not include any demographic questions. Instead, a unique ID number was generated for each participant and the data were uploaded to a secure server. In this way, app data were not linked to any personal information.
In this phase, the research team also considered assessment design. Studies reliant on real-time (or near real-time) monitoring of behaviour tend to employ event- or time-based sampling [37]. In the case of the former, participants record the behaviour as it happens or shortly thereafter; in the later, behaviour is recorded in response to signals that occur multiple times a day, often at random [37]. Occasionally, both assessment methods are utilised simultaneously [37]. As substance use is episodic in nature, event-based assessment is considered an appropriate method for tracking both frequency and timing of use [1]. We therefore determined this would be the most pertinent method of collecting alcohol consumption information. A disadvantage of event-based monitoring, however, is that it is difficult to assess compliance; that is, there is no way to verify participants record all events as required [37]. To combat this limitation, we required the app send participants twice-daily prompts asking them if they had consumed alcohol since the last submission. This served to remind participants to record drinking where they had forgotten to do so. In this regard, we adopted a similar assessment protocol as Dulin and colleagues; they also employed event-based assessment along with daily notifications reminding participants to record drinking [25]. We additionally required the app prevent participants from submitting data more than 24 hours in the past or more than 15 minutes into the future to prevent back/forward filling.

Finally, content for each assessment was based on the key variables of interest. Drink type formed the central element of each assessment, as, once selected, all subsequent items were dependent on this choice (see Figure 1). For instance, selecting mid strength beer as the drink type in the app automatically determined alcohol content
and serving size options. In Australia, mid strength beer has an alcohol content of 3.5% and is available, at least in licensed and retail venues, in various standard serving sizes [38,39]. Table 1 details drink type, alcohol content and serving size options available in the app (see Table 1). These options were not meant to be exhaustive, but were designed to capture typical drink types, average alcohol content, and standard serve sizes sold in Australia. The final aspect of assessment involved selecting a start/finish time for drinking.

Decisions made during the requirements analysis phase were documented via an iterative storyboard process. This ensured the research team and programmer had access to a common record.
Note. On opening, CNLab-A asks users if alcohol has been consumed in the last 24 hours. Thereafter, participants are asked if they have consumed alcohol since their last submission. If they indicate – by pressing “No” – that no drinking has occurred, the app can be closed. If participants indicate drinking has occurred – by pressing “Yes” – images of common alcoholic beverages (including beer, wine, cider/premix, spirit/liqueur, and cocktail) are displayed (i). Type of beverage consumed is selected by touching the appropriate image on the screen. Quantity and size consumed for each beverage is indicated via a simple scroll option menu (ii). Alcohol content as a function of beverage type is prefilled. This process is repeated by tapping “Back” in order to add as many drink types as required. Erroneously entered data can be deleted by swiping left. Prior to submitting data, the start and end time of drinking must be specified, again using a scroll option menu (iii). Participants are able to either report drinking in separate sessions or they can leave the app open so as to record beverages as they are consumed. The later option still allows participants to use other features on their phone. Participants can access a history of their submission dates and times (but not their drinking data) via the “History” button. At the conclusion of the experimental period, an automated message thanks participants; gives them simple feedback regarding the number of days they consumed alcohol, total standard drinks consumed, and average daily consumption; and, asks them to remove the app from their smartphone.
Table 1. Drink types, alcohol content and serving size options available in the CNLab-A app

<table>
<thead>
<tr>
<th>Drink type</th>
<th>Alcohol content (%)</th>
<th>Serving sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Beer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full strength</td>
<td>4.8</td>
<td>Glass 200ml</td>
</tr>
<tr>
<td>Mid strength</td>
<td>3.5</td>
<td>Pot/Middy 285ml</td>
</tr>
<tr>
<td>Light strength</td>
<td>2.7</td>
<td>Can/Bottle 375ml</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Schooner 425ml</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pint 570ml</td>
</tr>
<tr>
<td><strong>Wine</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White/Champagne</td>
<td>12.0</td>
<td>Glass 150ml</td>
</tr>
<tr>
<td>Red</td>
<td>13.0</td>
<td></td>
</tr>
<tr>
<td><strong>Fortified Wine</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.0</td>
<td>Glass 60ml</td>
</tr>
<tr>
<td><strong>Cider/Premix spirit</strong></td>
<td>5.0</td>
<td>Bottle 300ml</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bottle 330ml</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Can 375ml</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bottle 500ml</td>
</tr>
<tr>
<td><strong>Spirit/Liqueur</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40.0</td>
<td>Standard 30ml</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Double 60ml</td>
</tr>
<tr>
<td><strong>Cocktail</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40.0</td>
<td>One shot 30ml</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Two shots 60ml</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Three shots 90ml</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Four shots 120ml</td>
</tr>
</tbody>
</table>

**Feature and interface design**

Feature and interface design was informed by app software capabilities and in consultation with the programmer. During this phase, iOS app prototypes were developed and deployed in response to design decisions and feedback from the research team. Deployment was via Test Flight, which allowed invited users to test beta versions of iOS apps.

We determined response selection would be via touch screen and scroll menus. At the commencement of each assessment, for instance, when participants were asked if they have consumed alcohol since their last submission, they indicated “Yes” or “No” by touching either option on the screen. Similarly, it was decided drink types would be presented on one touch screen as a set of images with captions so participants could readily identify and select their beverage. Images were designed using Inkscape, a
freely available vector graphic software package. In the case of beer, images for mid and light strength were opaque versions of the full strength visual. Quantity and size options were available via a simple scroll option menu. Likewise, a scroll menu presented dates and times (in 15-minute intervals) for recording of drinking start/end time. Participants pressed “Submit” to upload (to the server) and end the assessment. Any submission that failed to upload due to lack of Internet connectivity would automatically be uploaded when connectivity was re-established. Figure 1 shows screenshots of the app interface (see Figure 1). The app icon was also designed during this phase. A discreet pattern and app name were chosen to minimise the likelihood friends or family of the participant would realise they were taking part in an alcohol-related study.

**App implementation**

One of the potential advantages of using electronic means of assessment in the behavioural sciences relates to penetration. Such methods can provide investigators with large, diverse samples if facilitated by appropriate implementation decisions related to development and deployment. To that end, the research team decided to develop a native app to be run locally on iOS and Android platforms with marketplace deployment via Apple iTunes and Google Play respectively in order to optimize app availability. Due to funding limitations, however, only an iOS version of the app was initially piloted.

There were two pilot studies. The first involved a small sample of testers ($N = 8$, $M_{age} = 38.13$, $SD = 16.54$, range: 18-68, 37.5% female) and was designed to identify programming bugs and oversights. The second comprised a slightly larger group ($N =$
SMARTPHONE APP DEVELOPMENT, COMPLIANCE AND REACTIVITY

19, $M_{\text{age}} = 37.37$, $SD = 9.73$; range: 22-68, 68.4% female) and focused on eliciting user feedback post piloting. During both pilot studies, testers were asked to keep a hardcopy record of any data submissions so that app data could be checked for accuracy. Average time to submit drinking information during piloting was 34 seconds. Suggestions provided by individuals involved in piloting that were incorporated into the final version of the app included making an instructional video explaining how to use CNLab-A available to participants [40]; providing post-experimental summary feedback regarding total standard drinks consumed and average daily intake; and, several minor changes to the app user interface. Otherwise, pilot testers indicated compliance with app protocols was not onerous. They all reported using a mix of real-time and prompt-based submissions.

The final production version of the app was eventually made available on both iOS (8.4+) and Android (Kitkat 4.1+) platforms. As operating systems evolve, both apps are audited to ensure they continue to function as required.

**Participants and Procedure**

This study was based on data from 671 participants ($M_{\text{age}} = 23.12$, $SD = 7.24$, range: 16-56, 70% female) that form a subset of an ongoing project – entitled CheckMyControl – investigating the relationship between alcohol use and various social and cognitive factors in the healthy population. This subsample completed the app component of the project. Participants were recruited via adverts posted in and around the University of Melbourne, researcher networks, and social media posts. They were fluent in English. The University of Melbourne Human Research Ethics
Committee approved the study in accordance with the standards for ethical research of the National Health and Medical Research Council.

After reading a plain language statement and providing informed consent, participants answered an online researcher-devised demographic survey and the Alcohol Use Disorders Identification Test (AUDIT; [34]). Participants were then required to download and use the CNLab-A smartphone app to record alcohol use for 21 days. They were compensated AU$10 for time spent completing online surveys and AU$0.50 each day information about alcohol consumption was submitted via the app (regardless of whether alcohol had been consumed or not). Participants received a bonus AU$9.50 if app data were submitted on all 21 days. The maximum participants could be reimbursed was AU$30.

Measures

**Alcohol Use Disorders Identification Test (AUDIT)**

The AUDIT is a 10-item screening measure that asks participants to respond to questions assessing alcohol intake, problems, and dependence with reference to the preceding six months. Participants were categorised into hazard \((n = 286)\) and non-hazard \((n = 385)\) groups based on their score, with scores of 8 or more indicative of hazardous alcohol consumption [41].

**CNLab-A app**

CNLab-A is a freely available custom-built app that can be used to record alcohol intake for research purposes. Once downloaded, CNLab-A requires participants to allow it to send them notifications. One notification is pre-set to 8 am while the other
can be set to suit the user. While participants are directed at the outset to record alcohol consumption as it happens (or as soon thereafter as possible), notifications serve to prompt individuals to input information twice daily in case they neglect to do so when drinking. Thus, alcohol intake data can be submitted at any time, either in response to notifications or while drinking. A unique ID code, provided to participants during the online component of the study, is required before the app opens. CNLab-A has previously been found to be a valid measure of alcohol intake [27].

**Statistical Analyses**

Independent $t$-tests and chi-square analyses were conducted to determine whether hazard and non-hazard groups were matched demographically. Homogeneity of variance was assessed using Levine’s Test; where this assumption was violated, adjusted $t$-values and associated degrees of freedom were reported. Effect sizes were computed for $t$-tests using $r$-values; they were interpreted according to Cohen’s guidelines: 0.10 = small, 0.30 = moderate, and 0.50 = large effect [42].

Participants were considered compliant with app protocols if they responded to at least one notification each day. Reactivity was assessed using linear growth model analyses. Conducted using SPSS 24, these analyses examined (1) change in drinks consumed each day and (2) change in daily (“yes”) responding over the 21-day experimental period as a function of AUDIT group membership. In the first analysis, drinks per day (within persons; level 1) were nested within individuals (between persons; level-2); in the second, daily (“yes”) response rate (within persons; level 1) was nested within individuals (between persons; level-2). Time was centred on the first day of data collection (Day 0); each unit of time represents an interval of one day.
SMARTPHONE APP DEVELOPMENT, COMPLIANCE AND REACTIVITY

[43]. All mixed models were estimated using restricted maximum likelihood [44]. Initially, slope variance and autocorrelation were included in both models but they were removed when parameter estimates for these effects were found to be very small (< 0.01).

Results

Descriptive Statistics

At the time of testing, 39.9% of the sample were aged under 20, 46.2% were 20 to 29 years, and 13.9% were aged 30 or over. Median age was 20 years. By comparison, median age in Australia is 38 years [45]. Most participants were born in Australia (66.8%), spoke English as their first language (80%), and lived in urban regions (88.1%). Census data shows 67% of the Australian population is locally born and 79% speak English as their first language [45]. Most Australians (86%) reside in urban regions [46]. A small number of participants indicated they identified as an indigenous Australian (0.6%). This was less than census data suggests is typical for the state (Victoria) in which this study took place (0.8%; [45]). This difference, however, is likely a product of our recruitment campaign. As advertisements for the study were posted in and around the University of Melbourne, the sample contained a large number of young tertiary-aged participants (85.7%). The proportion of indigenous students studying at this institute in 2016 stood at 0.57% [47].

The majority of participants consumed alcohol (92.8%). A small proportion indicated they had been diagnosed with an alcohol (0.3%) or substance (0.4%) use disorder. Hazard and non-hazard groups did not differ significantly with regard to age, \( t(669) = \)
1.02, \( P = 0.31 \), or years of education, \( t(654.11) = 1.65, P = 0.10 \). There was a significant association between gender and AUDIT group membership, \( \chi^2(1, N = 671) = 5.96, P = .02 \). The odds of being a hazardous drinker were 1.51 time higher for males than females (95% CI 0.22, 2.01). There were significant differences between the AUDIT scores of the hazard \( (M = 12.45, SD = 4.30) \) and non-hazard \( (M = 4.64, SD = 3.59) \) groups, \( t(547.67) = 24.94, P < .001 \). Mean alcohol intake indices recorded via CNLab-A are detailed in Table 2 (see Table 2).

Table 2. Average alcohol intake indices as recorded via CNLab-A app (21-days), including hazard \( (n = 286) \) and non-hazard \( (n = 385) \) group totals and differences

<table>
<thead>
<tr>
<th></th>
<th>Total M (SD)</th>
<th>Hazard M (SD)</th>
<th>Non-hazard M (SD)</th>
<th>( t(669) )</th>
<th>95% CF</th>
<th>( LL )</th>
<th>( UL )</th>
<th>( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days drinking</td>
<td>5.32 (4.17)</td>
<td>6.90 (4.02)</td>
<td>4.15 (3.90)</td>
<td>8.91</td>
<td>2.14</td>
<td>3.35</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Total drinks(^a)</td>
<td>24.26</td>
<td>37.45</td>
<td>14.47</td>
<td>12.16</td>
<td>19.27</td>
<td>26.70</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>(25.41)</td>
<td>(28.09)</td>
<td>(17.75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drinks per day</td>
<td>1.20 (1.25)</td>
<td>1.86 (1.38)</td>
<td>0.71 (0.86)</td>
<td>12.45</td>
<td>0.97</td>
<td>1.33</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>Drinks per drinking</td>
<td>3.98 (3.02)</td>
<td>5.53 (2.92)</td>
<td>2.83 (2.56)</td>
<td>12.45</td>
<td>2.27</td>
<td>3.12</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>day</td>
<td>2.20 (2.09)</td>
<td>2.60 (2.14)</td>
<td>1.89 (2.00)</td>
<td>4.41</td>
<td>0.40</td>
<td>1.04</td>
<td>0.17</td>
<td></td>
</tr>
</tbody>
</table>
Consumption

<table>
<thead>
<tr>
<th>Highest drink</th>
<th>4.14 (3.15)</th>
<th>6.03 (3.37)</th>
<th>2.79 (2.09)</th>
<th>14.34</th>
<th>2.80</th>
<th>3.69</th>
<th>0.48</th>
</tr>
</thead>
<tbody>
<tr>
<td>count in 2 hours</td>
<td>2.16 (2.58)</td>
<td>3.48 (2.83)</td>
<td>1.19 (1.86)</td>
<td>11.89</td>
<td>1.91</td>
<td>2.67</td>
<td>0.42</td>
</tr>
<tr>
<td>6/6+ intake</td>
<td>1.31 (1.92)</td>
<td>2.26 (2.27)</td>
<td>0.61 (1.20)</td>
<td>11.14</td>
<td>1.35</td>
<td>1.94</td>
<td>0.40</td>
</tr>
<tr>
<td>8/8+ intake</td>
<td>0.85 (1.44)</td>
<td>1.51 (1.70)</td>
<td>0.36 (0.94)</td>
<td>10.41</td>
<td>0.94</td>
<td>1.38</td>
<td>0.37</td>
</tr>
<tr>
<td>12/12+ intake</td>
<td>0.31 (0.80)</td>
<td>0.58 (1.06)</td>
<td>0.10 (0.43)</td>
<td>7.26</td>
<td>0.35</td>
<td>0.61</td>
<td>0.27</td>
</tr>
<tr>
<td>20/20+ intake</td>
<td>0.07 (0.31)</td>
<td>0.12 (0.41)</td>
<td>0.03 (0.20)</td>
<td>3.55</td>
<td>0.04</td>
<td>0.15</td>
<td>0.14</td>
</tr>
</tbody>
</table>

4/4+ intake refers to occasions where four or more drinks were consumed in one episode.

Drinks refer to self-reported alcohol consumption in Australian standard drinks (1 drink = 10 g alcohol). P-values all < .001. CI = confidence interval of the difference between groups; LL = lower limit; UL = upper limit.

Reactivity

Table 3 details parameter estimates for fixed and random effects of the linear growth analysis model for change in drinks consumed each day. With regard to fixed effects, the non-hazard group reported consuming significantly fewer standard drinks (0.67) at Day 0 than the hazard (0.67 + 1.48 = 2.15) group. The non-hazard group showed no significant decrease in consumption over the 21-day experimental period, whereas the hazard group demonstrated a slight though significant decrease in drinking over the same period (0.04 standard drinks per day). Within-person variance (7.54) equates to

Compliance

On average, participants used CNLab-A 20.27 (SD = 1.88) days out of 21. There were no significant differences between hazard and non-hazard groups with regard to number of days of use, t(669) = 0.69, P = .49. As data submission was either event- or notification-contingent, there was no upper limit to the number of drinking sessions participants could report using the app. Participants received a maximum of 42 notifications asking them to record information about their drinking. They submitted data, on average, 2.00 (SD = 0.41) times per day. There were no significant differences in average number of responses based on AUDIT group membership, t(669) = 0.61, P = .54. There were 27,355 data points captured via the app in total.
2.75 SD units; thus, 95% of observed residuals were between +/-5.49 units of their fitted values. The intercept variance (0.74) corresponds to 0.86 SD units; 95% of the population therefore varied between +/-1.72 units of the typical intercept for their group (see Table 3).

Table 4 shows parameter estimates for fixed and random effects of the linear growth model for change in daily (“yes”) responses. The non-hazard group responded (“yes”) significantly less often (0.32) at Day 0 than the hazard (0.32 + 0.15 = 0.47) group. A slight though significant decrease in responding was evident for both the non-hazard (-0.003 per day) and hazard (-0.003 + (-0.004) = -0.007 per day) groups. Within-person variance (0.14) equates to 0.37 SD units; thus, 95% of observed residuals were between +/-0.75 units of their fitted values. The intercept variance (0.08) corresponds to 0.28 SD units; 95% of the population therefore varied between +/-0.57 units of the typical intercept for their group (see Table 4).

Table 3. Parameter estimates for linear growth model of drinks per day as a function of hazard/non-hazard AUDIT group membership

<table>
<thead>
<tr>
<th>Estimate (SE)</th>
<th>t/z</th>
<th>df</th>
<th>p</th>
<th>95% CI*</th>
<th>LL</th>
<th>UL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(intercept, slopes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (level at Day 0)</td>
<td>0.67 (0.08)</td>
<td>8.68</td>
<td>2301.7</td>
<td>&lt;.001</td>
<td>0.52</td>
<td>0.82</td>
</tr>
<tr>
<td>Time</td>
<td>-0.01 (0.01)</td>
<td>-1.80</td>
<td>13418.</td>
<td>.07</td>
<td>-0.02</td>
<td>0.001</td>
</tr>
<tr>
<td>Hazard</td>
<td>1.48 (0.11)</td>
<td>13.26</td>
<td>2301.7</td>
<td>&lt;.001</td>
<td>1.26</td>
<td>1.70</td>
</tr>
<tr>
<td>Hazard by time</td>
<td>-0.03 (0.01)</td>
<td>-3.50</td>
<td>13418.</td>
<td>&lt;.001</td>
<td>-0.04</td>
<td>-0.01</td>
</tr>
</tbody>
</table>
Table 4. Parameter estimates for linear growth model of daily ("yes") responses per day as a function of hazard/non-hazard AUDIT group membership

<table>
<thead>
<tr>
<th></th>
<th>Estimate (SE)</th>
<th>t/z</th>
<th>df</th>
<th>p</th>
<th>95% CI^</th>
<th>LL</th>
<th>UL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(intercept, slopes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (level at Day 0)</td>
<td>0.32 (0.02)</td>
<td>18.24</td>
<td>982.85</td>
<td>&lt;.001</td>
<td>0.28</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-0.003</td>
<td>-3.63</td>
<td>13418.</td>
<td>&lt;.001</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hazard</td>
<td>0.15 (0.03)</td>
<td>5.99</td>
<td>982.85</td>
<td>&lt;.001</td>
<td>0.10</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>Hazard by time</td>
<td>-0.004</td>
<td>-3.53</td>
<td>13418.</td>
<td>&lt;.001</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(lags-variables)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between-person</td>
<td>0.08 (0.01)</td>
<td>16.86</td>
<td>&lt;.001</td>
<td>0.07</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(level 2) intercept</td>
<td>0.14</td>
<td>81.91</td>
<td>&lt;.001</td>
<td>0.14</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within-person (level 1) residual</td>
<td>(0.002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^CI = confidence interval; LL = lower limit; UL = upper limit.

Discussion

In this study, we aimed to describe the development and implementation of a custom-built smartphone app devised to measure real-time (or near real-time) alcohol consumption behaviour in the healthy population. Designed for use in the research
arena, the app was a product of an iterative process that included elements typical of agile software design. Decisions made during each phase of development were informed by a desire to create an app participants could download onto their own smartphones without ever having to visit the lab. We anticipated this might minimise disadvantages – such as equipment and training costs, participant burden, and missing data – often associated with using some types of electronic protocols, while simultaneously enhancing benefits pertaining to sample size and diversity. We additionally explored methodological factors related to app protocol compliance and reactivity as a function of hazard/non-hazard drinker status.

Compliance with app protocols was high. Participants were required to submit data about their drinking (regardless of whether they had consumed alcohol or not) at least once per day for 21 days. On average, they uploaded data on 96.5% of days and there were no differences between hazard and non-hazard groups with regard to the number of days of use or the number of responses per day. Presumably, the use of daily payments and an end of study bonus (for 21 consecutive days of data submissions) incentivized responding. In previous alcohol intake-based studies utilising various electronic methods of collecting data – including text messaging [48], hand-held electronic devices [49], and interactive voice response systems [8] – incentivized responding resulted in similarly high rates of compliance. It should be noted that when the measure is freely available via marketplace vendors, the size of the sample can quickly balloon. As such, it is important to not only balance the burden of protocol compliance with the incentive offered, but it is also necessary to consider overall budget constraints and to ensure there is some swift means of limiting access to the app if required.
Our results suggest there was some slight degree of reactivity – particularly among hazardous drinkers – to the app protocol. Though the effect was small, the hazard group decreased their intake significantly over the experimental period: 0.80 standard drinks in 21 days, which represented a 2% decrease in total standard drinks for this group. By contrast, non-hazard drinkers showed no significant reduction in consumption over the same period. This accords with evidence from other studies demonstrating some reduction in alcohol consumption due only to measurement among hazardous drinkers [50,51]. Even though participants received no feedback about their drinking during the assessment period, it is possible those in the hazard group were motivated to modify their intake because the act of recording it made them more aware of their behaviour. Equally, the knowledge they were being monitored may have induced them to drink less. Considered a manifestation of the Hawthorne Effect, social desirability is thought to underpin this type of assessment reactivity [52]. It is also possible reductions in consumption were the result of satisficing; that is, participants may have responded “yes” to drinking less often over time in order to avoid having to submit further information via the app. Both hazard and non-hazard groups showed a significant reduction in the frequency of responding “yes” to drinking over the 21-day period. The rate of this reduction was, however, very small: non-hazard participants decreased “yes” responding by 0.06 and hazard participants by 0.14 in 21 days. As such, this reduction might be a reflection of increased familiarity with the app over time, rather than satisficing; that is, participants might have summarized their drinking across a day into fewer submissions once they became more familiar with the app.
There is some debate in the literature regarding reactivity to real-time measures. Several investigators postulate such assessment reduces the likelihood of reactivity related to social desirability as participants record data in the absence of the researcher [28]. Bates and Cox found participants were, for example, more likely to reveal lifetime alcohol consumption details when they completed surveys outside – as opposed to inside – the lab [53]. Other researchers speculate real-time methods are particularly susceptible to reactivity effects because assessments are completed in close proximity to the behaviour, giving participants time to consider their actions [1], though it has also been suggested repeated surveying may reduce reactivity via habituation [54,55]. There is, nonetheless, consensus that reactivity is an overlooked facet of real-time research generally and further investigation is required [21,28,56]. It is possible, for instance, that reactivity differs according to the population. Our finding that hazardous drinkers reacted to the app protocol to a greater extent than non-hazard drinkers would certainly support this supposition. Our data also suggest app-based alcohol-related intervention studies would benefit from the inclusion of a measurement only control condition in order to disentangle effects related to reactivity from those linked to the intervention.

Several limitations to this study must be noted. Although representative of the population at large in terms of country of birth, first language, and usual place of residence, young university students predominated in our sample. Such individuals may have been more or less able/disposed to conform to the protocols of our study. It remains to be seen if similar rates of compliance and reactivity would be evident in other samples; older individuals and/or those
dependent on alcohol might, for example, respond in different ways to app-based protocols designed to assess alcohol consumption behaviour. Additionally, we did not examine if different assessment periods impact compliance and reactivity in diverse ways. A shorter experimental period might, for instance, diminish the effect of reactivity, though this may limit how effectively the app captures variability of intake.

In conclusion, we examined the feasibility of developing and employing an app – downloaded by participants to their own smartphones – designed to collect alcohol intake information for research purposes. We demonstrate how utilising apps such as CNLab-A can yield a potentially large sample representative of the population. Both hazard and non-hazard participants appeared highly compliant when using app protocols. Although there was some evidence of reactivity in our study, especially among hazardous drinkers, effect sizes were small. Our findings suggest the CNLab-A app – or potentially one similar – is a methodologically sound means of examining alcohol consumption behaviour across time. In future, such apps can be paired with those that chart cognition in real-time in order to facilitate a more thorough investigation of the antecedents and consequences of drinking behaviour.

Acknowledgements
The authors wish to thank Jiajie Li and Jemie Effendy who assisted with upgrades to the smartphone app as well as Cameron Patrick from the Melbourne Statistical Consulting Platform for advice in reviewing the analyses. This research was supported by an Australian National Health and Medical Research Council grant (1050766), and an Australian Research Council fellowship (FT110100088). The funding bodies had no role in the study design, collection, analysis or interpretation of data, writing the manuscript, or in the decision to submit the paper for publication.

**Conflict of Interest**

No conflict declared

**References**


22. Poushter J. Smartphone ownership and internet usage continues to climb in emerging economies but advanced economies still have higher rates of technology use. http://www.pewglobal.org/2016/02/22/smartphone-ownership-


   [doi:10.1146/annurev.clinpsy.3.022806.091415]


   [doi:10.1038/141613a0]


effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology* 2014;67:267-277.

[doi:doi.org/10.1016/j.jclinepi.2013.08.015]


[doi:doi.org/10.1016/j.chb.2007.02.021]


[doi:doi.org/10.1146/annurev.psych.54.101601.145030]