Word Count: 3,906/3,500

Tables: 0 (7 supplemental)

Figures: 4 (1 supplemental)

Mobile Health for Alcohol Use Assessment: Longitudinal Effects of Breathalyzer Self-Monitoring in Everyday Context

Yang Lu, M.S.[1], Catharine E. Fairbairn, Ph.D.[1], Jiaxu Han, M.A.[1], Nigel Bosch, Ph.D.[2,3]

[1]Department of Psychology, University of Illinois at Urbana-Champaign

[2]School of Information Sciences, University of Illinois at Urbana-Champaign

[3]Department of Educational Psychology, University of Illinois at Urbana-Champaign

Author Note

Correspondence concerning this article should be addressed to Yang Lu at the University of Illinois, 603 E. Daniel St., Champaign, IL 61820. Email: yanglu13@illinois.edu or to Catharine Fairbairn, 603 E. Daniel St., Champaign, IL 61820. Email: cfairbai@illinois.edu.

Funding Acknowledgement

This research was supported by NIH grants R01AA025969 and R01AA028488.

Declaration of Interests

Authors have no conflicts of interest to declare. Authors do not have financial or personal relationships with the people and organizations involved with this research.

Data Sharing

Data are proprietary and access requests should be directed to the manufacturer (research@bactrack.com). All code used in creating datasets, figures, and in performing analysis can be found at doi.org/10.17605/OSF.IO/ER9UH.

Abstract

Objective: Mobile health tracking technologies have burgeoned, offering objective health information to consumers on an unprecedented scale. But opportunities to directly test effects of such monitoring have been limited. Low-cost mobile breathalyzers are one tool commonly employed for blood alcohol concentration (BAC) assessment. The current study explores outcomes linked with BAC tracking technologies, examining effects on alcohol use and self-estimation in a large US-based sample.

Methods: Participants (*N*=32,179) include individuals who voluntarily purchased a mobile breathalyzer and provided repeated (≥3) adlib readings from 2016-2022. A paired smartphone application prompted a BAC self-estimate ("guess") before displaying measured BAC levels. Analyses included observations collected during active consumption (BAC>0.00%) from breathalyzer users who opted to share anonymized data. Breathalyzer users displaying inattentive patterns of guessing were excluded from self-estimation analyses. The final dataset comprised 787,393 BAC readings and 387,643 self-estimates.

Results: Across participants, the accuracy of BAC guesses increased by 2.38% over the course of breathalyzer use. Associations between breathalyzer use and BAC levels varied significantly according to participants' initial drinking levels, *b*=-0.0062, *95%CI* [-0.0065, -0.0059]. Among heavy drinking participants, BAC levels decreased on average from 0.106% to 0.096%, whereas the reverse trend emerged for lighter drinking participants (0.058% to 0.067%). A similar interaction emerged for BAC underestimation, *b*=-0.0058, *95%CI* [-0.0066, -0.0049], with odds of underestimation decreasing among heavy and increasing among light-drinking participants.

Conclusions: Results indicate promise for mobile BAC tracking technologies as a low-impact intervention with the potential to decrease drinking among individuals who drink heavily—a population particularly susceptible to alcohol-related problems. In contrast, inverted trends emerged for light drinking individuals, highlighting the need for empirical research in the fast-moving landscape of digital health.

Introduction

Alcohol is the seventh leading contributor to death and disability in the world.[1] In the United States, alcohol was implicated in 37.7% of traffic fatalities in 2010, and, from 1995-2019, was responsible for 37% of fall injuries, 24% of firearm injuries, and 21% of suicides.[2] Neurocognitive impairment of executive function occurs as a result of alcohol consumption, causing impulsive decision-making and prioritization of immediate rewards over delayed consequences.[3,4] Impulsiveness and greater willingness to perform risky acts, such as drunk driving, are characteristics of people who inaccurately self-estimate, especially underestimate, their blood alcohol concentration (BAC).[5,6] As a result, increasing awareness of consumption levels is considered integral to effective prevention and treatment efforts. Interventions for alcohol use disorder (AUD) frequently involve alcohol consumption self-monitoring to enhance clients' motivation for change, identify high-risk situations, and support behavior modification.[7,8]

Recent technological advances have laid the groundwork for new methods of investigating and treating AUD.[9,10] Consumer-facing network-connected devices and smartphone applications are promising tools for gathering objective, physiological data with high ecological validity.[9,11,12] Mobile breathalyzers are one such recently popularized instrument. They offer a well-validated proxy for BAC in the form of breath alcohol measurements, aiding self-monitoring by providing external cues to assess intoxication.[10,13] Though the effects of breathalyzer use alone have not been extensively studied, early research on breathalyzer-integrated telehealth programs found some success in reducing alcohol use,[14] attrition,[13] and relapse.[15] Moreover, the use of breathalyzers can facilitate self-reflection about intoxication during alcohol consumption and thus create opportunities for drivers to reconsider driving drunk.[2,16] Such opportunities may be especially valuable for individuals with heavier drinking patterns, including those who regularly consume alcohol at or above the binge consumption threshold (≥0.08% BAC).[17] This population—comprising individuals who tend to experience relatively impoverished internal signals for triggering awareness of excessive consumption—is one especially likely to experience negative consequences from alcohol intoxication.[18,19] Specifically, individuals who consume alcohol at heavy levels experience muted physiological and subjective effects from alcohol, including blunted activity in the amygdala, nucleus accumbens, and the prefrontal cortex, often in aggregate called tolerance.[20,21] Thus, given diminished internal signals, external cues to intoxication like BAC self-monitoring emerge as particularly important for this population.[18] However, prior research on the effects of BAC feedback has been characterized by methodological limitations, including predominantly laboratory-based interventions, limited time periods for self-assessment, and small heavy drinking samples (e.g., N=15; see [22] for a review). The promise of new technology calls for science to match the pace of a rapidly shifting digital landscape, including research exploring the consequences of BAC self-monitoring on a larger scale.

The current study analyzes BAC levels and BAC self-estimates within a large dataset of mobile breathalyzer users. We examine breathalyzer data from participants' everyday drinking contexts in a dataset spanning 6 years, tracing patterns in BAC levels and BAC self-estimates as they emerge within and across individuals. We focus on active drinking episodes (BAC>0.00%) as epochs most relevant to understanding self-perceptions of intoxication and associated harms. Hypotheses for this research were pre-registered prior to analysis: doi.org/10.17605/OSF.IO/ER9UH. We predicted that participants would provide more underestimates than overestimates. We further predicted that participants' BAC levels would decrease and BAC self-estimation accuracy would increase with repeated breathalyzer use.

Finally, we predicted the effects of repeated breathalyzer use would diverge significantly among heavier vs. lighter drinking individuals. Specifically, given diminished baseline internal cues to intoxication available to heavy drinking individuals, as well as legal and social reference points of intoxication (e.g., 0.08% BAC for driving), we predicted the effects of BAC self-monitoring over time would give rise to greater reductions in BAC levels, more pronounced increases in BAC estimation accuracy, and greater reductions in underestimation tendencies among individuals who drink heavily.

Method

Participants

Participants in this study represent individuals residing in the United States who purchased a smartphone breathalyzer and provided BAC readings between the years 2016 and 2022. The dataset consists of 787,393 measurements taken from 32,179 unique breathalyzer user accounts. Readings were taken in naturalistic contexts. Participants provided an average of 39.67 (*SD*=76.49) total breathalyzer readings over the course of 7.34 (*SD*=10.94) months of breathalyzer use. The current research is the first scientific report to issue from this participant sample (although see[23] for a driving mortality analysis employing data collected before 2017). All participants indicated consent for their de-identified data to be stored and used for research purposes. Location data revealed that a plurality of participants resided in California (13.97%), followed by Texas (7.72%), Colorado (6.84%), and Florida (4.79%). No individual user data, aside from location, was collected from this anonymous sample. Of participants, 38.9% (N=12,509) were classified as demonstrating initial heavy consumption patterns, consistent with binge drinking behavior (see also data analytic plan). Broad demographic information for a cross-section of breathalyzer users as provided by the device manufacturer (BACtrack) is presented in supplemental materials (Figure S1).

Procedure

The study explores effects of adlib alcohol monitoring, with participants selecting times and locations for breathalyzer self-assessment. Breathalyzers employed by study participants are pocket-sized (47-57g) devices designed to pair with participants' iPhone or Android smartphones via Bluetooth. Devices represent the full line of commercially available smartphone-integrated breathalyzers manufactured by the company BACtrack (i.e., BACtrack Mobile, C6, C8, and Vio devices), with retail costs for these devices typically ranging between $45-$150. BACtrack breathalyzers are in widespread use, available for purchase via online and physical retail outlets (e.g., Amazon, Walmart, CVS), and demonstrating comparable accuracy for assessing blood alcohol levels to evidential-grade breathalyzers. Research suggests that BACtrack Mobile devices display excellent reliability with police-grade breathalyzers (r = 0.916).[24] Prior to employing the breathalyzer, users are prompted to create an account via the BACtrack portal, linking individual user breathalyzer data across different smartphone and breathalyzer devices.

The smartphone application guides participants through the process of supplying a breath sample. In preparation for supplying a BAC reading, participants are required to undergo a wait period, in the course of which time they provide a BAC self-estimation (i.e., they "guess" their BAC level). The mobile application then instructs participants to blow into the breathalyzer until the displayed meter is filled. Information related to the time, date, and location of the BAC reading is automatically recorded together with the reading. After viewing BAC levels, participants are given the opportunity to indicate whether the reading was taken at least 15 minutes after the last sip of alcohol, so indexing mouth alcohol contamination. Finally, participants are prompted to report whether the breath samples originated from themselves or someone else. Participants indicate their agreement for anonymous data sharing via toggle switch available within the mobile application settings, which can be switched on and off at any point.

Data analysis plan

All code required for replicating dataset exclusions, statistical analysis, and figure creation is available at doi.org/10.17605/OSF.IO/ER9UH. The dataset comprised observations recorded between January 1, 2016, and February 28, 2022, reflecting the date at which BACtrack had accumulated sufficient data to permit appreciable sharing and the date of the research team's data use request, respectively. Primary analyses focused on positive BAC readings, with 0.00% BACs readings excluded. This exclusion was applied for several reasons including: 1) the study aims pertain to understanding severity of intoxication (vs. alcohol consumption frequency); 2) the rate of BAC self-assessment during sobriety was relatively low; and 3) the task of "guessing" BAC during sober moments was judged qualitatively different (i.e., less difficult). In addition, BAC readings marked by breathalyzer users as invalid were excluded, as were readings judged as supplied by institutional (vs. individual) users and observations from individuals who supplied insufficient BAC readings for longitudinal analysis ($N \leq 2$). For analyses specifically involving BAC guesses (vs. BAC levels), we also excluded breathalyzer users whose entries pointed to an inattentive or haphazard approach to providing BAC guesses, operationalized as those who consistently (>50% of instances) entered BAC guesses of 0.00% during moments of active consumption ($BAC > 0.02$) (see Figure 1). This resulted in a subsample of 18,800 participants (387,643 observations) employed for self-estimation analyses, compared to 32,179 participants (787,393 observations) for BAC level analyses. A full flow chart of study exclusions is provided in Figure 1.

Models evaluated the effects of repeated breathalyzer self-assessment and heavy drinking status on BAC level, self-estimation accuracy, and self-estimation pattern (i.e., odds of underestimation). Repeated breathalyzer self-assessment was operationalized as the order of BAC readings, ranked according to time stamps, within de-identified BACtrack user accounts. For primary analyses, participants were considered as exhibiting heavy drinking patterns if the average BAC of the first third of their supplied readings exceeded 0.08%. Alternative heavy drinking definitions are also explored in sensitivity analyses. Since the passage of time (e.g., age/maturity) is confounded with repeated breathalyzer self-assessment, time (i.e., days elapsed since first breathalyzer self-assessment) was included as a covariate in all models. BAC self-estimation accuracy was operationalized as the absolute difference between BAC level and BAC self-estimation, or as the inverse of estimation error. Given floor effects associated with BAC guesses at lower intoxication levels, and thus naturally higher accuracy at lower levels of consumption, all models exploring accuracy as outcome also integrated BAC level as covariate. Where absolute difference analyses reached significance, robustness checks repeated analyses substituting adjusted percent difference scores as outcome.

Analyses were conducted using multilevel models accounting for the clustering of observations within participants.[25] To account for non-normally distributed outcomes, generalized mixed models assuming a Poisson distribution and over-dispersed residuals were employed for BAC accuracy and BAC level analyses. Logistic models were employed in analyses predicting odds of underestimation, whereas standard linear models were used in predicting adjusted percent difference scores, which followed a normal distribution. Subject-level intercepts and slopes for the ranked assessment number variable were treated as random in all analyses. For other within-subject factors, all components were first estimated as random, adopting a maximal model structure, with results derived from the most complex model structure that reached convergence reported below.[26] As study hypotheses pertain to overall patterns of change, linear (vs quadratic/cubic) slopes were the focus of analysis, although sensitivity analyses explored the longevity of effects across epochs of breathalyzer self-assessment. Effect size estimates are provided in the form of Event Rate Ratios (*ERR*) for Poisson models and as Odds Ratios (*OR*) for logistic models. Change over the course of breathalyzer self-assessment was calculated by centering the ranked assessment number variable at the first and last (average maximum # readings across users) breathalyzer usage and, where relevant, calculating adjusted percent difference across the interval.

Results

Descriptives

See Figure 2 for sample descriptives. In the final sample used in BAC level analysis, mean BAC level was 0.078% (*SD*=0.057). Across the full course of breathalyzer usage, the average BAC among heavy drinking individuals was 0.101% (*SD*=0.061) and 0.061% (*SD*=0.048) among light drinking individuals. For observations used in self-estimation analyses (see Figure 2), the average absolute difference between BAC level and BAC self-estimation was 0.040% (*SD*=0.043). Of BAC self-estimates, 29.4% were above and 70.6% were below the measured BAC level.

Repeated breathalyzer self-assessment

Results indicated that, across the full sample of breathalyzer users included in the dataset, repeated breathalyzer self-assessment predicted increases in the accuracy of BAC self-estimation (i.e., decreases in BAC estimation error): absolute difference, b=-0.0008, *ERR*=0.9992, t=-7.65, p=< .0001, *95% CI* [-0.0010, -0.0006]; adjusted percentage difference, b=-0.0406, t=-9.24, *p*<.0001, *95% CI* [-0.0492, -0.0320] (see Figure 3). Over the full course of using the breathalyzer, the average breathalyzer user increased their BAC self-estimation accuracy by 2.38%. Repeated breathalyzer self-assessment was also linked with slight increases in BAC levels across all individuals who consumed alcohol in the sample, *b*=0.0006, *ERR*=1.0006, *t*=7.72, *p*<.0001, *95% CI* [0.0005, 0.0008]. Further, models examining self-estimation pattern indicated that participants were more likely to underestimate their BAC level with repeated BAC self-assessment, *b*=0.0023, *OR*=1.0023, *t*=9.44, *p*<.0001, *95% CI* [0.0018, 0.0028] (see Table S1). Of note, however, analyses indicated effects of repeated breathalyzer use were further moderated by participants' initial drinking levels (see below).

Initial heavy drinking levels

Models indicated a significant interaction between initial individual drinking levels and repeated BAC self-assessment in predicting BAC levels, *b*=-0.0062, *ERR*=0.9938, t=-37.53, *p*<.0001, *95% CI* [-0.0065, -0.0059] (see Table S2). Repeated breathalyzer self-assessment was associated with decreased BAC levels among heavy drinking participants, *b*=-0.0026, *ERR*=0.9974, *t*=-19.80, *p*<.0001, *95% CI* [-0.0028, -0.0023], and increased BAC levels among light drinking participants, *b*=0.0036, *ERR*=1.0036, *t*=30.61, *p*<.0001, *95% CI* [0.0034, 0.0038] (see Figure 4). Over the full course of using the breathalyzer, from initial to final self-assessment, BAC levels for the heavy drinking participants decreased, on average, from .106% to .096%. In contrast, the average light drinking participant's BAC level increased from .058% to .067%.

Results indicated a significant interaction between initial individual drinking levels and repeated breathalyzer self-assessment in predicting odds of BAC underestimation, *b*=-0.0058, *OR*=0.9942, *t*=-12.66, *p*<.0001, *95% CI* [-0.0066, -0.0049]. Among heavy drinking participants, repeated breathalyzer self-assessment was associated with decreasing odds of BAC underestimation, b=-0.0010, OR=0.9991, t=-2.68, p=.0074, *95% CI* [-0.0016, -0.0003]. The opposite trend was observed among light drinking participants with increasing odds of BAC underestimation, b=0.0048, OR=1.0048, t=14.91, p<.0001, *95% CI* [0.0042, 0.0054]. Whereas, at the initiation of BAC self-assessment, heavy drinking participants tended to underestimate their BACs to a greater extent than lighter drinking participants, b=0.2745, OR=1.3159, t =17.72, p<.0001, *95% CI* [0.2441, 0.3048]; by the end of self-assessment, this difference was significantly smaller, b=0.0978, OR =1.1027, t=6.67, p<.0001, *95% CI* [0.0690, 0.1265] (see Table S2).

In contrast, self-estimation accuracy increased (i.e., error diminished) for both heavy, b=-0.0011, ERR=0.9989, t=-7.56, p<.0001, *95% CI* [-0.0013, -0.0008], and light drinking participants, b=-0.0005, ERR=0.9995, t=-3.91, p<.0001, *95% CI* [-0.0008, -0.0003]. A significant interaction indicated that heavy drinking participants demonstrated a larger increase in self-estimation accuracy compared with light drinking participants, b=-0.0005, ERR=0.9995, t=-2.77, p=.0056, *95% CI* [-0.0009, -0.0002]. Of note, however, this interaction did not replicate in robustness checks using an adjusted percentage difference accuracy index (see Table S2).

Sensitivity analyses

Sampled breathalyzer users self-selected contexts for breathalyzer readings; thus, it is possible that BAC levels and BAC guesses would change not only with repeated breathalyzer use but also the time/context selected for self-assessment. Therefore, all models reaching significance were repeated, integrating contextual covariates (weekend/weekday; time of day; holiday; season) as a form of robustness check. Core results remained consistent in direction and significance level after the inclusion of contextual covariates.

For the purposes of primary analyses, we operationalized heavy drinking individuals according to average BAC levels observed within the first third of supplied breathalyzer readings. This operationalization was one aimed at reflecting users' initial drinking levels (all supplied $\geqslant$3 readings) while mitigating noise linked with operationalizations hinging on a single (initial) BAC reading. However, as a robustness check, analyses were repeated defining heavy drinking individuals as individuals displaying: 1) BAC >0.08% at initial reading; 2) Average BAC >0.08% across all readings provided. All models remained consistent in significance level and direction with these alternative heavy drinker definitions, with the exception of a single null finding obtained for the latter of these operationalizations in predicting overall BAC levels—a finding potentially attributable to the lack of sensitivity to change over time intrinsic to this particular heavy drinking definition, as well as potential redundancy between predictor and outcome (see Table S3).

Further, we tested the longevity of effects, exploring the extent to which significant effects were driven by early changes versus endured across epochs of breathalyzer use. Notably, interactions between initial drinking levels and breathalyzer self-assessments remained significant for BAC level and estimation pattern outcomes in models where the first 5 and 15 user-supplied BAC readings were excluded from analysis. In contrast, effects for accuracy appeared to emerge primarily within the first phase of breathalyzer use (see Table S4). Finally, key models were re-estimated using binomial rather than Poisson distributions to assess the robustness of distribution-related assumptions, and results remained consistent in magnitude, direction, and significance-level (see Table S5).

Discussion

Exploring observations from a large-scale dataset spanning six years, we investigated associations between BAC self-monitoring and drinking among mobile breathalyzer users who supplied readings across a range of real-world drinking environments. Results indicated that individuals tend to underestimate their BAC levels (69.6%), with the average difference between true BAC levels and BAC "guesses" amounting to the equivalent of two standard drinks (0.04% BAC). Breathalyzer self-monitoring was associated with significant increases in BAC estimation accuracy, with the accuracy of BAC guesses among all participants increasing by 2.38% over the course of breathalyzer use. In contrast to findings for estimation accuracy, which emerged as consistent in direction across heavy and light drinking individuals, associations between breathalyzer use and BAC levels varied significantly depending on initial drinking levels. Among heavy drinking individuals, BAC levels decreased from 0.106% to 0.096% over the course of BAC self-assessment, whereas the reverse trend emerged among lighter drinking individuals (0.058% to 0.067%). A similar pattern emerged in models examining BAC underestimation frequency, with heavy drinkers decreasing while light drinkers increased their odds of BAC underestimation over the course of breathalyzer self-assessment. All effects remained significant after parsing effects of the frequency of breathalyzer self-assessment from the passage of time, as well as models integrating covariates aimed at capturing potential changes in the context of self-assessment.

While opportunities for BAC self-monitoring in everyday drinking contexts have burgeoned in recent years, empirical research has not consistently kept pace. Research examining effects of BAC self-monitoring was mainly conducted several decades in the past, characterized by predominantly laboratory-based feedback training programs,[27,28] fixed alcohol dosing regimens,[29] and inadequately powered between-participant designs.[22] Using these methods, results of prior studies indicate effects of BAC self-assessment frequently fail to extend beyond the period of laboratory training.[22] In the decades since this research was originally conducted, technology has emerged permitting the assessment of BAC feedback in field contexts over the longer term.[9,12] In the present study, we examined a large sample of individuals engaging with novel mobile health technology, and, using these new methods, we found that gains from BAC self-monitoring also extend to real-world environments.

Individuals displaying heavy drinking patterns have long represented a population of interest in the assessment of BAC self-monitoring programs.[28–30] Those who drink heavily are more likely to experience negative consequences from alcohol, including across legal, health, economic and social domains.[2,31,32] In addition, heavy drinking individuals report fewer acute physiological and psychological effects of consuming alcohol,[18,19,33] and thus, in light of the dearth of interoceptive feedback, may derive particular benefit from externally-assisted programs of self-monitoring.[34] Thus, these individuals have the potential to evince particular gains in response to BAC-feedback training. Prior studies, however, have found mixed effects, with gains for heavy drinking individuals often failing to endure beyond the training context.[22] As noted above, prior studies have relied overwhelmingly on laboratory-based designs featuring fixed-dosing training procedures.[22,28] One limitation of this prior research is neglect of the contexts in which habitual alcohol use occurs, which could serve as cues that prime the body for consumption—automatically triggering physiological drives towards homeostasis—and thus subjective effects of the same dose of alcohol can vary dramatically across contexts.[31,35] For this reason, physiological sensations experienced in response to drinking in the lab can diverge substantially from those experienced in the pub.[35] Such context-dependent acute alcohol effects are likely to be especially pronounced for heavy drinking individuals, among whom learned associations between drinking and context are likely to be especially well established.[31,35] Further, given ethical and practical constraints impacting laboratory dosing procedures, alcohol doses administered within lab-based training procedures are likely to represent an atypical (i.e., mild) intoxication experience for this population.[9,12] In sum, while theory points to heavy drinking individuals as a population likely to benefit from BAC training, prior research has failed to offer support for such predictions, a disconnect that may be due to effects of atypical laboratory contextual and dosing procedures employed in these studies.

Findings of the current study, which examined BAC feedback delivered in everyday drinking contexts in response to ad-lib consumption, indicate BAC self-assessment is linked with decreasing BAC levels among heavy drinking individuals. These declines are accompanied by a diminished tendency to underestimate BAC, presenting the possibility that the decreasing tendency towards BAC underestimation may underlie reduced consumption. Such results indicate promise for BAC self-monitoring programs among a population in particular need of alcohol use intervention—individuals exhibiting heavy drinking patterns and thus particularly likely to experience negative consequences as a result of alcohol intoxication.[18,33] When provided objective BAC feedback, individuals who consume alcohol at higher levels may be alerted to their level of consumption, thus increasing attentiveness to the physical, social, and legal consequences of overconsumption

In contrast, effects of repeated self-assessment among lighter drinking individuals emerged as more complex, with breathalyzer self-assessment being linked to *increases* in BAC levels. One potential explanation for such results might be found in the familiar anchor point BAC (0.08%) linked with driving legality in the US—an anchor point that also represents a common motive driving independent BAC self-assessment for many.[1,21] Among some individuals, BAC self-assessment may lead to the recognition that their intoxication level regularly exceeds this 0.08% threshold whereas, among others (e.g., lighter drinking individuals), that they might consume additional drinks before exceeding this limit. Thus, results present the possibility that BAC self-assessment increases alcohol consumption for some. As digital health and consumer-driven interventions gain increasing popularity, findings of the current study point to the importance of empirical research to directly evaluate outcomes of these novel approaches, including studies powered to explore how effects might vary across individuals.[9,12]

Limitations of this research should be noted. Breathalyzer users examined here selected times for self-assessment. An advantage of this design is the potential to observe effects of BAC feedback within real-world drinking contexts in response to habitual consumption patterns. However, effects for both heavy and light drinking individuals might be influenced by a combination of factors, including not only effects of repeated BAC self-assessment but also (potentially) evolving choices surrounding contexts for self-assessment. Although results of the current research were unchanged with the inclusion of a variety of measured contextual covariates, the potential for influence by unmeasured factors remains. In addition, the sample of participants examined in this study reflects a self-selecting group who purchased a breathalyzer for individual use. Though a large population in their own right, these individuals may have characteristics that set them apart from general populations, and thus the generalizability of results here to other groups awaits further investigation. Additionally, the current investigation targets BAC readings taken during active drinking episodes. Results of this study carry information surrounding drinking quantity and extremity of intoxication, whereas any potential effects of BAC self-assessment on alcohol use frequency await future study. Further, participants in the current study were not explicitly informed of their exact level of accuracy/inaccuracy, but rather the temporal proximity between the time at which they provided a guess and the time at which readings were supplied was judged as facilitating internal comparison. Finally, there remains a possibility that the trends observed in this study are explained by a reversion to the mean among heavy and light-drinking individuals. However, given covariates included in models (i.e., passage of time) and results of robustness checks (i.e., alternative heavy drinker operationalizations comprising all readings), such an explanation seems unlikely as a primary explanation for observed effects.

In sum, recent years have seen a proliferation of tools available for technology-assisted mobile health interventions. Findings suggest that optimism surrounding the potential of these interventions should be accompanied by rigorous empirical investigation, as effects can emerge as complex and variable across individuals. Future research might explore semi-randomized sampling of responses from health monitoring tools in real-world contexts.

References

1. Griswold M, Fullman N, Hawley C, et al. Alcohol use and burden for 195 countries and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet*. 2018;392(10152):1015-1035. doi:10.1016/S0140-6736(18)31310-2

2. Alpert H, Slater M, Yoon Y, Chen C, Winstanley N, Esser M. Alcohol consumption and 15 causes of fatal injuries: A systematic review and meta-analysis. *Am J Prev Med*. 2022;63(2):286-300. doi:10.1016/j.amepre.2022.03.025

3. Fairbairn CE, Kang D, Federmeier KD. Alcohol and neural dynamics: A meta-analysis of acute alcohol effects on event-related brain potentials. *Biological Psychiatry*. 2021;89(10):990-1000.

4. Leigh BC. Peril, chance, adventure: concepts of risk, alcohol use and risky behavior in young adults. *Addiction*. 1999;94(3):371-383. doi:10.1046/j.1360-0443.1999.9433717.x

5. Laude J, Fillmore M. Drivers who self-estimate lower blood alcohol concentrations are riskier drivers after drinking. *Psychopharmacology*. 2016;233(8):1387-1394. doi:10.1007/s00213-016-4233-x

6. Stephan R, Alhassoon O, Allen K, et al. Meta-analyses of clinical neuropsychological tests of executive dysfunction and impulsivity in alcohol use disorder. *Am J Drug Alcohol Abuse*. 2017;43(1):24-43. doi:10.1080/00952990.2016.1206113

7. Miller WR, Rollnick S. *Motivational Interviewing: Helping People Change*. Guilford Press; 2012.

8. Miller WR. Behavioral treatment of problem drinkers: A comparative outcome study of three controlled drinking therapies. *Journal of Consulting and Clinical Psychology*. 1978;46(1):74.

9. Goldfine C, Lai J, Lucey E, Newcomb M, Carreiro S. Wearable and wireless mHealth technologies for substance use disorder. *Curr Addict Rep*. 2020;7(3):291-300. doi:10.1007/s40429-020-00318-8

10. Delgado MK, Huang Y, Wanner K, et al. Test accuracy of smartphone-paired breathalysers: A validation study. *Injury Prevention*. 2017;23.

11. Kazemi D, Borsari B, Levine M, Li S, Lamberson K, Matta L. A systematic review of the mHealth interventions to prevent alcohol and substance abuse. *J Health Commun*. 2017;22(5):413-432. doi:10.1080/10810730.2017.1303556

12. Steinhubl SR, Muse ED, Topol EJ. Can mobile health technologies transform health care? *JAMA*. 2013;310(22):2395-2396.

13. Hallgren K, Galloway G, Witkiewitz K, Linde P, Nix B, Mendelson J. Treatment retention and reductions in blood alcohol concentration (BAC) during the first 90 days of a telehealth

program for alcohol use disorder. *Am J Drug Alcohol Abuse*. 2023;49(2):249-259. doi:10.1080/00952990.2023.2175322

14. Mitchell M, Mendelson J, Gryczynski J, Carswell S, Schwartz R. A novel telehealth platform for alcohol use disorder treatment: Preliminary evidence of reductions in drinking. *Am J Drug Alcohol Abuse*. 2020;46(3):297-303. doi:10.1080/00952990.2019.1658197

15. Hämäläinen M, Zetterström A, Winkvist M, et al. Real-time monitoring using a breathalyzer-based eHealth system can identify lapse/relapse patterns in alcohol use disorder patients. *Alcohol Alcohol*. 2018;53(4):368-375. doi:10.1093/alcalc/agy011

16. Zaloshnja E, Miller T, Blincoe L. Costs of alcohol-involved crashes, United States, 2010. *Ann Adv Automot Med*. 2013;57:3-12.

17. US Department of Health & Human Services; NIH; National Institute on Alcohol Abuse & Alcoholism. NIAAA Council Approves Definition of Binge Drinking: (306662005-004). Published online 2004. doi:10.1037/e306662005-004

18. Brumback T, Cao D, King A. Effects of alcohol on psychomotor performance and perceived impairment in heavy binge social drinkers. *Drug and alcohol dependence*. 2007;91(1):10-17.

19. Paulus MP, Schuckit MA, Tapert SF, et al. High versus low level of response to alcohol: evidence of differential reactivity to emotional stimuli. *Biological psychiatry*. 2012;72(10):848-855.

20. Blaine S, Sinha R. Alcohol, stress, and glucocorticoids: From risk to dependence and relapse in alcohol use disorders. *Neuropharmacology*. 2017;122:136-147. doi:10.1016/j.neuropharm.2017.01.037

21. Trimmel M, Köpke E. Motivations to control drinking behavior in abstainers, moderate, and heavy drinkers. *Pharmacol Biochem Behav*. 2000;66(1):169-174. doi:10.1016/S0091-3057(00)00199-4

22. Aston E, Liguori A. Self-estimation of blood alcohol concentration: A review. *Addict Behav*. 2013;38(4):1944-1951. doi:10.1016/j.addbeh.2012.12.017

23. Aschbacher K, Hendershot CS, Tison G, et al. Machine learning prediction of blood alcohol concentration: A digital signature of smart-breathalyzer behavior. *npj Digit Med*. 2021;4(1):1-10. doi:10.1038/s41746-021-00441-4

24. Delgado M, Shofer F, Wetherill R, et al. Accuracy of consumer-marketed smartphone-paired alcohol breath testing devices: A laboratory validation study. *Alcohol Clin Exp Res*. 2021;45(5):1091-1099. doi:10.1111/acer.14597

25. Raudenbush SW, Bryk AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd ed. Sage Publications; 2002.

26. Barr D, Levy R, Scheepers C, Tily H. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang*. 2013;68(3):255-278. doi:10.1016/j.jml.2012.11.001

27. Köchling J, Geis B, Chao C, Dieks J, Wirth S, Hensel K. The hazardous (mis)perception of Self-estimated Alcohol intoxication and Fitness to drivE—an avoidable health risk: The SAFE randomised trial. *Harm Reduct J*. 2021;18(1):122. doi:10.1186/s12954-021-00567-4

28. Lipscomb TR, Nathan PE, Wilson GT, Abrams DB. Effects of tolerance on the anxiety-reducing function of alcohol. *Archives of general psychiatry*. 1980;37(5):577.

29. Shapiro A, Nathan P, Hay W, Lipscomb T. Influence of dosage level on blood alcohol level discrimination by alcoholics. *J Consult Clin Psychol*. 1980;48(5):655-656. doi:10.1037/0022-006X.48.5.655

30. Bullers S, Ennis M. Effects of blood-alcohol concentration (BAC) Feedback on BAC estimates over time. *J Alcohol Drug Educ*. 2006;50(2):66-87.

31. Carvalho AF, Heilig M, Perez A, Probst C, Rehm J. Alcohol use disorders. *The Lancet*. 2019;394(10200):781-792.

32. Bouchery EE, Harwood HJ, Sacks JJ, Simon CJ, Brewer RD. Economic costs of excessive alcohol consumption in the US, 2006. *American Journal of Preventive Medicine*. 2011;41(5):516-524. doi:10.1016/j.amepre.2011.06.045

33. Holdstock L, King A, De Wit H. Subjective and objective responses to ethanol in moderate/heavy and light social drinkers. *Alcohol Clin Exp Res*. 2000;24(6):789-794. doi:10.1111/j.1530-0277.2000.tb02057.x

34. Sanchez-Craig M, Annis H. 'Self-monitoring' and 'recall' measures of alcohol consumption: convergent validity with biochemical indices of liver function. *Alcohol Alcohol*. 1982;17(3):117-121. doi:10.1093/oxfordjournals.alcalc.a044315

35. Fairbairn CE, Bresin K, Kang D, et al. A multimodal investigation of contextual effects on alcohol's emotional rewards. *Journal of Abnormal Psychology*. 2018;127(4):359-373. doi:10.1037/abn0000346

Figures

Figure 1. Data Preprocessing Pipeline

Note. Final Data Set 1 was used to examine study aims related to BAC level; Final Data Set 2 (i.e., "sub-sample") was used to examine study aims related to BAC self-estimation accuracy (for details, see Data analysis plan).

Readings with Valid Dates and Time: Readings were excluded if the associated date or time (hour, minute, second) was missing. Individual Users: BACtrack breathalyzers are purchased for use within both personal and institutional settings. Accounts linked with >1,000 BAC readings were judged as reflecting institutional users and excluded from analysis. Valid Readings: Readings were excluded if potentially produced by user or device error (Reading with ⩾ 0.3 BAC), submitted at the same date and time (Duplicate Readings: the lowest BAC level reading is preserved in each set of duplicates), or participants did not wait 15 minutes since the last consumption of alcohol before submitting the breath sample, which elevated risk of reading contamination by mouth alcohol (Contaminated Readings). Readings Supplied by Self: Readings were excluded if marked by users as supplied by or for a friend. Readings with Positive BAC: Analyses focused on positive BAC readings (BAC>0.00%) for several reasons including; a) primary study aims centered around understanding intoxication extremity vs. consumption frequency; b) the task of guessing BAC during moments of sobriety was judged as substantively different; c) rate of sampling during sober moments was relatively sparse, and thus 0.00% readings were judged as comprising a non-representative sample of moments of sobriety. Users with ⩽ 2 Readings: Users were required to supply at least 3 readings for longitudinal analysis. Inattentive Guesses: Participants were excluded if they guessed a BAC of 0 in over half of the instances in which their BAC was actually above 0.02.

Alt text: Vertical flow diagram of exclusion criteria to create BAC level analyses dataset and a smaller BAC self-estimation accuracy dataset.

Figure 2. Descriptive Statistics

A. Geographic distribution of breathalyzer readings

Alt text: A map of the United States with red dots representing locations of reported BAC readings.

B. Distributions of BAC level, BAC guess, and estimation types

Alt text: Colored dots represent each BAC at different BAC levels and self-estimations, also depicted as a pie chart and frequency bars.

Figure 3. Repeated Breathalyzer Self-Assessment Predicted Increased BAC Self-estimation Accuracy

Note. Error bar represents 95% confidence interval for the median self-estimation error (adjusted percentage difference). Values displayed in the above figure are estimated directly from raw data (vs trend lines derived from multilevel models). Figure illustrates change across 40 instances of BAC feedback, which was the average number of breathalyzer self-assessments across participants in the larger dataset.

Alt text: A bar graph of median BAC self-estimation error across BAC self-assessment, with 95% confidence interval as error bars.

Figure 4. Initial Drinking Level (Light vs. Heavy) Moderated Relationship between Repeated BAC Self-assessment and Both BAC Levels and BAC Underestimation

Note. Error bar represents 95% confidence interval for the median BAC level. Blue bars and trendline represent participants who initially consumed alcohol at light levels. Purple bars and trendline represent participants who initially consumed alcohol at heavy levels. Values displayed in the above figure are estimated directly from raw data (vs trends derived from multilevel models). Figure illustrates change across 40 instances of BAC feedback, which was the average number of breathalyzer self-assessments across participants in the larger dataset.

Alt text: A double bar graph of median BAC on top of two trendlines of underestimation patterns across BAC self-assessments, split by heavy/light drinking.