Published in final edited form as:

Addict Res Theory. 2016 ; 24(4): 330–340. doi:10.3109/16066359.2016.1145674.

The Reliability of Naturalistic Observations of Social, Physical and Economic Environments of Bars

Christopher Morrison, M.P.H.^{1,2}, Juliet P. Lee, Ph.D^{1,*}, Paul J. Gruenewald, Ph.D¹, and Christina Mair, Ph.D.^{1,3}

¹ Prevention Research Center, Pacific Institute for Research and Evaluation, Oakland, California

² Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Australia

³ University of Pittsburgh Graduate School of Public Health, Department of Behavioral and Community Health Sciences, Pittsburgh Pennsylvania

Abstract

Drinking in bars contributes to numerous public health problems, including violence and motor vehicle crashes. In order to formulate effective preventive interventions it is essential to identify which specific features of bar environments are related to increased risks. Unobtrusive ethnographic observations are one approach that has been used to characterize these features; however no studies have assessed reliability in a representative sample of bars. We performed brief scouting assessments in all 165 bars in six purposively selected California cities, followed by unobtrusive observations from a subsequent representative sample of 97 bars which were located in low and high bar density areas of the cities. Inter-rater reliability between two independent observers assessed individual item reliability, and principal components analyses assessed the reliability of a series of scales describing the physical, social, and economic characteristics of the bars. For the scouting assessment, items exhibited at least moderate reliability (κ or $r \geq 0.40$). For the unobtrusive observations, items assessing physical and economic environments (e.g., pool table present, $\kappa = 0.90$; index beer cost, $r = 0.82$) had moderate to outstanding reliability (κ or $r > 0.80$). Items describing the social environment generally had poorer reliability, though group aspects (e.g., patron count, $r = 0.78$; patron circulation, $r = 0.64$) had better reliability than individual behaviors (e.g., derogatory speech, $\kappa = 0.12$). Scales constructed from specific sets of items exhibited modest reliability. The individual metrics and principal components we present will enable future studies seeking to disaggregate relationships between bar characteristics and public health problems.

Keywords

Bar; drinking; ethnography; quantitative; reliability; survey

* Corresponding author, Phone: +1-510-883-5772, Fax: +1-510-644-0594, jlee@prev.org.

CONFLICT OF INTEREST: None declared

INTRODUCTION

The World Health Organization estimates that 5.9% of deaths and 5.1% of the global burden of death and disease are attributable to alcohol consumption, including at least 15% of deaths from traffic fatalities and 22% of deaths resulting from interpersonal violence (World Health Organisation, 2014). In the USA, despite continued prevention and enforcement efforts, drunken driving and alcohol-related violence remain very common and resilient problems in communities. Drinking and driving continues to be one of the leading sources of mortality related to alcohol use in the United States. Although driving under the influence has reduced over the past several decades, the recent National Roadside Survey showed that 8.3% of weekend nighttime drivers were alcohol-positive, with an additional 1.5% over the legal limit (Berning, Compton, & Wochinger, 2015). In 2005, there were 16,885 alcohol-related fatalities, representing about 39% of all fatalities for that year. Young drivers aged 21-24 had the highest rate of fatal alcohol-related crashes, accounting for about 15% of all such incidents (Voas & Lacey, 2011). While alcohol involvement in violent crime is less well-assessed, US Department of Justice crime statistics show that of 5.2 million violent victimizations of residents age 12 or older in 2007, 14% of victims reported that their assailant was under the influence of alcohol (Department of Justice, 2005).

Although arising in many different community contexts, these problems have been specifically associated with drinking at venues licensed to sell alcohol for on-premise consumption (i.e., bars and restaurants). Statistical associations between alcohol outlets and alcohol-related problems have been demonstrated in ecological studies of outlet concentrations and motor vehicle crashes and violence (Campbell et al., 2009) and individual level studies of problem drinkers and their use of alcohol outlets (Livingston, Laslett, & Dietze, 2008; Treno, Grube, & Martin, 2003; Weitzman, Folkman, Folkman, & Wechsler, 2003). Understanding the characteristics of commercial alcohol outlets and their patrons that are specifically related to these problems is important for communities because (i) many operating characteristics of outlets are modifiable and can be regulated by local, county and state authorities and (ii) preventive intervention programs can be developed that focus upon particular population and individual risk behaviors that lead to problems in these commercial settings.

In order to assess the operating characteristics of bars, researchers have commonly utilized ethnographic methods, particularly naturalistic and participant observation (Bernard, 2006). These enable researchers to observe a wide range of micro-ecological phenomena that otherwise go unmeasured in quantitative survey studies. For example, Cavan's (1966) study of bar norms and bar types was based on participant observation in 100 San Francisco bars with data collection conducted recurrently over a period of three years. Spradley and Mann's (1975) study of gender and division of labor in bar work was based on 12 months of intensive participant observation and interviews in one US college bar. Roebuck and Frese's (1976) study of bar social ecology and patron and employee characteristics was based on 10 years of participant observation in one nightclub. While this level of intensive fieldwork is not always practical, bar characteristics observed in-depth may be simplified and codified in checklists which can be used across a larger number of bars. Ideally, this would enable statistical analyses of the characteristics of interest.

The quintessential example of this method was a study conducted by the research collective Mass-Observation in 1943 (Mass-Observation, 1943). A team of trained investigators and community volunteers compiled and reduced massive amounts of observational data on pubs throughout one English town. The researchers generated statistics on, among other things, types of pubs, types of drinkers, and levels of drinking, as well as qualitative data on diverse aspects of pub life including credit, abstention, and clothing styles. Subsequent studies have utilized quantifiable observation methods to focus on specific aspects of the bar environment theorized to affect risks for heavy drinking (e.g., drinking pace, drinking group characteristics (Harford, Feinhandler, O'Leary, & Dorman, 1983; Hennessy & Saltz, 1993; Kessler & Gomberg, 1974; Sykes, Rowley, & Schaefer, 1990; Sykes, Rowley, & Schaefer, 1993)).

A key reason to focus on barroom environments is that, among the many elements converging to create bar drinking, the barroom environment may be the most likely access point for interventions aiming to reduce and prevent problem drinking associated with bars. Since bars are socially constructed (i.e., the product of repeated confluences of human interactions at specific places) bar environments may plausibly be altered to maximize the likelihood of desired outcomes. Indeed, bar owners and staff may deliberately manipulate these environments to achieve outcomes which support their goals, which may be economic (food and drink sales, tips) as well as social (fun, sociability, status maintenance, reputation enhancement). Prior research has indicated that bar staff may well model and cue behaviors for patrons that are socially appropriate to the desired atmosphere for a bar. For example, the presence of ashtrays and observed bartender smoking were found to correlate with patron smoking in defiance of state smoke-free laws (Lee, Moore, & Martin, 2003; Moore, Lee, Martin, Todd, & Chu, 2009). Ethnographic and other qualitative studies have noted the propensity of bar staff to support heavy and potentially problematic drinking (Nusbaumer & Reiling, 2002; Tutenges, Bøggkjær, Witte, & Hesse, 2013). However, other studies (Turrisi, Nicholson, & Jaccard, 1999) have indicated the tendency for staff and managers to resist interventions attempting to alter their serving practice in particular, as these may impact their social and economic relationships with their patrons (Antin, Moore, Lee, & Satterlund, 2010; Lee, Antin, & Moore, 2008; Satterlund, Lee, & Moore, 2012).

Intervening on bar patrons is another plausible approach. Studies of drinking behaviors have associated risks for bar violence with more impulsive, heavy drinking, male patrons (Leonard, Quigley, & Collins, 2003; Quigley, Leonard, & Collins, 2003; Quigley & Leonard, 2004) and the few studies of bars and driving under the influence have indicated similar associations with heavier drinkers (Snow & Wells-Parker, 2001). However, studies of individuals have typically relied on data from surveys from the general population (Single & Wortley, 1993; Treno, Gruenewald, Remer, Johnson, & LaScala, 2008), from respondents selected for drinking outcomes (e.g., DUI offenders; Snow & Wells-Parker, 2001; Snow & Anderson, 1987), or respondents sampled opportunistically at or near to drinking venues (Duff & Rowland, 2006; Hughes, Anderson, Morleo, & Bellis, 2008). No studies have assessed individual-level characteristics using a generalizable sample of patrons from a generalizable sample of bars (Morrison, Lee, Gruenewald, & Marzell, 2015a). As the association between problem behaviors and problem locations is therefore difficult to establish, appropriate approaches to behavioral interventions on bar patrons are unclear.

In previous studies relating specific physical features of bars to instances of observed aggression, indicators of problem risks were physical decay; floor plans that promote greater or closer patron interaction, such as dance floors, pool tables, noise; and floor plans that encourage milling and crowding, e.g. small venues and those with little or no seating (Graham, Bernards, Osgood, & Wells, 2012; Graham, La Rocque, Yetman, Ross, & Guistra, 1980; Homel, Thomsen, & Thommeny, 1992; Lang, Stockwell, Rydon, & Lockwood, 1995; Morrison, Mair, Lee, & Gruenewald, 2015b; Quigley et al., 2003). Specific social and economic features associated with observed aggression included having a high percent of patrons who appeared to be intoxicated; providing cheap drinks (price discounting); and staff serving apparently intoxicated patrons (Graham, Osgood, Wells, & Stockwell, 2006; Homel, Carvolth, Hauritz, McIlwain, & Teague, 2004; Ker & Chinnock, 2008; Lang et al., 1995; Quigley et al., 2003; Stockwell, Lang, & Rydon, 1993; Warpenius, Holmila, & Mustonen, 2010).

The samples of bars used in these studies have, however, been limited by the study aims. For example, as the outcomes of interest in Graham and colleagues' studies were aggression and violence, their research focused on barrooms with a higher likelihood of observable aggressive incidents (i.e., large capacity bars and/or those identified as problematic by police; Graham, Bernards, Osgood, & Wells, 2006; Graham & Wells, 2001). The reliability of items used in these studies were typically assessed only for use in limited numbers of bars and bar types. No study has assessed the reliability such items across a representative sample of bars from more than one city.

Our study aimed to assess the operating characteristics of a representative sample of bars from six cities in California using observational tools developed through extensive experience with qualitative field methods in these settings. Our first goal was to assess the item reliabilities of the individual measures collected in the study; our second was to determine whether these could be used to construct reliable scales of bar operating characteristics. Assessing whether items and scales can reliably characterize a representative sample of bars is critical for research seeking to identify the specific aspects of bar environments that may contribute to public health problems.

METHODS

We operationally defined "bars" as on-premise drinking places licensed as bars, taverns, or pubs (types 23, 40, 42, 48, 61, and 75 from the California Department of Alcohol Beverage Control) and restaurants (type 47) with a separate bar area which focused on alcohol sales. We included restaurants with separate bar areas, as some such establishments have been found to function as drinking locations (Ponicki et al. 2014). We conducted two surveys: (1) a scouting assessment to determine the location and obviously visible operating characteristics of all bars in a sample of six California cities; and (2) unobtrusive observations in a randomly selected subsample of bars from stratified low and high bar density areas of the cities to provide detailed information on bar, staff and patron characteristics and interactions.

The cities we examined in this study were a subset of 50 cities between 50,000 and 500,000 population from an ongoing study of alcohol availability, drinking contexts and problems among youth and young adults in California. This ecological sample of cities is a representative subset of all 138 cities of this size in the state, randomly selected with the exclusion that no selected city could be spatially adjacent to any other (Gruenewald et al. 2013). Given the intensive fieldwork component of the current study, we excluded cities located more than 60 miles from our research center, ranked the remaining 14 cities according to bar density (per roadway mile), and selected the three highest density and three lowest density cities in order to have a geographic sample representing a full range of alcohol availability. In 2008, field teams conducted an assessment of all 799 on-premise outlets which were operating at that time in these six cities (Gruenewald et al. 2013). Of these establishments, 194 were identified as bars. Returning to these establishments in 2012 (and any others in the six cities that had obtained a license in the intervening four years), we further excluded nightclubs and liquor stores (which can have an on-premise license for wine tastings) and establishments that advertised closing times before 8:00pm on Friday and Saturday nights. In total, 165 eligible bars and restaurant-bars were identified for the current study. This study was approved by the research organization's Institutional Review Board.

Scouting Assessments

Three field scouts visited bars individually to perform eligibility checks on all establishments. In the 165 bars that met the inclusion criteria (i.e., those with a bar/pub license or a Restaurant-General license and a separate bar area), scouts remained on the premises for approximately 20 minutes. Each scout subsequently completed a 12-item hardcopy checklist using items adapted from the data collection form in the 2008 study characterizing the internal and external physical and economic environments of each outlet. Using data from the 128 (77.6%) bars for which we had observations from two scouts, we assessed inter-observer reliability using a Pearson product-moment correlation coefficient (r) for interval measures and Cohen's Kappa (κ) and percent agreement for dichotomous and categorical measures (Feinstein & Cicchetti, 1990). By convention, moderate inter-rater reliability is indicated when $0.40 \leq \kappa < 0.59$, substantial reliability when $0.60 \leq \kappa < 0.79$, and outstanding reliability when $\kappa \geq 0.80$ (Landis & Koch, 1977) with similar ranges generally applying to test-retest reliability assessments when these are not proximate in time (Bakeman, Quera, McArthur, & Robinson, 1997; Litwin, 2002).

Unobtrusive Observations

Subsequent to the scouting assessments, we grouped the 165 eligible bars into those which appeared in low ($n = 56$), medium ($n = 53$), and high ($n = 56$) density areas within each city. We operationally defined density as the median distance between each outlet and its five nearest neighbors. Excluding the middle group from further consideration, we had 112 outlets in low and high density areas. The average distance between outlets in low density areas ranged from 326.9 meters to 2,965.9 meters. The average distance between outlets in high density areas ranged from 21.1 meters to 363.7 meters. Although outlet densities varied from one city to another, in no case was the greatest density in any low density area greater than the lowest density in any high density area.

Research assistants who collected the more extensive observational data underwent 40 hours field training. Each person completed seven or more practice observations in out-of-sample bars prior to conducting his or her first observation. Each practice bar observation was followed by a debriefing session attended by all research assistants and project staff during which we discussed definitions of individual items. Coordinating their observations across a common set of outlets, Pearson correlations were greater than 0.70 and agreements on dichotomous items were greater than 80% at the conclusion of the training phase. Additionally we hired two specialist Spanish-speaking assistants to undertake observations in four bars in which scouts indicated patrons and staff were primarily Spanish-speaking.

Research assistants worked in pairs to conduct observations of 60 to 90 minutes at each bar after 8:00pm on Friday and Saturday evenings. Field teams conducted two observations per evening. The observers did not disclose their research objectives to bar staff and patrons and were both selected and instructed to blend in with the crowd (as characterized by the scouts) in each bar as best as possible. Like the scouts, they were allowed to purchase a menu or bar item, but were not to consume any alcoholic beverages. They were allowed to talk with staff and other patrons in a natural manner, but not to attempt to elicit any information beyond what would be normally expected of a bar patron. Observation teams attended each bar twice. When there were fewer than 5 people in a bar when the field teams arrived, the team was instructed to skip the observation. Two attempts were made to complete observations at all bars. In order to ensure that we had accurately characterized complex bar environments, 36 bars with more than 100 patrons or where the average drinking pace was greater than 2 drinks/person/hour (see below) during the first two visits were eligible for a third 120-minute observation up to the advertised closing time; we were able to complete 26 of these third observations.

Research assistants recorded observational data using a web-based quantitative instrument either after leaving the bar (to avoid drawing unwanted attention to the unobtrusive observations) or sometimes more publicly when they judged that use of mobile devices or brief note-taking would not be obtrusive. The observation instrument was designed to be as simple as possible, operating as a set of checklists programmed in electronic form. Field staff entered data remotely and uploaded these data to a secure password protected study website within 72 hours. The checklist items aimed to characterize the physical, social, and economic environment of each outlet (Holder et al., 1997; Lee et al., 2003; Moore et al., 2009; Treno, Gruenewald, Lee, & Remer, 2007). We developed items assessing social environmental elements observed in prior ethnographic bar studies (Antin et al., 2010; Lee et al., 2008), together with variables identified in prior research on bar-related risks for violence (Graham & Homel, 2008) and drunken driving (Roebuck & Murty, 1996). To record patron throughput the observers recorded the number of new patrons who entered the bar during 10-minute windows at the start and end of the observation period. Number of bar patrons was recorded as counts midway through the observation.

We assessed the intensity of drinking that took place in each bar using a drink pace measure adapted from previous studies (Aitken & Jahoda, 1983; Hennessy & Saltz, 1993; Hunter, Hannon, & Marchi, 1982). Observers sat at bar tables near a rear wall, and selected as many tables as they could reasonably observe in the line of sight between their seat and the bar.

The two research assistants did not always observe the same tables. From an agreed start time they observed the tables for 60 minutes, recording the number of people at each table, the number of drinks consumed by each person, and the length of time in minutes that the patrons remained at the table (0 to 60 minutes). When one group left a table and was replaced, this new group was counted as a new table. Drinks already on the table when the 60 minutes began were counted towards the total, and any single beverage was considered one drink, regardless of size. The metric obtained from these observations was average drink pace for the bar in drinks/person/hour.

Scaling

Similar to the scouting assessment, we assessed item reliability between the two unobtrusive observations conducted at the same visit (i.e., inter-observer reliability). Neglecting all items for which reliabilities were $r < 0.40$ or $\kappa < 0.40$ we also scaled subsets of items which we expected would indicate aggression (observed physical aggression, derogatory speech, profanity, yelling), levels of social interaction (buying rounds, proximity between groups, circulation of people within the bar, interactions across groups), social cohesion (presence of regulars, patrons helping staff, staff and patrons sitting together, staff addressing patrons by first name), games of skill (darts, pool table, video games), music and dancing (dance floor, stage, brightness, music volume), and food and beverage availability (wine, beer and spirits availability and index beer costs). The means of continuous and Likert scale items were averaged across observers. Among dichotomous items, features of the physical and economic environments were considered observed if observers agreed those features were present. Due to the potentially transient nature of the behaviors involved, we considered social environment features to have been observed if either of the two observers indicated as such (e.g., profanity). We expected fixed features (i.e., physical, economic environment) to have greater reliability than impermanent features (i.e., social environment) which may change during the course of an observation.

Finally, we used principal components analyses to assess the internal reliability of these scales. We scaled all items in their original sets, removed items with loadings less than 0.30 and report here the proportion of variance in item responses explained by first and second principal components and provide eigenvalues for each item (Carmines & Zeller, 1979). The aggression, social interaction and social cohesion items loaded onto a common component composing a single scale of social disorder and are reported as such here.

RESULTS

Scouting Assessments

At least one scout attended all 165 bars. Comparing scouting assessments for only bars in which both scouts were able to complete data collection ($n = 128$; 77.6%), inter-observer reliability varied substantially between bar features (Table I). Some items which reflected the external physical environments of bars could be reliably identified (e.g., location in a downtown area) while others could not (e.g., location in a shopping mall). Even among those external features of bar environments that one could assume would be easily assessed reliability could be quite poor (e.g., location in a residential area). Similarly, we found poor

reliability on some items which reflected the interior environments of bars (e.g., the presence of a stage) while other items reflecting the overall size and condition of bar interiors (e.g., number of rooms) could be assessed with moderate reliability. Among all scouting assessments those addressing aspects of the economic environment were most reliably assessed. The exception, beer availability, appeared unreliable only because beer was almost universally available across places.

Unobtrusive Observations

Of the 112 bars eligible for unobtrusive observations, we were able to complete data collection in 97 bars, including 45 in high density areas, and 52 in low density areas. The remaining 15 bars had fewer than five patrons on both occasions when visited and were therefore ineligible for further observations. In all there were 220 pairs of observations in the 97 bars. The observers attended 70 bars (72.2%) twice, 25 bars (25.8%) three times, erroneously attended one bar a fourth time, and in one bar could complete only one visit.

The inter-observer reliabilities of items from unobtrusive observations are presented in Table II. As shown there, the majority of items demonstrated at least moderate reliability. Across all groups of items, those assessing individual behaviors (e.g., buying rounds) were least reliably assessed. On the other hand, all exterior and interior aspects of bar physical environments, and the majority of items reflecting the behaviors of groups of patrons (e.g., patron count, interactions across groups) and bar staff and their interactions, could be reliably assessed. On the whole, the more particular and specific an item was with regard to individuals and their behaviors within bars, the lower the reliability of that item. For example, numbers of bartenders and servers could be assessed with substantial and moderate reliability respectively, but observations of staff addressing patrons by first names could not be reliably assessed. Evidence that staff addressed patrons by first names was likely to be observed by one but not another observer so, despite relatively high levels of agreement (e.g., 74.2% for staff addressing patrons by first name), the impacts of disagreements between observers on reliability assessments was substantial.

Table III presents the results of principal components analyses of sets of items representing internal characteristics of bars. First and second components are shown to provide a sense of the degree to which the first component is dominant. In each case the first component accounts for more than twice as much variance as the second, indicating the generally univariate nature of each derived scale. We report eigenvalues to provide a sense of the meaning of each measure, and we include loadings (eigenvectors) on each second component as a guide to the interpretation of second components. As shown in the table, items representing the availability of food and beverages form a single dominant component indicated by the availability of spirits and wine and whether food was on offer. The first component accounts for 90.8% of the generalized variance among these measures and shows that the expansiveness of availability of spirits, wine, and food can be reliably assessed. The second, correspondingly weak, component accounts for only 9.2% of generalized variance among the items and reflects the availability of wine. Other principal component analyses indicate that activities within bars can be assessed with substantial reliability (i.e., music and dancing, games of skill). For social disorder the first principal component accounts for

62.9% of the generalized variance among items and shows that active bar environments (circulation of people within the bar, interaction across groups) may also be those places in which rowdiness, profanity, and signs of aggression occur.

DISCUSSION

This study demonstrates that it is possible to reliably characterize certain aspects of the physical, social, and economic environments of a representative sample of bars using scouting assessments and unobtrusive observations. In general we found that the physical and economic characteristics of these environments could be assessed with reasonable reliability and this was particularly true with respect to assessment of available activities within establishments and food and beverage availability. In contrast, measures of the social environment had comparatively low reliability. This would appear to be due at least in part to the fact that a number of the aspects of the social environment were infrequently observed and may have been witnessed by only one observer. While most physical features of these environments were relatively fixed, many social features of these environments were not. It is notable in this regard that greater reliability in the assessments of social characteristics of these places was obtained with items that assessed relatively fixed characteristics of social environments. Those items with greatest reliability were also those which assessed group aspects (e.g., patron count, interaction across groups, crowdedness, number of bartenders) rather than individual behaviors (e.g., derogatory speech). Thus, as a general rule, it appears that while both the fixed characteristics of these outlets and public activities of social groups in these outlets can be readily observed and reliably assessed, individual characteristics and specific social interactions cannot.

In addition to establishing that some individual items were reliable in a representative sample of bars, we also developed four scales that characterized collective aspects of bar environments: music and dancing facilities, games of skill, social disorder, and food and beverage availability. Principal components analyses demonstrated that for each scale, the first principal component accounted for most of the variance between bars, suggesting that the items that comprise the scales tend to co-occur within bars. For example, in bars where patron circulation was high and interaction across groups was common, observers were also likely to witness patron aggression, rowdiness and profanity. Because this study used a representative sample of bars, these four scales may represent features common to bars everywhere. That they can be characterized reliably is an important contribution.

Nine of the individual items in our instrument had direct analogues to measures used in prior studies; seven from Graham and colleagues' (2006) study of 118 large nightclubs in Toronto, and two items from Clapp and colleagues' (2007, 2009) study of 32 student bars in southern California. Our results were generally similar to those reported by Graham and colleagues: we had slightly better outcomes for noisiness ($r = 0.68$ vs $r = 0.65$), rowdiness ($r = 0.57$ vs $r = 0.40$), and circulation ($r = 0.64$ vs $r = 0.56$), but slightly poorer outcomes for crowdedness ($r = 0.74$ vs $r = 0.80$), patron count ($r = 0.78$ vs $r = 0.94$), and staff count ($r = 0.68$ vs $r = 0.87$). These count and Likert scale items characterized the barrooms in both studies with moderate to outstanding reliability. By comparison, compared to Clapp and colleagues, our observation of liberal pouring ($\kappa = 0.36$ vs $\kappa = 1.00$) and meal availability ($\kappa = 0.75$ vs $\kappa =$

1.00) had poorer reliability. There are two possible explanations for this difference. First, the convenience sample of student-oriented bars in the Clapp study may have reduced the variability of the observers' experiences across the study sample, reducing the range of behaviors which they were required to code. Second, the use of a short instrument (sixteen items) may have allowed the research staff to better focus on the variables of interest.

The reliability of drinking pace measures published in other studies are for two observers monitoring individual drinkers, so direct comparison with our novel measure is not possible. For example, Hennessey and Saltz (1993) reported correlation coefficients for observations of the amount of time that patrons were observed ($r = 0.97$) and the number of drinks they consumed while being observed ($r = 0.93$). Hunter, Hannon and Marchi (1982) reported 82% agreement for the number of sips taken to consume a beverage, and Aitken and Jahoda (1983) reported over 90% agreement for drink types and counts of drinks consumed per person over time. As our observers may have monitored different tables it is not possible to isolate single drinkers or drink counts. Our drinking pace measure is intended to characterize the overall pace of drinking within specific barroom environments. It appears to do so with moderate reliability ($r = 0.55$) given the possible variance for individuals within bars.

The results of this study indicate that it is possible to systematically assess even very small social environments. Previous sociological research hypothesizing the influence of neighborhood social disorganization on various health outcomes resulted in the development of reliable methods for systematically assessing a range of neighborhood environments (Raudenbush & Sampson, 1999). Systematic Social Observation and similar types of instruments have subsequently been utilized to assess spatial effects in small scale settings (Bader et al., 2015; Cunningham-Myrie et al., 2015; Kwarteng, Schulz, Mentz, Zenk, & Opperman, 2014; Parsons et al., 2010; Schaefer-McDaniel, O'Brien Caughy, O'Campo, & Gearey, 2010). Although studies have indicated the importance of other types of micro-environments on health outcomes --for example, homes (Sirard, Laska, Patnode, Farbachsh, & Lytle, 2010) or recreational facilities (Shores & West, 2008)—very few studies have attempted to systematically assess such micro-environments in a way that could be replicated. While the impacts of drinking and drug use settings on drug use behaviors have been long hypothesized, widely-applicable instruments to assess environmental features of barrooms have not been available.

Understanding relationships between bars and outcomes such as violent behaviors and drink driving and identifying those characteristics which may be altered to achieve public health benefits requires reliable measures of barroom environments. Importantly, studies that use unreliable measures may lead to erroneous results at worst, and at best make little meaningful contribution to the scientific literature. Some individual measures and the principal components scales presented here address the need for reliable measures that can be used in a generalizable sample of bars. They may be useful in future research conducted in any sample of bars that seeks to disaggregate relationships between bar characteristics and public health problems. Further, these indicators may provide guidance to communities aiming to reduce and prevent bar-associated public health problems.

ACKNOWLEDGEMENTS

We wish to acknowledge the valuable contributions of Andy Steinburg, Neil O'Callaghan, Tim Cochran, Zev Hardman, Craig Wingate, Dayna Fisher, Phoenix Jackson, Scott Martin, and Rachelle Annechino to data collection for this study.

FUNDING: This research was supported by National Institute of Alcohol Abuse and Alcoholism (NIAAA) Grant R01-AA019773, and NIAAA Center Grant P60-AA006282.

REFERENCES

- Aitken PP, Jahoda G. An observational study of young adults' drinking groups: Drink preferences, demographic and structural variables as predictors of alcohol consumption. *Alcohol and Alcoholism*. 1983; 18(2):135–150.
- Antin TMJ, Moore RS, Lee JP, Satterlund TD. Law in practice: Obstacles to a Smokefree Workplace Policy in bars serving Asian patrons. *Journal of Immigrant and Minority Health*. 2010; 12(2):221–227. [PubMed: 18712482]
- Bader MDM, Mooney SJ, Lee YJ, Sheehan D, Neckerman KM, Rundle AG, Teitler JO. Development and deployment of the Computer Assisted Neighborhood Visual Assessment System (CANVAS) to measure health-related neighborhood conditions. *Health & Place*. 2015; 31:163–172. [PubMed: 25545769]
- Bakeman R, Quera V, McArthur D, Robinson BF. Detecting Sequential Patterns and Determining Their Reliability With Fallible Observers. *Psychological Methods*. 1997; 2(4):357–370.
- Bernard, HR. Research methods in anthropology: qualitative and quantitative approaches. Rowman Altamira; Lanham, MD: 2006.
- Berning, A.; Compton, R.; Wochinger, K. Results of the 2013-2014 National Roadside Survey of Alcohol and Drug Use by Drivers. Office of Behavioral Safety Research; Washington, DC: 2015.
- Campbell CA, Hahn RA, Elder R, Brewer R, Chattopadhyay S, Fielding J, Middleton JC. The Effectiveness of Limiting Alcohol Outlet Density As a Means of Reducing Excessive Alcohol Consumption and Alcohol-Related Harms. *American Journal of Preventive Medicine*. 2009; 37(6): 556–569. [PubMed: 19944925]
- Carmines, EG.; Zeller, RA. Reliability and Validity Assessment. Sage Publications, Inc.; Beverly Hills, CA: 1979.
- Cavan, S. Liquor License: An Ethnography of Bar Behavior. Aldine Pub. Co.; Chicago, Ill: 1966.
- Clapp JD, Holmes MR, Reed MB, Shillington AM, Freisthler B, Lange JE. Measuring college students' alcohol consumption in natural drinking environments: field methodologies for bars and parties. *Evaluation Rev*. 2007; 31:469.
- Clapp JD, Reed MB, Min JW, Shillington AM, Croff JM, Holmes MR, Trim RS. Blood alcohol concentrations among bar patrons: A multi-level study of drinking behavior. *Drug Alcohol Depen*. 2009; 102(1-3):41–48.
- Cunningham-Myrie CA, Theall KP, Younger NO, Mabile EA, Tulloch-Reid MK, Francis DK, Wilks RJ. Associations between neighborhood effects and physical activity, obesity, and diabetes: The Jamaica Health and Lifestyle Survey 2008. *Journal Of Clinical Epidemiology*. 2015; 68(9):970–978. [PubMed: 25910912]
- Duff C, Rowland B. 'Rushing behind the wheel': Investigating the prevalence of 'drug driving' among club and rave patrons in Melbourne, Australia. *Drugs: Education, Prevention and Policy*. 2006; 13(4):299–312.
- Feinstein AR, Cicchetti DV. High agreement but low Kappa: I. the problems of two paradoxes. *Journal of Clinical Epidemiology*. 1990; 43(6):543–549. [PubMed: 2348207]
- Graham K, Bernards S, Osgood DW, Wells S. Bad nights or bad bars? Multi-level analysis of environmental predictors of aggression in late-night large-capacity bars and clubs. *Addiction*. 2006; 101:1569–1580. [PubMed: 17034436]
- Graham K, Bernards S, Osgood DW, Wells S. 'Hotspots' for aggression in licensed drinking venues. *Drug and Alcohol Review*. 2012; 31(4):377–384. [PubMed: 22050319]

- Graham, K.; Homel, R. *Raising the Bar: Preventing Aggression in and Around Bars, Pubs and Clubs*. Willan Publishing; United Kingdom: 2008.
- Graham K, La Rocque L, Yetman R, Ross TJ, Guistra E. Aggression and barroom environments. *Journal of Studies on Alcohol*. 1980; 41(3):277–292. [PubMed: 7374145]
- Graham K, Osgood W, Wells S, Stockwell T. To What Extent is Intoxication Associated With Aggression in Bars? A Multilevel Analysis. *Journal of Studies on Alcohol*. 2006; 67:382–390. [PubMed: 16608147]
- Graham K, Wells S. Aggression among young adults in the social context of the bar. *Addiction Research & Theory*. 2001; 9(3):193–219.
- Gruenewald PJ, Remer LG, LaScala EA. Testing a social ecological model of alcohol use: the California 50-city study. *Addiction*. 2013; 109:736–745. [PubMed: 24304295]
- Harford TC, Feinhandler SJ, O'Leary J, Dorman N. Drinking in bars: An observational study of companion status and drinking behavior. *International Journal of the Addictions*. 1983; 18(7):937–950. [PubMed: 6642800]
- Hennessy M, Saltz RF. Modeling social influences on public drinking. *Journal of Studies on Alcohol*. 1993; 54:139–145. [PubMed: 8459707]
- Holder HD, Saltz RF, Grube JW, Voas RB, Gruenewald PJ, Treno AJ. A community prevention trial to reduce alcohol-involved accidental injury and death: overview. *Addiction*. 1997; 92:S155–S171. [PubMed: 9231442]
- Homel R, Carvolth R, Hauritz M, McIlwain G, Teague R. Making licensed venues safer for patrons: what environmental factors should be the focus of interventions? *Drug and Alcohol Review*. 2004; 23(1):19–29. [PubMed: 14965884]
- Homel R, Thomsen S, Thommeny JL. Public drinking and violence: not just an alcohol problem. *Journal of Drug Issues*. 1992; 22:679–697.
- Hughes K, Anderson Z, Morleo M, Bellis MA. Alcohol, nightlife and violence: the relative contributions of drinking before and during nights out to negative health and criminal justice outcomes. *Addiction*. 2008; 103(1):60–65. [PubMed: 17996008]
- Hunter PA, Hannon R, Marchi D. Alcohol consumption in natural settings as a function of sex, age and income level. *Journal of Studies on Alcohol*. 1982; 43(2):387–392. [PubMed: 7121005]
- Ker K, Chinnock P. Interventions in the alcohol server setting for preventing injuries. *Cochrane Database of Systematic Reviews*. 2008; (3) Art. No.: CD005244.
- Kessler M, Gomberg CA. Observations of barroom drinking; methodology and preliminary results. *Quarterly Journal of Studies on Alcohol*. 1974; 35:1392–1396. [PubMed: 4445472]
- Kwarteng JL, Schulz AJ, Mentz GB, Zenk SN, Opperman AA. Associations between observed neighborhood characteristics and physical activity: findings from a multiethnic urban community. *Journal Of Public Health (Oxford, England)*. 2014; 36(3):358–367.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159–174. [PubMed: 843571]
- Lang E, Stockwell T, Rydon P, Lockwood A. Drinking settings and problems of intoxication. *Addiction Research*. 1995; 3(2):141–149.
- Lee JP, Antin TMJ, Moore RS. Social organization in bars: Implications for tobacco control policy. *Contemporary Drug Problems*. 2008; 35(1):59–98. 52. [PubMed: 22522904]
- Lee JP, Moore RS, Martin SE. Unobtrusive observations of smoking in urban California bars. *Journal of Drug Issues*. 2003; 33(4):983–999.
- Leonard KE, Quigley BM, Collins RL. Drinking, personality, and bar environmental characteristics as predictors of involvement in barroom aggression. *Addictive Behaviors*. 2003; 28(9):1681–1700. [PubMed: 14656553]
- Litwin, MW. *How to assess and interpret survey psychometrics*. Vol. 8. Sage Publications; Thousand Oaks: CA: 2002.
- Livingston M, Laslett A-M, Dietze P. Individual and community correlates of young people's high-risk drinking in Victoria, Australia. *Drug and Alcohol Dependence*. 2008; 98(3):241–248. [PubMed: 18640793]
- Mass-Observation. *The Pub and the People*. Victor Gollancz Ltd.; London: 1943.

- Moore R, Lee J, Martin S, Todd M, Chu BC. Correlates of Persistent Smoking in Bars Subject to Smokefree Workplace Policy. *International Journal of Environmental Research and Public Health*. 2009; 6(4):1341–1357. [PubMed: 19440522]
- Morrison C, Mair CF, Lee JP, Gruenewald PJ. A critical assessment of bias in survey studies using location-based sampling to recruit patrons in bars. *Subst Use Misuse*. 2015; 50:1427–1436. [PubMed: 26574657]
- Morrison C, Mair CF, Lee JP, Gruenewald PJ. Are barroom and neighborhood characteristics independently related to local-area assaults?. *Alcohol Clin Exper Res*. 2015; 39(12):2463–2470. [PubMed: 26756799]
- Nusbaumer MR, Reiling DM. Environmental Influences on Alcohol Consumption Practices of Alcoholic Beverage Servers. *The American Journal of Drug and Alcohol Abuse*. 2002; 28(4):733–742. [PubMed: 12492267]
- Parsons JA, Singh G, Scott AN, Nisenbaum R, Balasubramaniam P, Jabbar A, Dunn J. Standardized observation of neighbourhood disorder: does it work in Canada? *International Journal Of Health Geographics*. 2010; 9:6–6. [PubMed: 20146821]
- Ponicki WE, Gruenewald PJ, Remer L, Martin SE, Treno AJ. Assessing the validity of on-premise alcohol license data in six communities in California: operating characteristics and outlet densities. *Subst Use Misuse*. 2014; 49:51–58.
- Quigley B, Leonard K, Collins RL. Characteristics of Violent Bars and Bar Patrons. *Journal of Studies on Alcohol*. 2003; 64:765–772. [PubMed: 14743938]
- Quigley BM, Leonard KE. Alcohol use and violence among young adults. *Alcohol Research and Health*. 2004; 28(4):191–194.
- Raudenbush SW, Sampson RJ. *Ecometrics: Toward a Science of Assessing Ecological Settings, With Application to the Systematic Social Observation of Neighborhoods*. Sociological Methodology. 1999; 29(1):1.
- Roebuck, J.; Frese, W. *The Rendezvous; A Case Study of an After-hours Club*. The Free Press; New York: 1976.
- Roebuck, JB.; Murty, KS. *The Southern Subculture of Drinking and Driving: A Generalized Deviance Model for the Southern White Male*. Garland Publishing, Inc.; New York: 1996.
- Satterlund TD, Lee JP, Moore RS. Changes in smoking-related norms in bars resulting from California's smoke-free workplace act. *Journal of Drug Education*. 2012; 42(3):315–326. [PubMed: 23705511]
- Schaefer-McDaniel N, O'Brien Caughy M, O'Campo P, Gearey W. Examining methodological details of neighbourhood observations and the relationship to health: A literature review. *Social Science & Medicine*. 2010; 70(2):277–292. [PubMed: 19883966]
- Shores KA, West ST. The relationship between built park environments and physical activity in four park locations. *Journal Of Public Health Management And Practice*. 2008; 14(3):e9–e16. [PubMed: 18408541]
- Single E, Wortley S. Drinking in various settings as it relates to demographic variables and level of consumption: findings from a national survey in Canada. *Journal of Studies on Alcohol*. 1993; 54(5):590–599. [PubMed: 8412149]
- Sirard JR, Laska MN, Patnode CD, Farbaksh K, Lytle LA. Adolescent physical activity and screen time: Associations with the physical home environment. *The International Journal of Behavioral Nutrition and Physical Activity*. 2010; 7
- Snow R, Wells-Parker E. Relationships between drinking problems and drinking locations among convicted drinking drivers. *American Journal of Drug and Alcohol Abuse*. 2001; 27(3):531–542. [PubMed: 11506267]
- Snow RW, Anderson BJ. Drinking Place Selection Factors Among Drunk Drivers. *British Journal of Addiction*. 1987; 82(1):85–95. [PubMed: 3470046]
- Spradley, JP.; Mann, BJ. *The Cocktail Waitress: Woman's Work in a Man's World*. John Wiley & Sons, Inc.; New York: 1975.
- Stockwell T, Lang E, Rydon P. High risk drinking settings: the association of serving and promotional practices with harmful drinking. *Addiction*. 1993; 88(11):1519–1526. [PubMed: 8286997]

- Sykes R, Rowley R, Schaefer J. Effects of group participation on drinking behaviors in public bars: An observational survey. *Journal of Social Behavior and Personality*. 1990; 5(4):385–402.
- Sykes RE, Rowley RD, Schaefer JM. The influence of time, gender and group size on heavy drinking in public bars. *Journal of Studies on Alcohol*. 1993; 54:133–138. [PubMed: 8459706]
- Treno AJ, Grube JW, Martin SE. Alcohol Availability as a Predictor of Youth Drinking and Driving: A Hierarchical Analysis of Survey and Archival Data. *Alcoholism: Clinical and Experimental Research*. 2003; 27(5):835–840.
- Treno AJ, Gruenewald PJ, Lee JP, Remer LG. The Sacramento Neighborhood Alcohol Prevention Project: Outcomes From a Community Prevention Trial. *Journal of studies on Alcohol and Drugs*. 2007; 68:197–207. [PubMed: 17286338]
- Treno AJ, Gruenewald PJ, Remer LG, Johnson F, LaScala EA. Examining multilevel relationships between bars, hostility and aggression: social selection and social influence. [Article]. *Addiction*. 2008; 103(1):66–77. [PubMed: 18028523]
- Turrisi R, Nicholson B, Jaccard J. A Cognitive Analysis of Server Intervention Policies: Perceptions of Bar Owners and Servers. *Journal of Studies on Alcohol*. 1999; 60:37–46. [PubMed: 10096307]
- Tutenges S, Bøggkjær T, Witte M, Hesse M. Drunken Environments: A Survey of Bartenders Working in Pubs, Bars and Nightclubs. *International Journal of Environmental Research and Public Health*. 2013; 10(10):4896–4906. [PubMed: 24157511]
- Department of Justice. Criminal Victimization in the United States, 2005 *StatisticTables*. Bureau of Justice Statistics; Washington, D.C.: 2005.
- Voas, RB.; Lacey, JC. Alcohol and Highway Safety 2006: A Review of the State of Knowledge. Office of Behavioral Safety Research; Washington, D.C.: 2011.
- Warpenius K, Holmila M, Mustonen H. Effects of a community intervention to reduce the serving of alcohol to intoxicated patrons. *Addiction*. 2010; 105(6):1032–1040. [PubMed: 20219058]
- Weitzman ER, Folkman A, Kerry Lemieux Folkman MPH, Wechsler H. The relationship of alcohol outlet density to heavy and frequent drinking and drinking-related problems among college students at eight universities. *Health & Place*. 2003; 9(1):1–6. [PubMed: 12609468]
- World Health Organisation. Global Status Report on Alcohol and Health. World Health Organisation. , editor. Geneva, Switzerland: 2014.

Table 1

Reliability of scouting assessment items (128 bars)

Item	Item Type	Units	κ	Reliability %	r
Physical Environment (External)					
Parking lot	Binary	Yes/No	0.78	90.8	**
Other bars within sight	Binary	Yes/No	0.40	70.1	*
Location:					
Strip mall	Binary	Yes/No	0.62	81.5	*
Downtown	Binary	Yes/No	0.95	97.7	***
Shopping mall	Binary	Yes/No	0.31	90.8	
Residential	Binary	Yes/No	0.35	90.8	
Physical Environment (Internal)					
Stage	Binary	Yes/No	0.06	41.4	
Exits	Discrete	Count			0.11
Rooms	Discrete	Count			0.55 *
Estimated bartop length	Discrete	Feet			0.52 *
Conditions	Likert	0 (Poor) to 4 (Excellent)			0.59 *
Economic Environment					
Beer available	Binary	Yes/No	0.00	99.0	
Wine available	Binary	Yes/No	0.74	97.7	**
Spirits available	Binary	Yes/No	0.65	96.8	**
Index beer cost	Continuous	\$			0.85

Nb. Cohen's κ and percent agreement used for binary and categorical items; Pearson's r used for continuous, discrete and Likert items.

* Moderate reliability

** Substantial reliability

Outstanding reliability

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Inter-observer reliability for 220 paired unobtrusive observation items (97 bars)

Item	Item Type	Units	κ	Reliability %	r
Physical Environment					
<i>Exterior</i>					
Parking lot	Binary	Yes/No	0.80	92.17	***
Other bars within sight	Binary	Yes/No	0.55	79.82	*
<i>Interior</i>					
Board games	Binary	Yes/No	0.43	97.72	*
Dance floor	Binary	Yes/No	0.68	89.91	**
Darts	Binary	Yes/No	0.65	94.06	**
Karaoke	Binary	Yes/No	0.77	97.71	**
Pool table	Binary	Yes/No	0.90	95.41	***
Restaurant	Binary	Yes/No	0.68	85.32	**
Stage	Binary	Yes/No	0.84	96.35	***
Video games	Binary	Yes/No	0.51	85.78	*
Television	Binary	Yes/No	0.47	96.35	*
Ashtrays	Binary	Yes/No	0.66	97.69	**
<i>Lighting</i>					
Brightness	Likert	0 (low) to 10 (high)			0.64 **
Social Environment					
<i>Patron Characteristics</i>					
Predominant Gender:	Categorical		0.32	62.21	
Male					
Female					

Item	Item Type	Units	κ	Reliability %	r
None predominates					
Predominant Ethnicity:	Categorical		0.50	73.85	*
African American					
Asian					
Latino/a					
None predominates					
White					
Regulars					
Patron count	Discrete	Count			0.78 **
20 minute patron throughput count	Discrete	Count			0.48 *
<i>Patron Behavior</i>					
Aggression	Binary	Yes/No	0.45	90.78	*
Smoking	Binary	Yes/No	0.35	94.04	
Buying rounds	Binary	Yes/No	0.28	67.58	
Proximity between groups > arms' length	Binary	Yes/No	0.17	60.45	
Derogatory speech	Binary	Yes/No	0.12	87.67	
Profanity	Binary	Yes/No	0.40	79.91	*
Shoving	Binary	Yes/No	-0.01	97.72	
Yelling	Binary	Yes/No	0.25	71.89	
Sex trade	Binary	Yes/No	-0.02	96.8	
Circulation of people within bar	Likert	0 (low) to 10 (high)			0.64 **
Interaction across groups	Likert	0 (low) to 10 (high)			0.65 **
<i>Interior Atmosphere</i>					
Crowdedness	Likert	0 (low) to 10 (high)			0.74 **
Noisiness	Likert	0 (low) to 10 (high)			0.68 **
Rowdiness	Likert	0 (low) to 10 (high)			0.57 *
Music volume	Likert	0 (low) to 10 (high)			0.71 **

Item	Item Type	Units	κ	Reliability %	r
<i>Staff</i>					
Police presence	Binary	Yes/No	0.39	97.26	
Bouncer	Binary	Yes/No	0.64	90.87	**
Uniformed security	Binary	Yes/No	0.57	94.50	*
Smoking	Binary	Yes/No	0.52	96.77	*
Drinking	Binary	Yes/No	0.60	90.41	**
<i>Patron-Staff Interaction</i>					
Number of bartenders	Discrete	Count			0.68 **
Number of servers	Discrete	Count			0.45 *
Patrons helping staff	Binary	Yes/No	0.34	91.28	
Staff and patrons sitting together	Binary	Yes/No	0.43	86.76	*
Staff addressing patrons by first name	Binary	Yes/No	0.31	74.19	
Drinks on the house	Binary	Yes/No	0.47	93.48	*
Liberal pouring	Binary	Yes/No	0.36	70.33	
<i>Alcohol Consumption</i>					
Drinking pace	Continuous	Drinks/Person/Hour			0.55 *
Economic Environment					
Meals available	Binary	Yes/No	0.75	89.10	**
Beer available	Binary	Yes/No	0.00	98.16	
Wine available	Binary	Yes/No	0.41	91.24	*
Spirits available	Binary	Yes/No	0.64	95.37	**
Index beer available	Binary	Yes/No	0.46	83.64	*
Index beer cost	Continuous	\$			0.82 ***

Nb. Cohen's κ and percent agreement used for binary and categorical items; Pearson's r used for continuous, discrete and Likert items.

* Moderate reliability

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Outstanding reliability

Substantial reliability
**

Table 3

Factor loadings (eigenvectors) from principal components analysis for unobtrusive assessment items (97 bars)

Component	Music and Dancing		Games of Skill		Social Disorder		Food and Beverage	
	1	2	1	2	1	2	1	2
Physical Environment								
Brightness	-0.496	0.427						
Dance floor	0.533	0.141						
Music volume	0.515	-0.426						
Stage	0.453	0.785						
Darts			0.517	0.745				
Pool table			0.658	-0.031				
Video games			0.548	-0.666				
Social Environment								
Aggression					0.399	0.490		
Circulation of people within bar					0.504	-0.316		
Interaction across groups					0.447	-0.513		
Profanity					0.360	0.627		
Rowdiness					0.507	-0.064		
Economic Environment								
Meals available							0.594	-0.384
Spirits available							0.594	-0.384
Wine available							0.544	0.839
<i>Eigenvalue</i>	<i>2.569</i>	<i>0.661</i>	<i>1.682</i>	<i>0.838</i>	<i>3.146</i>	<i>0.895</i>	<i>2.723</i>	<i>0.277</i>
<i>Proportion of variance explained</i>	<i>0.642</i>	<i>0.165</i>	<i>0.561</i>	<i>0.280</i>	<i>0.629</i>	<i>0.179</i>	<i>0.908</i>	<i>0.092</i>