Validation of the Color of Drinking Survey instrument: A mixed methods study measuring the secondhand impacts of high-risk drinking in college settings.

Agustina Marconi, Reonda Washington, Amanda Jovaag, Ashley Knobeloch, Vilma Irazola, Carolina Muros Cortés, Laura Gutierrez, Courtney Blomme, Natalia Elorriaga

Submitted to: Interactive Journal of Medical Research
on: July 24, 2024

Table of Contents

Validation of the Color of Drinking Survey instrument: A mixed methods study measuring the secondhand impacts of high-risk drinking in college settings.

Agustina Marconi[1*] MD, MPH; Reonda Washington[1*] MPH, CHES; Amanda Jovaag[1*] MS; Ashley Knobeloch[1*] MPH, BNS, RN; Vilma Irazola[2*] MD, MSc; Carolina Muros Cortés[2*] MD; Laura Gutierrez[2*] MSc, BSc; Courtney Blomme[1*] MS, RD; Natalia Elorriaga[2*] PhD, MSc, LN

[1]UW- Madison Madison US
[2]Instituto de Efectividad Clinica (IECS) Dr. Emilio Ravignani 2024 Buenos Aires AR
[*]these authors contributed equally

Corresponding Author:
Agustina Marconi MD, MPH
UW- Madison
333 East Campus Mall
Madison
US

Abstract

Background: The “Color of Drinking” is an influential study in the alcohol consumption field that looks at secondhand harms of high-risk drinking on college students of color (microaggressions, fear of safety, harms in the classroom, and decreased belonging), and studies the connection between alcohol use and the campus racial climate. Since the release of the study findings in 2018, the Color of Drinking has received a lot of attention from other college settings, media coverage, and many requests of the research team to replicate the study around the country. As this instrument gained prominence, we decided to validate the Color of Drinking instrument.

Objective: This study aims to describe the development of the most recent version of the Color of Drinking Questionnaire and to assess its reliability and validity in a sample of undergraduate students attending UW-Madison.

Methods: Observational, analytic study that included both qualitative and quantitative approaches. We conducted in-depth cognitive interviews with students to evaluate comprehensibility and acceptability. Then, internal consistency, test-retest reliability and construct validity were assessed in a sample of UW-Madison undergraduate students. The revised version of the questionnaire was administered on two occasions. Internal consistency was evaluated for sets of items using data from the first administration. Test-retest reliability was evaluated by comparing the responses to the questionnaire administered at the beginning of the study and between 3 and 4 weeks later. Construct validity was assessed using data administrated at the beginning and other validated instruments administered at baseline.

Results: A total of 181 students completed the first administration of the questionnaire between June and November 2022. Of those, 177 responses were included for the analysis of internal consistency, 115 for test-retest reliability assessment and 98 for construct validity. The four dimensions evaluated, “impact of alcohol consumption on academics”, “impact of microaggressions”, “witnessing of microaggressions” and “alcohol intoxication and Bystanders’ interventions on alcohol intoxication” presented good internal consistency, with Cronbach’s Alpha coefficients ranging from 0.723 to 0.898. For the test-retest, the sections “Alcohol use”, “Areas avoided”, “Impact of by other students' alcohol consumption” and most items on the section “Alcohol culture and academics” showed moderate to substantial reliability. “Experiencing, witnessing and bystander intervention of microaggressions and alcohol intoxications” and most of the items from the section “Impact on health and sense of belonging” also showed moderate to substantial test-retest reliability. For the construct validity, correlations between the number of drinking days, the maximum number of drinks in a day and the Audit score as a continuous variable were moderate to high, r=0.630 (95% CI 0.533, 0.719) and r=0.647(95% CI 0.548, 0.741) respectively.

Conclusions: The tool was found to be valid and reliable in most dimensions. The three areas that were found to have lower reliability and validity were Alcohol and academics, Bystander intervention and Health impacts. Clinical Trial: NA

(JMIR Preprints 24/07/2024:64720)
DOI: https://doi.org/10.2196/preprints.64720

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Validation of the Color of Drinking Survey instrument: A mixed methods study measuring the secondhand impacts of high-risk drinking in college settings.

Agustina M. Marconi MD, MPH[1]; Reonda Washington MPH, CHES[1]; Amanda Jovaag, MS[1]; Courtney Blomme MS, RD[1]; Ashley Knobeloch MPH, BSN, RN[1]; Vilma Irazola MSc, MD[2]; Carolina Muros Cortés MD[2]; Laura Gutierrez MSc, BSc[2]; Natalia Elorriaga PHD, MSc, LN[2].

Abstract

Introduction: The “Color of Drinking” is an influential study in the alcohol consumption field that looks at secondhand harms of high-risk drinking on college students of color (microaggressions, fear of safety, harms in the classroom, and decreased belonging), and studies the connection between alcohol use and the campus racial climate. Since the release of the study findings in 2018, the Color of Drinking has received a lot of attention from other college settings, media coverage, and many requests of the research team to replicate the study around the country. As this instrument gained prominence, we decided to validate the Color of Drinking instrument. This study aims to describe the development of the most recent version of the Color of Drinking Questionnaire and to assess its reliability and validity in a sample of undergraduate students attending UW-Madison. Methods: Observational, analytic study that included both qualitative and quantitative approaches. We conducted in-depth cognitive interviews with students to evaluate comprehensibility and acceptability. Then, internal consistency, test-retest reliability and construct validity were assessed in a sample of UW-Madison undergraduate students. The revised version of the questionnaire was administered on two occasions. Internal consistency was evaluated for sets of items using data from the first administration. Test-retest reliability was evaluated by comparing the responses to the questionnaire administered at the beginning of the study and between 3 and 4 weeks later. Construct validity was assessed using data administrated at the beginning and other validated instruments administered at baseline. Results: A total of 181 students completed the first administration of the questionnaire between June and November 2022. Of those, 177 responses were included for the analysis of internal consistency, 115 for test-retest reliability assessment and 98 for construct validity. The four dimensions evaluated, “impact of alcohol consumption on academics”, “impact of microaggressions”, “witnessing of microaggressions” and “alcohol intoxication and Bystanders’ interventions on alcohol intoxication” presented good internal consistency, with Cronbach’s Alpha coefficients ranging from 0.723 to 0.898. For the test-retest, the sections “Alcohol use”, “Areas avoided”, “Impact of by other students' alcohol consumption” and most items on the section “Alcohol culture and academics” showed moderate to substantial reliability. “Experiencing, witnessing and bystander intervention of microaggressions and alcohol intoxications” and most of the items from the section “Impact on health and sense of belonging” also showed moderate to substantial test-retest reliability. For the construct validity, correlations between the number of drinking days, the maximum number of drinks in a day and the Audit score as a continuous variable were moderate to high, r=0.630 (95% CI 0.533, 0.719) and r=0.647(95% CI 0.548, 0.741) respectively. Conclusion: The tool was found to be valid and reliable in most dimensions. The three areas that were found to have lower reliability and validity were Alcohol and academics, Bystander intervention and Health impacts.

Key words: Validation Study; Alcohol drinking in College; Microaggression.

Introduction

[1] University Health Services (UHS), University of Wisconsin-Madison.
[2] Instituto de Efectividad Clínica y Sanitaria (IECS)

Alcohol during college life

The association between alcohol consumption and being a college student is broadly described in the literature. Drinking during college is often understood by students as part of their higher education experience and a way to socialize with peers [1]. The first years of college include major transitions in family bonds, living arrangements and socialization with peers. That type of transition is often related to risky behaviors, including alcohol consumption at high rates [2].

The Color of Drinking study

The “Color of Drinking” (COD) is an influential study in the alcohol consumption field that looks at secondhand harms of high-risk drinking on students of color (microaggressions, fear of safety, harms in the classroom, and decreased belonging), and connects alcohol use to racial climate [3]. COD examines differences in narratives and perspectives related to alcohol culture, high-risk drinking and racial climate for students of color and white students at a predominately white institution. A study questionnaire was designed by the UW-Madison team and was used to collect both quantitative and qualitative data. It was launched for the first time in 2015 and the initial administration included only students of color. In 2017, a second administration of the study included both white students and students of color. The questionnaire was delivered to 4,736 undergraduate students of color and 2,000 white students between November 2017 and January 2018. A student of color was defined as someone who did not self-identify solely as white. The response rates were 25.2% (N = 1,195) for undergraduate students of color and 16.7% (N=333) for white undergraduate students.

The main findings of the 2017/2018 cohort were the following: Safety for students of color in high-risk drinking environments is impacted more than their white peers, high-risk drinking is positively tied to student connection and belonging; Students of color reported higher rates of abstaining and non-drinking than white students (15.3% vs 8.4%); African American/Black students considered leaving UW–Madison at twice the rate of white students (43.6% vs 20.9%); Students of color are more likely to report being “financially struggling” (14.5% vs 7%); White, problematic drinkers scored higher on the Diener Flourishing Scale than students of color (27.3% vs 10.2%); Both groups commonly reported having to find alternative study spaces because of the alcohol culture (46%), and scheduled group meetings around alcohol consumption (40%).

During 2020 the COD team published an in-depth analysis of racial microaggressions and alcohol use in this sample. Results showed African-American/Black was the racial group that witnessed (82.9%) and suffered more microaggressions 79.8%). The percentage of students witnessing microaggressions increased with the year in school (49.3% in first year vs 75.8% in fourth year students). Alcohol use in the last 30 days, feeling impacted by other’s use of alcohol, and avoiding certain areas due to alcohol consumption were significantly associated with experiencing microaggressions among students of color [4].

Since the release of the study findings in 2018/2019, the Color of Drinking has received a lot of attention from other college settings, media coverage, and many requests of the research team to replicate the study around the country [5]. As this instrument gained greater prominence, both as a research tool and as a way for campuses to understand their student experiences, the decision was made to undertake a validation study of the Color of Drinking instrument. This exploration of the tool allows UW-Madison to better interpret results from any administration of the survey, and to clearly demonstrate to other campuses what they will learn from their own fielding of the survey.

This study aims to describe the development of the last version of the Color of Drinking

Questionnaire and to assess its reliability and validity in a sample of undergraduate students attending UW-Madison.

Methods

Study Design

In 2022 we developed the COD validation protocol. This study was an observational, analytic study that included bot. First, the original version of the questionnaire was reviewed by experts to assess content validity. Second, we conducted in-depth cognitive interviews with students to evaluate comprehensibility and acceptability. Then, the internal consistency, test-retest reliability and construct validity of the tool was assessed in a sample of UW-Madison undergraduate students. The revised version of the questionnaire was administered on two occasions. Internal consistency was evaluated for sets of items using data from the first administration. Test-retest reliability was evaluated by comparing the responses to the questionnaire administered at the beginning of the study (COD1) and between 3 and 4 weeks later (COD2). Construct validity was assessed using data from the COD1 and other validated instruments administered at baseline.

Participants

The study was conducted in an opportunistic sample of undergraduate students ≥ 18 years old at UW-Madison, ensuring variability in age, gender, and ethnicity/ race among students with Freshmen, Sophomore, Junior and Senior standing [6]. Students were invited to participate by e-mail. Five hundred invitations were sent for the quantitative phase, and interested students were also individually invited to complete the survey.

The Color of Drinking Questionnaire (COD)

The instrument was developed by the University Health Services (UHS) at the University of Wisconsin- Madison, based on information gathered during interviews with students. To ensure face and content validity, the questionnaire was revised by experts in the field and modified according to their indications. The instrument explores the following dimensions: a) Alcohol use and culture, b) Academics and alcohol culture (general, own use and drinking culture), c) Locations (areas avoided due to concerns of the alcohol use of others or where feeling unsafe), d) The impact of others' alcohol use, e) Microaggressions (Students of color only), f) Bystander intervention (witnessing and intervention), g) Impact on personal health and wellbeing and h) Overall experience and sense of belonging at the campus. The tool also includes questions about socio-demographic characteristics such as age, gender, race/ethnicity and financial situation. The questionnaire included both open-ended and closed-ended questions. Questions about experiencing microaggressions were only asked of students of color, due to the definition of microaggressions used by the authors: "Racial microaggressions are an everyday manifestation of oppression that brings psychological consequences to target groups" [7]. The wellbeing section includes a validated scale, the Diener Flourishing Scale, consisting of 8 items [8].

During the first phase of the validation study, the original version was revised by a group of researchers with expertise in questionnaire design obtaining a preliminary version. Later, a series of cognitive interviews were carried out with a sample of eight students of 19 to 22 years, ensuring variability in gender, ethnicity/race, and stage of academic study. They consisted of the flexible administration of the draft questionnaire and relied heavily on verbal probing to improve the instrument [9,10]. An "inspect-and-repair" model was used to inspect the items, detect flaws, and

then repair them. If the interviewer identified any problem with one or more of the questions (about comprehension, recall, difficulty to answer, response, or other) they would rephrase the items to test a potential solution [11]. The interviews were focused on the set of questions developed for the questionnaire and did not include the already validated Diener Flourishing Scale nor the section on sociodemographic characteristics. After the series of interviews, a definitive version of the questionnaire was produced [8].

The final version consists of 31 items (19 core questions + 12 skip questions). In addition, the survey includes 12 questions about sociodemographic characteristics.

Assessment of the psychometric properties

Internal consistency. This property is a measure of the inter-correlation among items and hence the consistency in the measurement of the intended construct [12]. The internal consistency was evaluated in the following sets of items:

-Impact of alcohol consumption on academics (Q8. How often have you experienced the following during the current semester? items: 8.1 to 8.5)

-Impact of microaggressions (Q17. How much did the microaggressions impact your… items: 17.1 to 17.3?)

-Witnessed microaggressions (Q25. Have you witnessed any of the following? ...items: 25.1 to 25.8)

-Bystander intervention (Q26. Would you intervene in the following situations? ...items: 26.1 to 26.4)

Test-retest reliability. Test-retest reliability, which is a measurement of reproducibility, was evaluated for closed-ended questions.

Construct validity. In order to assess construct validity, participants also completed three validated scales: the Alcohol Use Disorders Identification Test (AUDIT), the “Sense of Belonging Index” (SBI) and the “Workplace and School Microaggressions” Scale, from the Racial and Ethnic Microaggressions Scale (REMS) [13,14,15]. They also reported their Grade Point Average (GPA).

The AUDIT is a 10-item questionnaire with a range of possible scores from 0 to 40. Scores ≥7 and 8 have been proposed as “Hazardous and harmful alcohol use” among women and men, respectively [13, 16]. It was expected that a) the median number of drinking days and the maximum number of drinks in a day during the last 2 weeks measured by the COD questionnaire are greater among those scoring ≥7 than those scoring <7 in the AUDIT instrument, and b) the number of drinking days and the maximum number of drinks per day during the last 2 weeks are correlated with the AUDIT score.

The validated “Workplace and School Microaggressions” sub-scale, from the REMS included 5 items describing experienced microaggressions in the last 6 months, with possible score ranges from 0 to 5. It was expected that those who answered that they had experienced a microaggression on campus or surrounding areas (COD, Q14) score higher in the REMS subscale than those who answered that they had not such experience [15].

The SBI is a validated 6-item questionnaire with possible scores ranging from 5 to 30 (the greater the score, the more sense of belonging). It was hypothesized that the SBI score is lower among those who reported that the alcohol culture had impacted their overall sense of belonging at UW-Madison

in comparison to those that reported that the alcohol culture had not impacted it [14].

Finally, the academic performance was operationalized as self-reported accumulated and last-term GPA. It was expected that academic performance is better among those students who responded that they never received a poor grade because they had chosen to drink instead of study during the semester.

Data analysis

Descriptive statistics were calculated for the population and for each item. Hypothesis tests were selected according to the variable distribution. Wilcoxon rank sum test with continuity correction and Kendall correlation coefficients were used to test the hypotheses related to construct validity.

The internal consistency was evaluated by calculating the Cronbach's Alpha coefficient by dimension. Additionally, we calculated the variation in the Alpha coefficient after excluding each item from the analysis to assess the effect of dropping a particular item. A Cronbach's alpha value of at least 0.70 is usually interpreted as adequate internal consistency [17]. Low values of Cronbach's alpha represent poor inter-relatedness between items.

To assess test-retest reliability, the Cohen's Kappa coefficient (dichotomous data), Cohen's weighted Kappa coefficient (ordinal data) and the Intra-class correlation coefficient (ICC), for continuous data) and their 95% confidence interval were calculated. The higher the coefficient, the stronger the test-retest reliability [18]. Kappa values between 0.41 and 0.60 were considered as moderate reliability, between 0.61 and 0.80 as substantial reliability and between 0.81 and 1.00 very high reliability [10]. Kappa is influenced by the trait prevalence and basal rates. Kappa may be low even though there are high levels of agreement and even though individual ratings are accurate, thus we have also calculated the proportion of observed and expected agreement only by chance. For the interpretation of the Intraclass correlation coefficient, values lower than 0.5 were considered indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 good reliability, and values higher than 0.90 excellent reliability [19].

All the analyses considered alpha < 0.05 as statistically significant. Analyses were performed using R Statistical Software (v4.1.3; R Core Team 2022) and RStudio statistical software version 2022.02.3 for Windows.

Ethics

This study was approved by the University Institutional Review Board (IRB).

The study was conducted according to the guidelines established in the Declaration of Helsinki for studies involving human subjects. Written informed consent was obtained from all participants.

Results

A total of 181 students completed the first administration of the questionnaire between June and November 2022. Of these, 116 also completed the second administration. Six observations were excluded from the analysis due to missing answers in >50% of the questions (4 from the first administration and 2 from the second administration. All of them were from the group of students of color. A total of 177 responses were available for the analysis of internal consistency, 115 for test-retest reliability assessment and 98 for construct validity.

Participants' characteristics for each sub-study are shown below in Table 1. Students of color represented 48,6%, 43,5% and 57,1% in sub-studies of assessment of Internal consistency, test-retest assessment, and construct validity, respectively. Most of the students of color were Asian or Asian American, followed by Hispanic, Latinx or Spanish origin. There were students representing all stages of academic studies. A larger number of students lived off campus, and the most frequent current financial situation was described as "tight but doing fine".

Table 1

Internal consistency.

The four dimensions evaluated (*Impact of alcohol consumption on academics, Impact of microaggressions, witnessing of microaggressions and Alcohol intoxication and Bystanders' interventions on alcohol intoxication)* presented good internal consistency, with Cronbach's Alpha coefficients ranging from 0.723 to 0.898 (Table 2).

Table 2

Test-retest reliability

The results from the analyses of test-retest reliability are shown in Table S1 and Table S2. Approximately 33% and 30% of the students reported that they do not drink alcohol or did not drink during the last 30 days in test and retest surveys, respectively. The mean number of drinking days within the last two weeks was 1.4 and 1.8 for test and retest, respectively. In summary, the section "Alcohol use" showed moderate to substantial reliability (Table S1).

On test and retest surveys, 9.6% of the participants reported that the alcohol culture at UW-Madison had impacted their academics (Table S2). Most items on the section "Alcohol culture and academics" showed moderate to substantial test-retest reliability, but one of them showed low test-retest reliability.

On test and retest surveys, 51.8% of the participants reported that during the semester they had avoided specific areas on or off campus due to concerns of alcohol use by others. The proportion in which participants were negatively impacted by other students' alcohol consumption was 35% for test and 28% for retest. Items from these two sections, "Areas avoided" and "Impact of by other students' alcohol consumption" also showed mostly moderate to substantial test-retest reliability (Table S2).

Among the students of color, 46.8% and 40.3% reported having experienced a microaggression on campus or surrounding areas on test and retest surveys. On test and retest surveys, 71% and 73% of the students had witnessed at least once a microaggression on campus or surrounding areas during their time at UW-Madison. A summary of test-retest reliability assessment for experiencing, witnessing and bystander intervention of microaggressions and alcohol intoxications is presented in Table S2. Most of the items from these dimensions showed moderate to substantial test-retest reliability.

From test and retest surveys, 33% and 25% of the participants answered that their own or alcohol consumption by other students impacted their health in the time they have attended UW-Madison. At test and retest, 43% and 42% of the students reported that the alcohol culture had impacted their overall sense of belonging at UW-Madison. Most of the items from the section "Impact on health and

sense of belonging" showed moderate to substantial test-retest reliability, except for one (Table S2).

Supplementary material Table S1 and Table S2

Construct validity

Results from Q5 were directly associated with the AUDIT score. Students with AUDIT scores ≥7 reported a higher number of drinking days within the last two weeks than those with lower AUDIT scores. A similar association was found regarding the maximum number of drinks in a day and the AUDIT score with a cut-off point of 7 (Table 3). Correlations between the number of drinking days, the maximum number of drinks in a day and the Audit score as a continuous variable were moderate to high, r=0.630 (95% CI 0.533, 0.719) and r=0.647(95% CI 0.548, 0.741) respectively.

A few students reported that they received a poor grade because they had chosen to drink instead of study. Their GPA was lower than the other students (Table 4).

In addition, 27 out of 67 students of color answered that in their time at UW-Madison they had experienced a microaggression on campus or surrounding areas. They scored higher in the Racial and Ethnic Microaggression Scale than the rest of the students.

Lastly, 41 out of 98 students answered that the alcohol culture impacted their overall sense of belonging at UW-Madison. These students scored lower in the Sense of Belonging Index in comparison with the rest of the students (Table 4).

Table 3

Table 4

Discussion

The validation process outlined here included the assessment of reliability and validity of the Color of Drinking questionnaire, used in an exploratory study of the impact of UW–Madison's alcohol culture on students of color [3]. The tool was found to be valid and reliable in most dimensions. The three areas that were found to have lower reliability and validity were Alcohol and academics, Bystander intervention and Health impacts.

In the Alcohol and academics dimension, we found that questions related to personal academic consequences had higher internal consistency and a higher Cronbach's Alpha than the other questions in the same dimension. Questions about the secondhand negative academic consequences were found to be highly relevant but not highly correlated with each other. Question 8.5 "My professors or TAs invited me out to places where alcohol is served" tested poorly when placed with the other secondhand academic consequences. For the finalized questionnaire we decided to separate the personal academic consequences, and secondhand negative consequences as two separate subdomains. We also removed question 8.5 due to some unclearness about the scope and meaning of the item identified during the qualitative interview and little consistency with other items in the expected subdomain.

In the bystander intervention dimension, there were issues with test-retest reliability among 4 questions. Bystander discrimination questions from the questionnaire had low test-retest reliability. The questions were more prone to variation over time than the other bystander questions. The

decision was made to remove these 4 questions from the finalized version of the questionnaire. Other validation processes in college settings show adequate validity and reliability when assessing other types of bystander interventions [20].

The health impacts section showed a low test-retest reliability. When the qualitative responses were evaluated, students brought up different interpretations, but also very profound answers and concepts such as mental health struggles or experiencing sexual assault. Based on these results, the decision was made to rewrite and retest the validity of this question grounded on the evaluation of content validity.

Questionnaires are widely used instruments utilized to collect data in different fields [21, 22]. A questionnaire designed for populational health studies should capture what it was meant to collect when designed in a reliable manner. The validation process will not only allow us to measure what we want to measure, but also will make our data comparable over time and with other college settings [23].

References

1. National Institute on Alcohol Abuse and Alcoholism. Alcohol's effects on health; harmful and underage college drinking. March 2023. Accessed November 12, 2023. https://www.niaaa.nih.gov/publications/brochures-and-fact-sheets/college-drinking
2. Lorant V, Nicaise P, Soto VE, d'Hoore W. Alcohol drinking among college students: college responsibility for personal troubles. *BMC Public Health.* 2013;13(615). https://doi.org/10.1186/1471-2458-13-615
3. Washington R, Marconi A, Reeves M, Jardas E. The color of drinking: an exploratory study of the impact of UW–Madison's alcohol culture on students of color. University Health Services. 2020. Accessed September 1, 2023. https://www.uhs.wisc.edu/wp-content/uploads/2020/01/10_1_2018_ColorOfDrinkingBooklet_RLG.pdf
4. Marconi A, Washington R, Reeves M, Bradley Q, Ayala A, Griggs C. Examining racial microaggressions and alcohol use among marginalized populations at a predominately white institution. *J Am Coll Health*. 2022. doi:10.1080/07448481.2022.2098027
5. Rosciglione A. 'If you don't drink, you don't belong': a deep dive into UW's Color of Drinking survey. The Daily Cardinal. November 10, 2022. Accessed September 1, 2023. https://www.dailycardinal.com/article/2022/11/if-you-dont-drink-you-dont-belong-a-deep-dive-into-uws-color-of-drinking-survey
6. MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. *Psychol Methods.* 1999;4(1):84–99. https://doi.org/10.1037/1082-989X.4.1.84
7. Sue DW, Capodilupo CM, Holder AMB. Racial microaggressions in the life experience of Black Americans. *Prof Psychol Res Pract*. 2008;39(3):329–336. doi:10.1037/0735-7028.39.3.329
8. Diener, E, Wirtz D, Tov W, et al. New well-being measures: short scales to assess flourishing and positive and negative feelings. *Soc Indic Res.* 2010;97(2):143–156. doi:10.1007/s11205-009-9493-y
9. Beatty PC, Willis GB. Research synthesis: the practice of cognitive interviewing. *Public Opin. Q*. 2007;71(2):287–311. doi:10.1093/poq/nfm006
10. National Research Council. Cognitive aspects of survey methodology: building a bridge between disciplines. Washington, DC: The National Academies Press; 1984. doi:10.17226/930
11. Willis GB. *Analysis of the Cognitive Interview in Questionnaire Design (Understanding Qualitative Research).* Oxford University Press; 2015.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159–174. https://doi.org/10.2307/2529310
13. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the alcohol use

disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. *Addiction.* 1993;88(6):791–804. doi:10.1111/j.1360-0443.1993.tb02093.x
14. Pedler ML, Willis R, Nieuwoudt JE. A sense of belonging at university: student retention, motivation and enjoyment. *Journal of Further and Higher Education*. 2022;46(3):397–408. doi:10.1080/0309877X.2021.1955844
15. Nadal KL. The Racial and Ethnic Microaggressions Scale (REMS): construction, reliability, and validity. *J Couns Psychol.* 2011;58(4):470–480. doi:10.1037/a0025193
16. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG, de la Salud OM. AUDIT: cuestionario de identificación de los trastornos debidos al consumo de alcohol: pautas para su utilización en atención primaria. 2001. Accessed September 1, 2023. https://apps.who.int/iris/handle/10665/331321
17. Nunnally JC. *Psychometric Theory.* 2nd ed. McGraw-Hill; 1978.
18. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use.* 4th ed. Oxford University Press; 2008.
19. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155–163. doi:10.1016/j.jcm.2016.02.012
20. Banyard VL, Plante EG, Moynihan MM. Rape prevention through bystander education: bringing a broader community perspective to sexual violence prevention. February 2005. Accessed September 1, 2023. https://www.ncjrs.gov/pdffiles1/nij/grants/208701.pdf
21. Taherdoost H. Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. *IJARM.* 2016;5(3):28–36. http://dx.doi.org/10.2139/ssrn.3205040
22. Aithal A, Aithal PS. Development and validation of survey questionnaire & experimental data – a systematical review-based statistical approach. *IJMTS.* 2020;5(2):233–251. doi:10.47992/IJMTS.2581.6012.0116
23. Kazi A, Khalid W. Questionnaire designing and validation. *J Pak Med Assoc*. 2012;62(5):514-516. http://ecommons.aku.edu/pakistan_fhs_mc_women_childhealth_paediatr/14

Supplementary Files