Experimental and Clinical Psychopharmacology

Naturalistic Substance Use Before/During MTurk Research Participation Is Associated With Increased Substance Demand and Craving

Shahar Almog, Liana S. E. Hone, Chiara M. Licata, Jillian M. Rung, and Meredith S. Berry

Online First Publication, October 10, 2024. https://dx.doi.org/10.1037/pha0000743

AMERICAN PSYCHOLOGICAL ASSOCIATION

Experimental and Clinical Psychopharmacology

ISSN: 1064-1297

https://doi.org/10.1037/pha0000743

Naturalistic Substance Use Before/During MTurk Research Participation Is Associated With Increased Substance Demand and Craving

Shahar Almog[1], Liana S. E. Hone[1], Chiara M. Licata[1], Jillian M. Rung[1, 2], and Meredith S. Berry[1, 3]
[1] Department of Health Education and Behavior, University of Florida
[2] Department of Epidemiology, University of Florida
[3] Department of Psychology, University of Florida

Although crowdsourcing platforms are widely used in substance-use research, it is unclear what percentage of participants use substances at the time of participation and how this might affect data quality, behavioral outcomes, or decision making. We conducted a secondary analysis of data collected on MTurk for a two-session, within-subject experiment recruiting individuals who regularly use alcohol, cannabis, cigarettes, or opioids. We analyzed 527 observations collected across two sessions (Session 1: $n = 303$, Session 2: $n = 224$) on measures of substance use before (within 3 hr)/during participation, data quality, demand in hypothetical purchase tasks, delay discounting, and craving. Substance use before/during participation was common (35.7%). Some participants reported substance use before/during both (25.4%) or only one (20.1%) of the sessions. Between-subject analyses of the first session data revealed that participants who used substances before/during participation did not differ on quality measures yet were slower to complete the survey. Controlling for individual differences in demographic variables and typical substance use, using a substance before/during participation was associated with increased hypothetical consumption of substances when the substance was free (demand intensity) and higher craving for substances, but not delay discounting. Substance use before/during MTurk participation among individuals who regularly use substances is prevalent and may impact outcome measures or standardization across sessions in repeated measures designs. Several implications have emerged, including statistically or experimentally controlling for substance use occurring before/during participation, which could improve the validity and rigor of online substance use research, and should be considered a part of best practices.

Public Health Significance
The present study showed that substance use just before/during participation in online remote research is prevalent among individuals who regularly use substances. Substance use before/during participation was associated with increased hypothetical drug consumption and craving for substances. These results inform specific recommendations to statistically or experimentally control for substance use occurring before/during remote research, to improve validity and rigor of online substance use and addiction research.

Keywords: MTurk, acute substance use, behavioral economic demand, delay discounting, craving

Supplemental materials: https://doi.org/10.1037/pha0000743.supp

Kelly E. Dunn served as action editor.
Shahar Almog https://orcid.org/0000-0003-3093-3947
Liana S. E. Hone https://orcid.org/0000-0002-6777-978X
Jillian M. Rung https://orcid.org/0000-0002-5945-6454
Meredith S. Berry https://orcid.org/0000-0002-5526-7437
Jillian M. Rung is now at Highmark Health, Pittsburgh, Pennsylvania, United States. All authors contributed in a significant way to the article, and all authors have read and approved the final article. This study was not preregistered. Materials and analysis code for this study are available by emailing Shahar Almog or Meredith S. Berry. Portions of this study were presented at the 2023 Annual Meeting of the American Psychological Association. The authors have no conflicts of interest to disclose. This research was supported in part by the National Institute on Drug Abuse Grant R21DA056813 awarded to Meredith S. Berry. Meredith S. Berry gratefully acknowledges that her time was also supported in part by the National Institute on Drug Abuse Grant K01DA052673. This research was also supported by the National Institute on Alcohol Abuse and Alcoholism Grant L30AA027013 awarded to Liana S. E. Hone. Support for Jillian M. Rung was also previously provided by the National Institute on Drug Abuse Grant T32DA035167. The funding sources had no role other than financial support.

Shahar Almog played a lead role in formal analysis, investigation, project administration, visualization, and writing–original draft and an equal role in conceptualization, data curation, methodology, software, and validation. Liana S. E. Hone played a supporting role in writing–review and editing and an equal role in conceptualization. Chiara M. Licata played a supporting role in investigation, software, and writing–review and editing. Jillian M. Rung played a supporting role in data curation, formal analysis, software, and validation and an equal role in methodology and writing–review and editing. Meredith S. Berry played a lead role in funding acquisition, resources, supervision, and writing–review and editing, a supporting role in visualization and writing–original draft, and an equal role in conceptualization, methodology, and validation.

Correspondence concerning this article should be addressed to Shahar Almog or Meredith S. Berry, Department of Health Education and Behavior, University of Florida, PO Box 118210, Gainesville, FL 32611, United States. Email: almogs@ufl.edu or mberry@ufl.edu

The use of crowdsourcing platforms generally and for substance use research specifically has markedly increased in the past decade and, especially, during the COVID-19 pandemic (Arditte et al., 2016; Paolacci & Chandler, 2014; Stanton et al., 2022). Crowdsourcing platforms such as Amazon Mechanical Turk (MTurk) and Prolific allow for cost-efficiency, increased sample diversity, access to hard-to-reach populations, and wider geographic areas of data collection (Strickland & Stoops, 2019; Strickland et al., 2022). Given the limited control over administration conditions and increased use of these platforms, general best practices for psychological research have been developed (e.g., screening recommendations, use of attention checks), which facilitate higher quality data collection. However, information on and best practices for specific populations, such as those who regularly use substances, appears to be lacking.

Limited data exist regarding rates and effects of naturalistic substance use immediately before and during survey completion to guide best practices, particularly among regular substance-using populations who may be more likely to use substances in uncontrolled conditions. Such data may be especially important given acute drug consumption and intoxication can influence attention and decision making on various tasks, including behavioral economic tasks frequently employed in substance use research. For example, acute alcohol consumption may result in increased behavioral economic demand and craving for alcohol among moderate drinkers (e.g., Amlung et al., 2015; Motschman et al., 2022). Similarly, alcohol administration reduced the likelihood of condom use with increasing delay to condom availability on the Sexual Delay Discounting Task (e.g., P. S. Johnson et al., 2016). However, alcohol administration may, or may not, affect monetary delay discounting as the literature suggests mixed results (Bidwell et al., 2013; Reed et al., 2012; Reynolds et al., 2006). Regarding cannabis, laboratory administration studies suggested that cannabis administration decreased demand and craving for cannabis immediately following the administration compared to placebo control (demand and craving: Hindocha et al., 2017, craving: Metrik et al., 2015). The converse may also be true, in which deprivation from regularly used substances can potentially influence decision making. For example, nicotine deprivation might affect delay discounting (e.g., Field et al., 2006).

These considerations are further underscored by the potential of current intoxication, substance deprivation, and/or withdrawal to affect outcomes which may be differentially impacted across various study designs. Current substance use or deprivation not accounted for could contaminate group-based inferences or associations between variables in cross-sectional studies, or interfere with condition effects and response consistency in within-subjects studies. Therefore, data on prevalence of substance use immediately before or during survey completion and the effect on behavioral and decision-making tasks in online samples will be critical to inform specific best practices for substance use researchers. Findings will inform the potential need to experimentally or statistically control for acute substance use or intoxication and for specific instructions provided to the participants that may aim to standardize substance use before or during participation.

Thus, the purpose of the present secondary analysis was to characterize rates of substance use within 3 hr prior to and during survey completion among MTurk participants who reported regular use of alcohol, cannabis, cigarettes, or opioids. Specifically, we aimed to (a) describe substance use before/during participation overall, across repeated assessments, and per specific substance-based subsamples; (b) characterize and compare demographics, data quality, and typical substance use across participants who used a substance before/during participation versus participants who did not use any substance before/during participation; and (c) determine the effects of naturalistic substance use before/during participation (vs. none) on behavioral economic demand for substances, delay discounting, and craving for substances. We then discuss the implications and recommend best practices based on these findings.

Method

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the present study. This secondary analysis was not preregistered. Materials and analysis code for this study are available by emailing Shahar Almog or Meredith S. Berry.

Participants and Procedure

Participants were recruited on MTurk from September 2021 to September 2022. The parent study was a within-subject repeated measures experimental study. It consisted of one screener and two experimental sessions separated by a minimum of 5 days, and each participant completed both sessions. MTurk workers, age 18 and older, who reported U.S. residency per MTurk account, with a history of at least 100 approved tasks and 95% approval rate, were first screened by a separate screener designed to identify individuals who demonstrated adequate English proficiency and reported regular substance use (i.e., alcohol, cannabis, nicotine cigarettes, opioids). The parent study aimed to assess behavioral economic measures across a wide substance-using sample and thus focused on several commonly used substances with associated well-characterized behavioral economic purchase tasks and common methods of administration (e.g., for nicotine, combustible cigarettes). Regular substance use was defined as consumption of at least 10 times in the past month for alcohol or cigarettes and five times for cannabis or opioids. MTurk workers who passed all qualifications were able to see the first of two sessions. In the first session, participants were asked to choose the substance they used most frequently. Based on their choice of most frequently used substance (e.g., cannabis), participants were presented with the appropriate questions and tasks (e.g., a cannabis purchase task).

In the first session, participants were randomized to one of two experimental conditions, in which they viewed either natural or built environment images on the computer screen. The second session was completed at least 5 days after the first and was identical in procedure with the exception of the stimuli (e.g., if the participant viewed natural environment images in the first session, they would view built environment images in the second session). A first set of 13 images was presented full screen for approximately 5 min before the delay discounting task, and a second set of another six images was presented for 2 min before the drug purchase task.

The survey started with questions on state craving, followed by the slideshows of the stimuli (either nature or built stimuli, depending on the session), the delay discounting and drug purchase tasks, three scales (i.e., self-compassion, empathy, nature relatedness, data are not relevant to these analyses and therefore not presented here), questions on typical substance use and substance use before and during the survey, and demographic questions. Only participants who met our definition of regular use, adhered to study procedures (e.g., fully watched the slideshows, based on Qualtrics time stamps recording durations), and provided systematic delay discounting or systematic demand data were invited to the second session. The compensation was $0.15 for the screener and $3.50 for each study session (total of $7.15). The parent study aimed to have a final data set of 50 participants per substance type (i.e., alcohol, cannabis, cigarettes, opioids). Despite recruiting efforts, the opioid subgroup remained extremely small ($n = 11$), and as a result, data collection for this group was discontinued. In the present study, the opioid subsample was included in the overall sample analyses; however, we did not perform any additional exploratory analysis among that subsample separately due to the small sample. For the current analysis, we used all available data from participants who completed one or both sessions and provided complete surveys. All study procedures were approved by the Institutional Review Board of the University of Florida under protocol IRB201902033.

Measures

Substance Use Measures

Substance Use Before/During Participation. Participants were asked whether they had used any substance in the past 3 hr. If answered "yes," they were asked whether they had used any substance since they had started the survey, which substance(s) they had used, and how much of the substance/s they had used (the last item is not presented in the present analysis). Herein, Use Before/During refers to the group of participants who used a substance within 3 hr of participation, including those who used a substance during survey completion. No Use Before/During refers to the group of participants who did not use any substance before or during survey completion.

Typical Substance Use. Participants' most frequently used substance defined their substance-based subsample affiliation: either the alcohol subsample, cannabis subsample, cigarette subsample, or opioid subsample. For this substance, participants were asked to indicate (a) the number of days in the past month they used the substance and (b) the quantity used on a typical day of use (e.g., number of drinks/grams of cannabis/cigarettes/pills).

Demand for Alcohol/Cannabis/Cigarettes/Opioids

Participants' demand for their most frequently used substance (a measure of motivation to purchase and consume alcohol, cannabis, cigarettes, or opioids) was assessed using a state purchase task. The task for each substance was based on those previously used in the literature (Murphy & MacKillop, 2006 for alcohol; Aston et al., 2021 for cannabis; MacKillop et al., 2008 for cigarettes; and Strickland et al., 2019 for opioids). Participants in the alcohol subsample were asked how many of their favorite standard alcoholic drinks they would purchase and consume over the next 5 hr across 17 prices (from $0 to $30 per drink). Participants in the cannabis subsample were asked how many grams of their typical cannabis they would purchase at that present moment to consume over the next week across 19 prices (from $0–$60 per gram). Participants in the cigarette subsample were asked how many of their favorite cigarettes they would purchase at that present moment to smoke over the next 24 hr, across 16 prices (from $0 to $140 per cigarette). Participants in the opioid subsample were asked how many of their typical opioid pills they would purchase at that present moment to consume over the next 24 hr, across 17 prices (from $0 to $20 per pill). Similar assumptions across the four tasks were presented. That is, participants were asked to imagine that during the given period they could only get the substance from this source, the substance is similar in quality to their typical substance, they did not use any substance prior to making their decisions, and the substance they decided to purchase is for their own consumption during the given period as they cannot store, share, or sell any substance they choose to purchase. After reading the instructions, participants were asked three questions to confirm understanding of the main assumptions. Full task instructions are provided in the Supplemental Material.

Delay Discounting

To assess degree of delay discounting (i.e., the subjective decay in a reward's value due to a delay in receiving it), we used a hypothetical monetary adjusting amount task (Du et al., 2002). In this task, participants were asked to choose between two rewards, a larger delayed reward or a smaller immediate reward. The larger later reward was always $100, and the smaller immediate reward started at $50 and was adjusted based on the participant's choice. Participants were presented with six choices in each of seven delay blocks. The seven delays were 1 week, 2 weeks, 1 month, 6 months, 1 year, 5 years, and 25 years. Responses in each delay block yield an indifference point (i.e., the equal subjective value of the immediate and delayed rewards). After graphing the indifference points of the seven delays, the area under the connecting line, divided by the total rectangular area, was calculated and used as the measure of discounting (i.e., the area under the curve [AUC]; Myerson et al., 2001). In the present analysis, we used the ordinal AUC in which the delays are numbered by their order (Borges et al., 2016). The ordinal AUC allows equal weight of the delays in the AUC calculation. The AUC value ranges from 0 to 1, where lower values indicate greater discounting (i.e., greater preference for the immediate but smaller rewards, or more "impulsive" decision making). The AUC is considered a-theoretical and can quantify degree of discounting with data that are considered nonsystematic (according to conventional algorithms by M. W. Johnson & Bickel, 2008). Because experimental manipulations might lead to nonsystematic discounting patterns (Stein et al., 2016), and the discounting task was presented after engagement with stimuli intended to influence discounting, we used the AUC as the delay discounting outcome.

State Craving

Participants were asked to rate their current desire for their most frequently used substance (i.e., the substance that determined the substance-based subsample and specific tasks) on a scale of 1 (*not at all*) to 10 (*extremely*). The scale was presented twice: at the beginning of the survey and after the behavioral economic tasks. In the present analysis, only the first rating, which was the baseline assessment, was used.

Data Quality Measures

Performance on Five Data Quality Indicators. The study included five data quality indicators. The five indicators were two attention checks embedded in the delay discounting task (e.g., "Would you prefer $0 now or $100 in a month"), one instructional item that asked the participants to remember a word to be entered in a later part of the survey, and whether the participant provided systematic demand data (according to Stein et al., 2015) or systematic delay discounting data (according to M. W. Johnson & Bickel, 2008). More details on the criteria for nonsystematic demand and discounting data are provided in the data analysis section below.

Performance on the five data quality indicators was used in two ways. The first was in data preparation. Best practices of crowdsourcing research suggest treating attention as a continuum and exhibiting caution when excluding participants based on a single indicator (Almog et al., 2023; Nichols & Edlund, 2020). Thus, we included participants who passed all or failed only one quality indicator (a similar approach of allowing one failed indicator was used in Phung et al., 2019). Accordingly, we included 45 participants who failed one quality indicator and excluded nine participants who failed two or more. Of the nine excluded participants, eight participants were from the cannabis group (one used a substance before/during participation), and one participant was from the opioid group (who did not use any substance before/during participation). Of these nine, three failed three quality indicators, and six failed two quality indicators. Of the 45 participants who failed only one quality indicator, 10 failed a delay discounting attention check, two failed the instructional attention check, 22 provided nonsystematic demand data, and 11 provided nonsystematic delay discounting data.

Second, after excluding the nine participants, we used the performance on the five quality indicators to examine whether a higher percentage of participants who used a substance before/during participation provided overall poorer data quality (i.e., failed a quality indicator) compared to participants who did not use any substance before/during participation. Poor-quality data are often flagged for exclusion. If recent substance use impairs performance on data quality items, then such data would likely be excluded, and additional assessment of proximate substance use may not be necessary.

Duration of Survey Completion. We used a Qualtrics-derived timing variable of duration of survey completion (in seconds), as a measure that captures how long it took the participant to complete the survey. Commonly, extremely fast response times and survey completion times indicate lack of attention or careless responding, although extremely long response times can also be problematic (Wood et al., 2017).

Demographics

Participants were asked for their age, sex, race, ethnicity, annual income, and highest educational degree.

Data Analysis

We used R Statistical Software (Version 4.1.2, R Core Team, 2020) run within RStudio (RStudio Team, 2019) to describe the sample and SPSS (Version 29.0.0.0) to perform the regression modeling. We used GraphPad Prism (Version 10.0.2) to analyze behavioral measures and produce the demand curves and other graphs.

Substance Use Before/During Participation (Aim 1)

We described the numbers (and percentages) of observations completed by participants who used alcohol, cannabis, cigarettes, other substances, or more than one substance before/during participation, in the full sample (i.e., of total number of observations in the data set, $n = 527$), and in the alcohol, cannabis, cigarette, and opioid subsamples separately. Additionally, we described percentages of substance use before/during participation across sessions, whether participants used in one, both, or neither of the sessions. This was the only use of data from the second session; all other analyses employ data from the first session only, described in detail below.

Characterizing Use Before/During Versus No Use Before/During Groups (Aim 2)

Focusing on the first session only, we compared characteristics of those in the Use Before/During and No Use Before/During groups. We compared these groups on demographic characteristics, performance on data quality indicators (i.e., passing/failing quality indicators, duration to complete the survey), condition assignment in the parent study, and the typical substance use (i.e., frequency and quantity) in each substance-based subsample. We used Fisher's exact tests to compare categorical variables and parametric or nonparametric *t* tests (Mann–Whitney U test) for continuous variables based on normality of distributions. Distributions were determined as nonnormal if skewness or kurtosis had a value greater than |2|.

Substance Use Before/During Participation and Behavioral Measures and Craving (Aim 3)

To evaluate the potential association between using a substance before/during participation and the outcomes of demand for substances, delay discounting, and craving, we used the data from the first session only, using independent observations (i.e., we did not include within-subject data). We first presented the mean demand for alcohol, cannabis, cigarettes, and opioids, delay discounting, and craving for both groups: Use Before/During group and No Use Before/During group. Next, using multiple regression models, we evaluated the association between using a substance before/during participation and demand/delay discounting/craving, controlling for demographic characteristics (i.e., age, education, income), typical substance use (i.e., frequency and quantity), and variables related to the parent study (substance-based subsample and condition assignment). Specifically, we ran four multiple regressions with dependent variables of (a) demand intensity, (b) demand rate of change of elasticity, (c) delay discounting, and (d) craving. Regression models were conducted with all participants, standardizing variables across substance subsamples as necessary. Specific data preparation and analysis procedures are reported herein.

Demand for Substances. We first assessed whether any of the purchase task data were nonsystematic following conventional algorithms (Stein et al., 2015) and excluded nonsystematic data from the demand analysis. Consumption was identified as nonsystematic if it violated any one of the three criteria: (1) violating the "trend" criterion, that is, if relative change scores, calculated with $(\log Q_1 - \log Q_n)/(\log P_n - \log P_1)$ where Q_1 is consumption at the first price P_1 and Q_n is consumption at last price P_n, were lower than 0.025 indicating overall nondecreasing consumption across escalating prices; (2) violating the "bounce" criterion, that is, if there were more than two increases in consumption from one price to the next, with at least one increase greater than 25% of the initial consumption when the substance was free, or three or more increases in consumption even if lower than 25% of the initial consumption; or (3) violating the "reversal from zero" criterion, that is, if consumption was greater than zero following at least two consecutive prices with zero consumption. After inspecting the data, we also excluded two participants with extreme values of consumption that were unrealistic for the period of time presented in the task (i.e., 100 alcoholic drinks for 5 hr, 100 opioid pills for 24 hr).

A nonlinear regression equation was fit to the individual and mean consumption data of those who used a substance before/during participation (Use Before/During) and those who did not (No Use Before/During) within the alcohol, cannabis, cigarette, and opioid subsamples separately. We used the exponentiated demand model (Koffarnus et al., 2015):

$$Q = Q_0 \times 10^{k(e^{-\alpha Q_0 C}-1)}, \tag{1}$$

where Q was consumption at price C, Q_0 was the derived intensity (consumption as price reaches zero, left unconstrained), and α was the derived rate of change of elasticity. The span parameter k was calculated as the range of consumption across the data set in logarithmic units adding 0.5, that is, log(max − min) + 0.05. The k parameter was set to 2.2 for the alcohol model, 2.5 for the cannabis and cigarette models, and 1.9 for the opioid model. To use the nonlinear regression and based on the specific purchase task, the price of $0 was adjusted to 0.01 (for alcohol, cannabis, opioids) or 0.001 (for cigarettes). The demand curves were visually inspected.

For the subsequent regression analyses, we used the observed intensity metric (i.e., the hypothetical consumption if the substance was free) and the nonlinear regression-derived metric of rate of change of elasticity (i.e., alpha, reflects sensitivity to price increases). Higher intensity and lower alpha reflect greater demand. Both metrics were standardized within the alcohol, cannabis, cigarette, and opioid subsamples. To evaluate the association between using a substance before/during participation and demand for substances while controlling for personal characteristics and experiment-related factors, we conducted two multiple regression models. One model predicted intensity (i.e., consumption when the substance is free), and the other predicted rate of change of elasticity (i.e., alpha). The variables that were entered into the model were age, education, income, number of days of substance use in the past month, typical number of substance units used on a day of use (z-score), substance-based subsample, condition assignment in the parent study, and substance use before/during participation (yes/no). All regression assumptions were met.

Delay Discounting. To assess data quality, we first evaluated the rates of nonsystematic data following conventional algorithms (M. W. Johnson & Bickel, 2008). Per the present task, discounting data were described as nonsystematic if (a) any indifference point was at least $20 greater than the immediately preceding indifference point, or (b) the last indifference point was lower than the first indifference point by $10 or less. We characterized data based on these criteria but did not exclude nonsystematic delay discounting data (e.g., Stein et al., 2016). For presentation purposes, we graphed the mean delay discounting AUC across the two groups, Use Before/During and No Use Before/During. The mean delay discounting AUCs across the two groups (Use Before/During and No Use Before/During) were also evaluated for potential differences using a t test. The AUC distributions were relatively normal and therefore were not transformed prior to the statistical analyses. Next, a multiple regression evaluated the association between using a substance before/during participation and delay discounting while controlling for age, education, income, typical substance use (frequency and quantity), substance-based subsample, and condition assignment in the parent study. All regression assumptions were met.

Craving. For presentation purposes, we graphed the mean craving scores across the groups Use Before/During and No Use Before/During and evaluated potential differences in mean craving scores using a t test. A multiple regression model, with the craving score as the dependent variable evaluated the association with using a substance before/during participation, controlling for the same variables as previously described with one exception. Because craving was assessed prior to the manipulation (i.e., the slideshows), the condition assignment of the parent study was not entered into the model. All regression assumptions were met.

Exploratory Analysis Within Individual Substance-Based Subsamples. To ensure collapsing the groups into one model was appropriate, and to allow interested researchers to learn about specific substance-using populations, we ran the multiple regressions with intensity, delay discounting, and craving as outcomes within the alcohol, cannabis, and cigarette subsamples separately. In these models, we used the original nonstandardized variables of observed intensity and typical number of substance units used per day of use as predictors. Results for these exploratory analyses are presented in the Supplemental Material.

Results

Participants

After excluding nine participants who failed two or more quality indicators, the analytic sample (first session only) included 303 participants with mean age of 38.8 (SD = 11.8), of whom 56.8% were female. The sample was mostly White (85.1%), and almost half (47.7%) obtained a bachelor's degree or higher. See Table 1 for characteristics of the full sample and by the Use Before/During and No Use Before/During groups.

Substance Use Before/During Participation (Aim 1)

Overall, we analyzed 527 observations (i.e., completed sessions) that were collected across two sessions (first session n = 303; second session n = 224). Across all observations, 35.7% (n = 188) were completed with substance use occurring before participation (within the past 3 hr), and 9.9% (n = 52) were completed while the participant was using a substance during survey completion. In the first session only, of 303 participants, 35.6% (n = 108) completed the survey with substance use occurring before/during participation. Of the 303 participants who completed the initial session, 224 completed the second and provided additional information on recent/current substance use.

Analyzing the data of the 224 participants who completed the two study sessions, 25.4% (n = 57) used a substance before/during both sessions, 20.1% (n = 45) used a substance before/during one of the sessions only (22 used before/during the first session only; 23 used before/during the second session only), and 54.5% (n = 122) did not use any substance before/during either session.

Data from the first session (n = 303) showed that 104 participants chose alcohol as their most frequently used substance, 120 chose cannabis, 69 chose cigarettes, and 10 chose opioids. Participants

Table 1
First Session Sample Characteristics

Variable	Full sample $N = 303$	Use before/during $n = 108$	No use before/during $n = 195$	p
Age (years), *M* (*SD*)	38.8 (11.8)	36.6 (10.0)	40.0 (12.5)	**0.009**
Sex, *n* (%)				0.213
Female	172 (56.8%)	65 (60.2%)	107 (54.9%)	
Male	130 (42.9%)	42 (38.9%)	88 (45.1%)	
Prefer not to respond	1 (0.3%)	1 (0.9%)	0 (0.0%)	
Race, *n* (%)				0.869
White	258 (85.1%)	92 (85.2%)	166 (85.1%)	
Black or African American	17 (5.6%)	7 (6.5%)	10 (5.1%)	
Asian	14 (4.6%)	6 (5.6%)	8 (4.1%)	
Mixed	7 (2.2%)	1 (0.9%)	6 (3.1%)	
Other	3 (1.0%)	1 (0.9%)	2 (1.0%)	
NA	4 (1.3%)	1 (0.9%)	3 (1.5%)	
Ethnicity Hispanic Latinx, *n* (%)				0.843
No	273 (90.1%)	98 (90.7%)	175 (89.7%)	
Yes	30 (9.9%)	10 (9.3%)	20 (10.3%)	
Education, *n* (%)				**<.001**
High school or less	38 (12.5%)	24 (22.2%)	14 (7.2%)	
Some college but no degree	84 (27.7%)	37 (34.3%)	47 (24.1%)	
Associate degree (2-year)	37 (12.2%)	19 (17.6%)	18 (9.2%)	
Bachelor's degree (4-year)	102 (33.7%)	24 (22.2%)	78 (40.0%)	
Master's degree and higher	42 (14.0%)	4 (3.7%)	38 (19.6%)	
Income, *n* (%)				**.012**
\$25,000 or less	73 (24.5%)	34 (32.1%)	39 (20.3%)	
\$26,000–\$50,000	82 (27.5%)	32 (30.2%)	50 (26.0%)	
\$51,000–\$75,000	69 (23.2%)	24 (22.6%)	45 (23.4%)	
\$76,000–\$100,000	41 (13.8%)	12 (11.3%)	29 (15.1%)	
\$101,000–\$125,000	16 (5.4%)	4 (3.8%)	12 (6.3%)	
\$126,000–\$150,000	9 (3.0%)	0 (0.0%)	9 (4.7%)	
\$151,000 or more	8 (2.7%)	0 (0.0%)	8 (4.2%)	
Prefer not to say	5 (1.7%)	2 (1.9%)	3 (1.5%)	
Data quality, *n* (%)				.739
Passing all	258 (85.1%)	91 (84.3%)	167 (85.6%)	
Failing one	45 (14.9%)	17 (15.7%)	28 (14.4%)	
Survey duration (sec), *M* (*SD*)	1600.5 (408.6)	1710.2 (458.2)	1536.9 (364.8)	**<.001**
Condition assignment, *n* (%)				0.474
Nature	157 (51.8%)	59 (54.6%)	98 (50.3%)	
Built	146 (48.2%)	49 (45.4%)	97 (49.7%)	

Note. Data quality includes five indicators. Differences in categorical variables assessed with Fisher's exact test, continuous variables with *t* test or Mann–Whitney U test based on normality of distribution. To allow comparison, the three highest income categories were collapsed. Significant p values in bold represent significant differences between the Use Before/During and No Use Before/During groups. NA = not applicable.

who used a substance before/during participation typically used the substance that defined their substance-based subsample affiliation (i.e., most frequently used substance), but not exclusively. Across both sessions, 8.5% ($n = 45$) reported using a substance that was different than their most frequently used substance (i.e., cross-substance use). Lastly, across both sessions, 3.0% ($n = 16$) reported polysubstance use (i.e., more than one substance) in proximity to survey completion. Table 2 presents the counts and percentages of the substances used before/during participation across all observations in the full sample and per substance-based subsamples.

Characterizing Use Before/During Versus No Use Before/During Groups (Aim 2)

Table 1 presents characteristics of the full sample of the first session and the groups of Use Before/During ($n = 108$) and No Use Before/During ($n = 195$), including comparisons on demographics, data quality, and condition assignment between the groups.

Demographics

There were no significant differences between the groups of those engaging in substance use before/during participation versus those who did not, in sex, race, and ethnicity ($ps > .213$). However, participants in the Use Before/During group were younger ($p = .009$), obtained lower levels of education ($p < .001$), and had lower income ($p = .012$).

Typical Substance Use

Overall, in the full sample, those reporting using a substance before/during participation also reported using a substance more frequently in the past month ($M = 27.53$ days, $SD = 5.29$) compared to those reporting no substance use before/during participation ($M = 22.16$, $SD = 7.75$), $t(288.1) = -7.13$, $p < .001$. Comparing the z scores of the typical quantity of substance units per day of use, participants who used a substance before/during the survey reported

Table 2
Counts and Percentages of Substances Used Before/During Participation Across All Observations, in the Full Sample and per Substance-Based Subsamples

Substance used before/during participation	All observations $N = 527$	Alcohol subsample $n = 185$	Cannabis subsample $n = 204$	Cigarette subsample $n = 125$	Opioid subsample $n = 13$
Alcohol, *n* (%)	27 (5.1%)	20 (10.8%)	4 (2.0%)	3 (2.4%)	0 (0.0%)
Cannabis, *n* (%)	61 (11.6%)	7 (3.8%)	51 (25.0%)	1 (0.8%)	2 (15.4%)
Cigarettes, *n* (%)	111 (21.1%)	7 (3.8%)	18 (8.8%)	83 (66.4%)	3 (23.1%)
Other, *n* (%)	6 (1.1%)	0 (0.0%)	2 (1.0%)	1 (0.8%)	3 (23.1%)
More than one, *n* (%)	16 (3.0%)	5 (2.7%)	6 (2.9%)	4 (3.2%)	1 (7.7%)

Note. Observations N = number of observations across the two sessions. Subsample = based on participant choice of most frequently used substance. "Other" includes opioids, stimulants, and others. Percentages refer to the total number of observations per sample/subsample (column).

consuming higher quantities; however, the difference did not reach significance, $t(201.4) = -1.76$, $p = .079$. Table 3 summarizes the average typical consumption of Use Before/During and No Use Before/During groups within the three larger substance-based subsamples (i.e., alcohol, cannabis, and cigarette; the opioid subsample was not analyzed separately due to extremely small sample size). Those reporting proximate substance use in the alcohol and cannabis subsamples reported a higher number of days of use in the past month compared to those who reported no use before/during participation (ps < .006). Those reporting use before/during participation in the alcohol subsample also reported drinking more alcoholic drinks per day ($p = .004$). For cigarette smokers, the typical consumption did not differ between those reporting use or no substance use before/during participation (ps > .109).

Data Quality and Condition Assignment

The two groups, Use Before/During and No Use Before/During did not differ in percentage of failing any data quality indicator ($p = .739$). However, they did differ in survey duration, in which those who reported substance use before/during participation were significantly slower (median of 27.01 min, Q1–Q3: 23.67–33.17) compared to those who reported no use before/during participation (median of 24.65 min, Q1–Q3: 21.63–28.63), Mann–Whitney U = 13079.5, $p < .001$. On average, the duration of survey completion for those who reported substance use before/during participation was 173.3 s longer (i.e., 11.3% longer). Lastly, the groups did not differ in the condition assignment in the first session of the parent study ($p = .474$).

Behavioral Measures (Demand, Delay Discounting) and Craving (Aim 3)

Demand for Substances

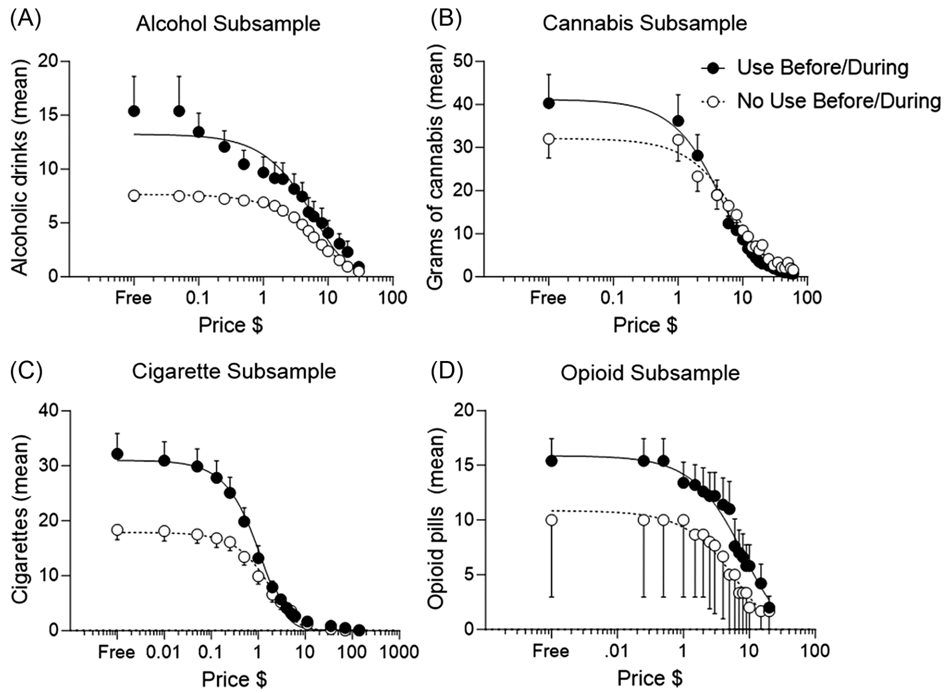
Alcohol. For the alcohol subsample, we identified and excluded from analyses two participants with nonsystematic purchase data, who failed the second criterion ("bounce"), and another participant with unrealistic consumption data (with consumption of 100 drinks). After excluding the three participants from the demand analysis, 88 participants reported not using any substance before/during participation. Of 13 participants who reported use before/during participation, six drank alcohol, two used cannabis, two smoked tobacco cigarettes, and three used more than one substance. The exponentiated model described the mean and individual consumption data of the two groups relatively well: Use Before/During ($R^2 = .93$, root-mean square error [RMSE] = 1.17, Q1–Q3 $R^2 = .88$–.97) and

Table 3
Typical Substance Use of Use Before/During and No Use Before/During Groups, in Substance-Based Subsamples (Alcohol, Cannabis, Cigarette)

Variable	Use before/during	No use before/during	*p*
Alcohol subsample, *n*	14	90	
Drinking days per month, *M* (*SD*)	25.6 (5.1)	19.7 (7.1)	**.005**
Drinks per day, *M* (*SD*)	7.9 (5.2)	4.8 (4.4)	**.004**
Cannabis subsample, *n*	43	77	
Using days per month, *M* (*SD*)	26.7 (6.6)	23.3 (10.3)	**.006**
Grams per day, *M* (*SD*)	1.0 (1.2)	1.2 (1.7)	.667
Cigarette subsample, *n*	45	24	
Smoking days per month, *M* (*SD*)	29.0 (3.7)	29.4 (2.3)	.642
Cigarettes per day, *M* (*SD*)	18.4 (11.6)	13.5 (7.7)	.109

Note. Differences between groups were assessed with Mann–Whitney U tests due to nonnormal distributions. Significant p values in bold represent significant differences between the Use Before/During and No Use Before/During groups.

Figure 1
Demand for Substances



Note. Hypothetical mean consumption as a function of price for participants who reported using, or not using a substance before/during participation. The *X* axis (logarithmic) represents the price of a single substance unit; the *Y* axis represents the number of substance units purchased. Circles and error bars represent the mean observed data and standard error of the mean, respectively. Curves represent the fit of the nonlinear regression equation to the data points (Koffarnus et al., 2015). To allow the nonlinear regression analysis and based on the specific purchase task, the price of \$0 was adjusted to 0.01 (alcohol, cannabis, opioids) and 0.001 (cigarettes). The *k* parameter in the nonlinear regression equation (see text for details) was set to 2.2 for alcohol, 2.5 for cannabis and cigarettes, and 1.9 for opioids. Some of the error bars are hidden by the data points.

No Use Before/During (R^2 > .99, RMSE = .14, Q1–Q3 R^2 = .87–.95). Figure 1 (Panel A) depicts the demand curves for alcohol for those in the Use Before/During (n = 13) and No Use Before/During (n = 88) groups. Visual inspection of the mean demand curves suggests higher demand for alcohol for those using a substance before/during participation.

Cannabis. For the cannabis subsample, 19 participants provided nonsystematic purchase data and were excluded from the demand analysis; 13 participants failed the first criterion ("trend"), and six participants failed the second criterion ("bounce"), from whom four had two increases in consumption (at least one higher than 25% of the initial consumption), and two had three or more increases (lower than 25% of the initial consumption). Of those providing systematic data, 64 participants reported no use before/during participation. Of the 37 participants reporting substance use before/during participation, 24 participants used cannabis, seven used nicotine, two used alcohol, and four used more than one substance. The exponentiated model described the mean and individual consumption data of the two groups well: Use Before/During (R^2 = .99, RMSE = .92, Q1–Q3 R^2 = .84–.96) and No Use Before/During (R^2 = .99, RMSE = .71, Q1–Q3 R^2 = .85–.96). Figure 1 (Panel B) depicts the demand curves of the Use Before/During (n = 37) and No Use Before/During (n = 64) groups. Visual inspection of the mean demand curves suggests higher demand for cannabis for those using a substance before/during participation.

Cigarettes. For the cigarette subsample, all purchase data were systematic and were included in the analysis. Of 69 participants, 45 reported substance use before/during participation. Of those, 43 participants smoked nicotine cigarettes, and two used more than one substance. The exponentiated model described the mean and individual consumption data of the two groups well: Use Before/During (R^2 = .99, RMSE = 1.03, Q1–Q3 R^2 = .96–.98) and No Use Before/During (R^2 = .99, RMSE = .64, Q1–Q3 R^2 = .94–.98). Figure 1 (Panel C) depicts the mean demand curves for cigarettes for the Use Before/During (n = 45) and No Use Before/During (n = 24) groups. Visual inspection of the mean demand curves suggests higher demand for cigarettes for those using a substance before/during participation.

Opioids. For the opioid subsample, we excluded one participant who provided nonsystematic purchase data, failing the first criterion ("trend"), and one who provided unrealistic consumption data (i.e., 100 pills). Of the eight included in the final analysis, five reported before/during use; two used cannabis, one used opioids, one smoked cigarettes, and one used more than one substance. Three participants did not use any substance before/during

Table 4
Regression Models Results

Predictor	Demand intensity $N = 274$ $F(10, 263) = 8.17$, $R^2 = .24$, $p < .001$					Delay discounting $N = 298$ $F(10, 287) = 4.63$, $R^2 = .14$, $p < .001$					Craving $N = 298$ $F(9, 288) = 9.57$, $R^2 = .23$, $p < .001$				
	b	*SE b*	β	*p*	sr²	*b*	*SE b*	β	*p*	sr²	*b*	*SE b*	β	*p*	sr²
Constant	.38	.46		.410		.41	.09		<.001		6.85	1.22		<.001	
Age	−.01	.01	−.10	.079	.01	.00	.00	**.23*** **	<.001	.05	−.04	.01	**−.17** **	.002	.03
Education	.22	.13	.11	.076	.01	.06	.03	**.15***	.020	.02	−.30	.35	−.05	.389	.00
Income	−.02	.04	−.03	.591	.00	−.01	.01	−.07	.276	.00	.14	.11	.07	.206	.00
Days/month	−.00	.01	−.02	.729	.00	.00	.00	.01	.881	.00	.03	.02	.07	.253	.00
Substance/day	.46	.06	**.40*** **	<.001	.15	−.03	.01	**−.13***	.022	.02	.71	.18	**.21*** **	<.001	.04
Substance (alcohol)	−.01	.34	−.01	.970	.00	.06	.07	.14	.387	.00	−1.94	.89	**−.32***	.030	.01
Substance (cannabis)	−.13	.33	−.07	.688	.00	.05	.07	.12	.438	.00	−.81	.87	−.14	.348	.00
Substance (cigarette)	−.14	.33	−.06	.684	.00	.06	.07	.12	.386	.00	−.99	.88	−.14	.257	.00
Condition	−.13	.11	−.06	.239	.00	−.05	.02	**−.12***	.033	.01					
Before/during use	.47	.13	**.23*** **	<.001	.04	−.04	.03	−.09	.138	.01	1.22	.36	**.20*** **	<.001	.03

Note. Full sample results for regression models evaluating the association between before/during substance use and demand intensity (observed values of hypothetical consumption when substance is free), delay discounting (ordinal AUC), and craving, controlling for individual characteristics and experiment-related variables. Significant β values in bold. Age = age in years; Education = dichotomously coded as obtained less than 4-year degree (0) or Bachelor's degree or higher (1); Income = ordinal categories (1–7, representing annual income from less than \$25 K and up to \$151 K and higher, respectively); Days/month = number of days in past month in which the main substance was consumed; Substance/day = *z* score of number of substance units consumed during a typical day of use (six outliers with a *z* score higher than 3 were winsorized to 2.61; higher than the maximal nonoutlier *z* score of 2.60; three from the alcohol subsample, two from the cannabis subsample, and one from the cigarette subsample); Substance = substance-based subsample affiliation with dummy variables referenced to opioid subsample; Condition = dichotomous variable for condition assignment in the first session in the parent study, Built (0) or Nature (1); Before/during use = not using a substance before/during participation (0) or using a substance before/during participation (1). Constant = Outcome for mean age, less than bachelor's degree, mean income category 2.56 (between 50 and 51 K), mean days per month of main substance consumption, mean quantity of substance used on a typical day of use, member of the opioid subsample, experimental built condition (not included in the craving model), and not using a substance before/during survey completion. To prepare the intensity variable, we first identified outliers within the alcohol, cannabis, cigarette, and opioid subsamples separately. Five outliers with *z* scores higher than 3 were winsorized to one raw unit higher than the maximal nonoutlier score in the group's observed consumption data: one alcohol score of 50 drinks was adjusted to 26, and four cigarette scores of 99 were adjusted to 61 cigarettes. The adjusted *z* scores were used as the dependent variable in the intensity model. Data from five participants were excluded list-wise from the models due to missing income. From the demand intensity model, 24 participants with nonsystematic/unrealistic consumption data were excluded. *SE* = standard error; AUC = area under the curve.
* $p < .05$. ** $p < .01$. *** $p < .001$.

participation. The exponentiated model described the mean consumption data of the two groups relatively well: Use Before/During ($R^2 = .97$, RMSE = .70, Q1–Q3 $R^2 = .55$–.94) and No Use Before/During ($R^2 = .97$, RMSE = .56, Q1–Q3 $R^2 = .76$–.96). Figure 1 (Panel D) depicts the mean demand curves for opioids for the Use Before/During ($n = 5$) and No Use Before/During ($n = 3$) groups. Visual inspection of the mean demand curves suggests higher demand for opioids for those using a substance before/during participation; however, we urge caution in interpreting the results due to the small sample size.

Regression Analyses. The multiple regression model evaluating the associations between before/during substance use and demand intensity (i.e., consumption when the substance is free), controlling for age, education, income, typical substance use, substance-based subsample, and condition assignment in the parent study, was significant, $F(10, 263) = 8.17$, $R^2 = .24$, $p < .001$. Results showed that using a substance before/during participation was associated with increased demand intensity ($b = .47$, $\beta = .23$, $p < .001$). See Table 4 for the regression results and details on variable coding and variable standardization. The regression model assessing rate of change of elasticity (i.e., natural logged alpha) for the overall sample was significant, $F(10, 263) = 2.15$, $R^2 = .08$, $p = .022$. However, no association was found between using a substance before/during participation and rate of change of elasticity (alpha) ($b = -.02$, $\beta = -.08$, $p = .235$). See Supplemental Table S1 for regression results.

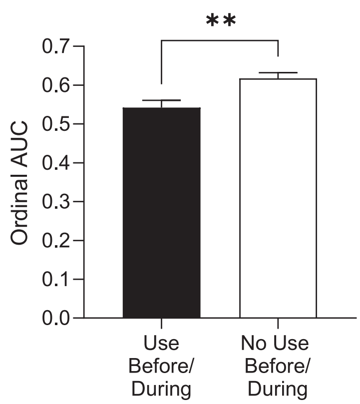
Delay Discounting

Figure 2 presents the delay discounting ordinal AUC results of the two groups, Use Before/During and No Use Before/During in the first session ($n = 303$). Results include nonsystematic discounting data from 11 participants who failed the first criterion of exhibiting an increase of at least \$20 from one indifference point to the next. Participants who reported use before/during participation had a lower mean delay discounting AUC value, representing steeper discounting (i.e., greater preference for immediate, smaller rewards) than those who reported no use before/during participation (0.54 and 0.62, respectively, $p = .001$). However, the regression model, $F(10, 287) = 4.63$, $R^2 = .14$, $p < .001$, showed that after controlling for age, education, income, typical substance use, substance-based subsample, and condition assignment, the association between using a substance before/during participation and discounting was not significant ($b = -.04$, $\beta = -.09$, $p = .138$). See Table 4 for regression results. The same regression was applied while excluding participants who failed any attention checks ($n = 12$) or provided nonsystematic data ($n = 11$), and results remained similar, confirming the results were not biased due to inattentive responding.

Craving for Substances

Overall, participants who reported using a substance before/during participation had higher mean craving for their most frequently used

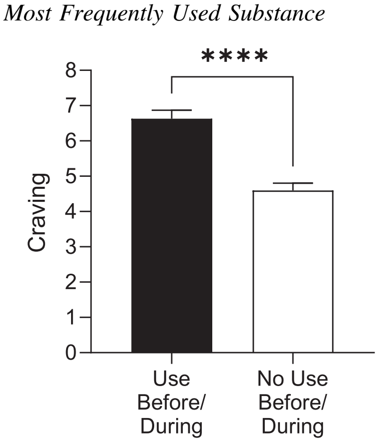
Figure 2
Delay Discounting



Note. Mean ordinal AUC values for participants reporting using versus not using a substance before/during participation. Lower AUC value represents greater delay discounting. Although the difference was significant, when controlling for individual characteristics and experiment-related variables in regression analysis, the association between before/during substance use and greater delay discounting did not reach standard levels of significance. Error bars represent the standard error of the mean. AUC = area under the curve.
** $p < .01$.

substance than those who reported no use (6.63 and 4.60, respectively, $p < .0001$; see Figure 3). The regression model evaluating the association between using a substance before/during participation and craving while controlling for the relevant covariates was significant, $F(9, 288) = 9.57$, $R^2 = .23$, $p < .001$. Using a substance before/during participation was significantly associated with higher craving for the most frequently used substance ($b = 1.22$, $\beta = .20$, $p < .001$). See Table 4 for regression results.

Figure 3
Craving for Most Frequently Used Substance



Note. Mean craving for participants reporting using versus not using a substance before/during participation. The association between before/during substance use and craving remained significant, even when controlling for individual characteristics in regression analysis. Error bars represent the standard error of the mean.
**** $p < .0001$.

Discussion

The present secondary analysis of 527 observations with 303 individuals who regularly use substances, collected on MTurk across two sessions, found that 35.7% of all observations were completed by participants who used a substance just before participation and 9.9% were completed while the participants were using a substance. Participants who used a substance before/during participation mostly used the substance they reported as most frequently used, but not exclusively. Among the alcohol and cannabis subsamples, those who used a substance before/during participation (i.e., very recent use) reported heavier typical use than those who did not, but not among cigarette smokers. Nevertheless, we found that, overall, using a substance before/during participation was associated with increased demand and craving for the most frequently used substance after controlling for group differences in age, education, income, and typical substance use severity. Moreover, most traditional measures of data quality were not sensitive to recent substance use; the one exception was longer survey completion time, which is not on its own an unambiguous indicator. Taken together, our data suggested that very recent substance use might increase demand and craving. These patterns persisted after accounting for typical frequency and quantity of use, suggesting a state (acute) effect of recent substance use on demand and craving beyond any trait-related (or typical, more stable) characteristics. We recognize that follow-up remote experimental studies with prospectively designed control conditions are needed to mechanistically confirm these results. However, the present results provide evidence that substance use before/during participation may represent a unique variable that could influence study results if undetected. Hence, naturalistic self-administration in proximity to survey completion is an important factor to consider and include in crowdsourcing research.

The present findings complement past research, extending from laboratory placebo-controlled administration studies to online samples of participants engaging in naturalistic drug self-administration. Our results showed that substance use before/during participation was associated with increased demand and craving for substances. Specifically, our results of higher demand and craving for alcohol within the alcohol subsample are consistent with past alcohol administration studies in laboratory conditions, which found increased demand and craving for alcohol after acute administration (Amlung et al., 2015; Motschman et al., 2022). In contrast to our online study, a laboratory cannabis–tobacco administration study (Hindocha et al., 2017) showed that immediately after active cannabis administration, craving for cannabis was reduced compared to the placebo group. However, over the following hour, craving for cannabis increased for all groups. In the present online cannabis subsample, the exact interval between the last cannabis use and survey completion is unknown; thus, timing of measurements may be a potential explanation and/or that craving itself changes over time. Hindocha et al. (2017) also found decreased demand using the elasticity and breakpoint indices following cannabis administration, although their findings also showed nonsignificant, increased intensity following cannabis use (i.e., consumption when cannabis is free), which aligns with our findings.

In our sample, participants who used a substance before/during participation were younger, had lower education and lower income, and they discounted hypothetical money more steeply (i.e., had greater preference for immediate but smaller monetary amounts) than those reporting no substance use before participation. However, the association between using a substance before/during participation and steeper discounting was not significant when we controlled for these individual differences, which aligns with the results of Bidwell et al. (2013) and Reynolds et al. (2006). Thus, the differences found in discounting among these participants did not appear to be a state-like effect related to the naturalistic substance use in proximity to participation but rather related to age, education, and typical substance use (which were significant predictors in our model). These findings align with past research on the associations between steeper discounting and severity of substance use (Amlung et al., 2017), younger age, lower education, and lower income (Reimers et al., 2009). More research on the acute effect of substances on behavioral measures is needed in laboratory and naturalistic conditions, and specifically assessing polysubstance and cross-substance effects (i.e., using a substance before participation that is different from the substance under study or subsample affiliation).

Polysubstance and cross-substance use in online samples appear to be prevalent and call for attention. Our data showed that 3% of the surveys were completed by participants who used more than one substance before/during study participation. Further, some participants (8.5%) reported using a substance prior to survey completion that did not match their self-reported, most regularly used substance (which determined their subsample affiliation and tasks presented). For example, from our data, some alcohol-subsample participants reported using cannabis or cigarettes prior to survey completion. Although not explicitly examined as an independent contributor in our analysis, past laboratory research suggested that administration of one substance may increase consumption of another substance. For example, in Barrett et al. (2006), smoking nicotine cigarettes (compared to denicotinized cigarettes) elevated self-administration of alcohol. The cross-substance effect in which consumption of one substance increases demand/craving for another substance may also explain the increased demand and craving as a state-related increase beyond typical trait characteristics. Therefore, it is necessary to control for proximate substance use of any substance, not only the primary substance under study.

Our results suggest that substance use before or during participation among crowdsourcing samples of individuals who regularly use substances is not likely to be detected (or screened out) by traditional quality measures used in crowdsourcing research. In our sample, participants who used substances before/during participation did not perform differently from those who did not use any substance on attention checks, nor in providing nonsystematic delay discounting or demand data. The only measure related to data quality in which differences were detected was survey completion time. Commonly, speed of response is used to flag participants for exclusion because high speed may indicate lack of attention (Wood et al., 2017). However, in our study, participants who used substances before/during survey completion were slower to complete the survey. This may result from the pharmacological effects of the drugs lengthening response time and impairing cognitive functions including decision making (e.g., alcohol: Tzambazis & Stough, 2000; cannabis: Crean et al., 2011). It may also result from the physical action of using a substance (e.g., drinking) during the survey and task completion. Although durations to complete the survey among participants who used a substance before/during survey completion were, on average, 11.3% longer, this difference would not flag them as outliers. Regardless, this is especially relevant for research assessing timing variables (e.g., response time) in decision making or other tasks, warranting controlling for very recent and acute substance use. Taken together, participants who reported using a substance prior to participation provided valid and good quality data; however, they performed differently on the behavioral measures that are frequently used in substance use and addiction research. Thus, our findings indicate that naturalistic substance use before/during participation is prevalent and might affect overall results if not detected and controlled for.

Given that using substances before/during participation was associated with higher demand and craving for substances, substance use before/during participation might pose further problems in different experimental studies, potentially leading to errors in study conclusions. First, the pharmacological effects of different substances (e.g., alcohol and cannabis) may reduce or eliminate manipulation effects, or the attention required to properly and actively engage with instructions or scenarios (e.g., reading a vignette with a specific scenario, episodic future thinking). Second, considering repeated measures designs, our results suggest that people are not consistently consuming substances before or during study participation, which would render differential pharmacological or attentional influences across one session but not the other, resulting in standardization issues with further impact on study results. Current substance use or deprivation not accounted for could also contaminate group-based inferences or associations between variables in cross-sectional studies. These concerns will also vary across different substances, for example, our findings suggest that the influence of recent smoking among cigarette smokers on craving appeared to be less evident, as might be reasonably anticipated (see Supplemental Material for more details). Lastly, the influence of very recent substance use might interact differently with different experimental manipulations or even different conditions within the same study. As our study included participants who used different substances and completed different purchase tasks (i.e., limited sample size for each drug condition), we did not examine a possible interaction. Still, researchers should be aware of the possibility that recent substance use may affect experimental manipulations/conditions differently. Targeted experimental research should be designed to test whether the pharmacological effects and different manipulations interact to differentially affect decision-making processes.

As online crowdsourcing research is likely to remain popular and even expand, especially after the COVID-19 era, best practices are warranted, specifically for regular substance use populations. Details about recent substance use among participants are needed: the last time any substance was used, the quantity/dose used, and the method of administration. The latter is especially relevant to cannabis as differences exist in the onset and duration of effect across methods of administration (e.g., smoking, edibles). Other details such as food intake in proximity to substance use and self-reported current level of "intoxication" may be informative and useful in order to control for in subsequent analysis. In contrast, researchers could ask participants to not use any drug in a specific time period before participation or stay consistent with their self-administration in proximity to study participation in repeated measures designs. Moreover, researchers may specifically target this population, or even develop systematized at-home self-administration studies (e.g., St. Pierre et al., 2023) once meeting all ethical approvals (e.g., Institutional Review Board). Researchers should also plan for larger samples that will enable exclusion of participants based on substance use in proximity to study participation or analysis of separate subsamples, especially if targeting heavy substance-using samples.

The present study has three main limitations. First, this is a secondary analysis of a repeated measures experimental study. It is possible that our manipulation was confounded with the naturalistic substance use before or during participation. To address this limitation, we verified that there were similar percentages of participants with and without substance use before/during the survey in both conditions and controlled for condition assignment in the regression models. Second, in our study, substance use Before/During was defined as using a substance within the past 3 hr including during survey completion. We cannot determine the level of intoxication (if any existed) following participants' self-reported naturalistic self-administration, which depends on many other factors that we did not assess (e.g., timing, dose and quantity administered, method of administration, body mass index, consumption of food, etc.). Third, our sample was heterogeneous and included participants who reported using different substances. Analyzing each substance-based subsample separately reduced the sample size dramatically, calling for more replication research with different substance using populations and larger samples. Specific research could also include different methods of administration of different substances. For example, we recruited in the parent study cigarette smokers; however, future research could also investigate nicotine vaping samples. For limited power considerations, we also did not pursue any analysis within-subjects among participants that used a substance before/during one of the sessions but not the other. More specific and targeted research is needed, with larger samples, on the rates and effects of recent substance use in general and on behavioral measures specifically, in online samples.

In summary, our analysis suggests that naturalistic substance use before or during study participation is prevalent in online samples of individuals who regularly use substances. Substance use researchers recruiting on crowdsourcing platforms should expect a subgroup of participants who use substances in proximity to or even during study participation. These participants overall may provide good quality data but may respond differently on behavioral measures, which are frequently used in substance use and addiction research. Collecting data on substance use just before or during participation will improve the rigor of online substance use research. This line of research could also advance the field to develop studies that target this population, for example, by developing at-home systematized self-administration studies.

References

Almog, S., Vásquez Ferreiro, A., Berry, M. S., & Rung, J. M. (2023). Are the attention checks embedded in delay discounting tasks a valid marker for data quality? *Experimental and Clinical Psychopharmacology*, *31*(5), 908–919. https://doi.org/10.1037/pha0000645

Amlung, M., McCarty, K. N., Morris, D. H., Tsai, C. L., & McCarthy, D. M. (2015). Increased behavioral economic demand and craving for alcohol following a laboratory alcohol challenge. *Addiction*, *110*(9), 1421–1428. https://doi.org/10.1111/add.12897

Amlung, M., Vedelago, L., Acker, J., Balodis, I., & MacKillop, J. (2017). Steep delay discounting and addictive behavior: A meta-analysis of continuous associations. *Addiction*, *112*(1), 51–62. https://doi.org/10.1111/add.13535

Arditte, K. A., Çek, D., Shaw, A. M., & Timpano, K. R. (2016). The importance of assessing clinical phenomena in Mechanical Turk research. *Psychological Assessment*, *28*(6), 684–691. https://doi.org/10.1037/pas0000217

Aston, E. R., Metrik, J., Rosen, R. K., Swift, R., & MacKillop, J. (2021). Refining the marijuana purchase task: Using qualitative methods to inform measure development. *Experimental and Clinical Psychopharmacology*, *29*(1), 23–35. https://doi.org/10.1037/pha0000355

Barrett, S. P., Tichauer, M., Leyton, M., & Pihl, R. O. (2006). Nicotine increases alcohol self-administration in non-dependent male smokers. *Drug and Alcohol Dependence*, *81*(2), 197–204. https://doi.org/10.1016/j.drugalcdep.2005.06.009

Bidwell, L. C., MacKillop, J., Murphy, J. G., Grenga, A., Swift, R. M., & McGeary, J. E. (2013). Biphasic effects of alcohol on delay and probability discounting. *Experimental and Clinical Psychopharmacology*, *21*(3), 214–221. https://doi.org/10.1037/a0032284

Borges, A. M., Kuang, J., Milhorn, H., & Yi, R. (2016). An alternative approach to calculating Area-Under-the-Curve (AUC) in delay discounting research. *Journal of the Experimental Analysis of Behavior*, *106*(2), 145–155. https://doi.org/10.1002/jeab.219

Crean, R. D., Crane, N. A., & Mason, B. J. (2011). An evidence based review of acute and long-term effects of cannabis use on executive cognitive functions. *Journal of Addiction Medicine*, *5*(1), 1–8. https://doi.org/10.1097/ADM.0b013e31820c23fa

Du, W., Green, L., & Myerson, J. (2002). Cross-cultural comparisons of discounting delayed and probabilistic rewards. *The Psychological Record*, *52*(4), 479–492. https://doi.org/10.1007/BF03395199

Field, M., Santarcangelo, M., Sumnall, H., Goudie, A., & Cole, J. (2006). Delay discounting and the behavioural economics of cigarette purchases in smokers: The effects of nicotine deprivation. *Psychopharmacology*, *186*(2), 255–263. https://doi.org/10.1007/s00213-006-0385-4

Hindocha, C., Lawn, W., Freeman, T. P., & Curran, H. V. (2017). Individual and combined effects of cannabis and tobacco on drug reward processing in non-dependent users. *Psychopharmacology*, *234*(21), 3153–3163. https://doi.org/10.1007/s00213-017-4698-2

Johnson, M. W., & Bickel, W. K. (2008). An algorithm for identifying nonsystematic delay-discounting data. *Experimental and Clinical Psychopharmacology*, *16*(3), 264–274. https://doi.org/10.1037/1064-1297.16.3.264

Johnson, P. S., Sweeney, M. M., Herrmann, E. S., & Johnson, M. W. (2016). Alcohol increases delay and probability discounting of condom-protected sex: A novel vector for alcohol-related HIV transmission. *Alcoholism: Clinical and Experimental Research*, *40*(6), 1339–1350. https://doi.org/10.1111/acer.13079

Koffarnus, M. N., Franck, C. T., Stein, J. S., & Bickel, W. K. (2015). A modified exponential behavioral economic demand model to better describe consumption data. *Experimental and Clinical Psychopharmacology*, *23*(6), 504–512. https://doi.org/10.1037/pha0000045

MacKillop, J., Murphy, J. G., Ray, L. A., Eisenberg, D. T., Lisman, S. A., Lum, J. K., & Wilson, D. S. (2008). Further validation of a cigarette purchase task for assessing the relative reinforcing efficacy of nicotine in college smokers. *Experimental and Clinical Psychopharmacology*, *16*(1), 57–65. https://doi.org/10.1037/1064-1297.16.1.57

Metrik, J., Aston, E. R., Kahler, C. W., Rohsenow, D. J., McGeary, J. E., & Knopik, V. S. (2015). Marijuana's acute effects on cognitive bias for affective and marijuana cues. *Experimental and Clinical Psychopharmacology*, *23*(5), 339–350. https://doi.org/10.1037/pha0000030

Motschman, C. A., Griffin, S. A., Warner, O. M., & McCarthy, D. M. (2022). Craving, stimulation, and sedation as predictors of alcohol demand under intoxication. *Psychology of Addictive Behaviors*, *36*(1), 20–27. https://doi.org/10.1037/adb0000719

Murphy, J. G., & MacKillop, J. (2006). Relative reinforcing efficacy of alcohol among college student drinkers. *Experimental and Clinical Psychopharmacology*, *14*(2), 219–227. https://doi.org/10.1037/1064-1297.14.2.219

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental*

Analysis of Behavior, *76*(2), 235–243. https://doi.org/10.1901/jeab.2001.76-235

Nichols, A. L., & Edlund, J. E. (2020). Why don't we care more about carelessness? Understanding the causes and consequences of careless participants. *International Journal of Social Research Methodology*, *23*(6), 625–638. https://doi.org/10.1080/13645579.2020.1719618

Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, *23*(3), 184–188. https://doi.org/10.1177/0963721414531598

Phung, Q. H., Snider, S. E., Tegge, A. N., & Bickel, W. K. (2019). Willing to work but not to wait: Individuals with greater alcohol use disorder show increased delay discounting across commodities and less effort discounting for alcohol. *Alcoholism: Clinical and Experimental Research*, *43*(5), 927–936. https://doi.org/10.1111/acer.13996

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reed, S. C., Levin, F. R., & Evans, S. M. (2012). Alcohol increases impulsivity and abuse liability in heavy drinking women. *Experimental and Clinical Psychopharmacology*, *20*(6), 454–465. https://doi.org/10.1037/a0029087

Reimers, S., Maylor, E. A., Stewart, N., & Chater, N. (2009). Associations between a one-shot delay discounting measure and age, income, education and real-world impulsive behavior. *Personality and Individual Differences*, *47*(8), 973–978. https://doi.org/10.1016/j.paid.2009.07.026

Reynolds, B., Richards, J. B., & de Wit, H. (2006). Acute-alcohol effects on the Experiential Discounting Task (EDT) and a question-based measure of delay discounting. *Pharmacology, Biochemistry, and Behavior*, *83*(2), 194–202. https://doi.org/10.1016/j.pbb.2006.01.007

RStudio Team. (2019). *RStudio: Integrated development for R*. https://www.rstudio.com/

Stanton, K., Carpenter, R. W., Nance, M., Sturgeon, T., & Villalongo Andino, M. (2022). A multisample demonstration of using the prolific platform for repeated assessment and psychometric substance use research. *Experimental and Clinical Psychopharmacology*, *30*(4), 432–443. https://doi.org/10.1037/pha0000545

St. Pierre, M., Daniels, S., Sanchez, T. A., Holtzman, S., Russo, E. B., & Walsh, Z. (2023). The Naturalistic Cannabis Administration Protocol (NCAP): A proof-of-concept study. *Journal of Psychoactive Drugs*, *55*(4), 389–401. https://doi.org/10.1080/02791072.2022.2125466

Stein, J. S., Koffarnus, M. N., Snider, S. E., Quisenberry, A. J., & Bickel, W. K. (2015). Identification and management of nonsystematic purchase task data: Toward best practice. *Experimental and Clinical Psychopharmacology*, *23*(5), 377–386. https://doi.org/10.1037/pha0000020

Stein, J. S., Wilson, A. G., Koffarnus, M. N., Daniel, T. O., Epstein, L. H., & Bickel, W. K. (2016). Unstuck in time: Episodic future thinking reduces delay discounting and cigarette smoking. *Psychopharmacology*, *233*(21–22), 3771–3778. https://doi.org/10.1007/s00213-016-4410-y

Strickland, J. C., Amlung, M., & Reed, D. D. (2022). Crowdsourcing methods in addiction science: Emerging research and best practices. *Experimental and Clinical Psychopharmacology*, *30*(4), 379–380. https://doi.org/10.1037/pha0000582

Strickland, J. C., Lile, J. A., & Stoops, W. W. (2019). Evaluating nonmedical prescription opioid demand using commodity purchase tasks: Test–retest reliability and incremental validity. *Psychopharmacology*, *236*(9), 2641–2652. https://doi.org/10.1007/s00213-019-05234-y

Strickland, J. C., & Stoops, W. W. (2019). The use of crowdsourcing in addiction science research: Amazon Mechanical Turk. *Experimental and Clinical Psychopharmacology*, *27*(1), 1–18. https://doi.org/10.1037/pha0000235

Tzambazis, K., & Stough, C. (2000). Alcohol impairs speed of information processing and simple and choice reaction time and differentially impairs higher-order cognitive abilities. *Alcohol and Alcoholism*, *35*(2), 197–201. https://doi.org/10.1093/alcalc/35.2.197

Wood, D., Harms, P. D., Lowman, G. H., & DeSimone, J. A. (2017). Response speed and response consistency as mutually validating indicators of data quality in online samples. *Social Psychological and Personality Science*, *8*(4), 454–464. https://doi.org/10.1177/1948550617703168

Received April 1, 2024
Revision received June 25, 2024
Accepted June 27, 2024 ■