Beyond Frequency and $\overline{}$ **Students: Context of Using Cannabis Relates to Cannabis-** \overline{a} **Related Outcomes Quantity of Cannabis Consumption among College**

Cannabis 2024 © Author(s) 2024 researchmj.org 10.26828/cannabis/2024/000225

Matison W. McCool1 , Matthew R. Pearson1, the Marijuana Outcomes Study Team*, the Protective Strategies Study Team, & the Addictions Research Team*****

¹Center on Alcohol, Substance Use, and Addictions, University of New Mexico

ABSTRACT

Objective: Due to little knowledge regarding the contextual factors related to cannabis use, we aimed to provide descriptive statistics regarding contextual factors related to use and examine the predictive ability of contextual factors. **Method:** We included college student participants ($n = 5700$; male = 2893, female = 3702, other gender identity = 48, missing = 57) from three multi-site studies in our analyses. We examined the means and standard deviations of contextual factors related to cannabis use (social context/setting, form of cannabis, route of administration, source of purchase, and proxies of use). Additionally, we tested the predictive ability of the contextual factors on cannabis use consequences, protective behavioral strategies, and severity of cannabis use disorder, via an exploratory machine learning model (random forest). Results: Descriptive statistics and the correlations between the contextual factors and the three outcomes are provided. Exploratory random forests indicated that contextual factors may be helpful in predicting consequences and protective behavioral strategies and especially useful in predicting the severity of cannabis use disorder. **Conclusions:** Contextual factors of cannabis use warrants further exploration, especially considering the difficulty in assessing dosage when individuals are likely to consume in a group context. We propose considering measuring contextual factors along with use in the past 30 days and consequences of use.

Key words: = cannabis use; cannabis-related consequences; social context; route of administration; college students; cannabis protective behavioral strategies

In the context of a massively growing legal cannabis market throughout the United States, a harm reduction approach to understanding cannabis necessitates the consideration of any relevant characteristic of one's cannabis use that may contribute to cannabis-related harms. In a meta-analysis, Pearson (2019) found a mediumsized association between cannabis use indicators and consequences $(r_w=0.367)$, demonstrating that most of the variance in cannabis-related negative consequences are not explained by any single indicator of cannabis use. This finding suggests that additional characteristics of cannabis use are needed to account for the likelihood of experiencing cannabis-related harms beyond frequency and quantity of use. Social contexts of use, or the temporal, motivational, and situational factors surrounding use, are additional characteristics of cannabis use that predict cannabis use outcomes (Beck et al., 2009).

Corresponding Author: Matison McCool, Ph.D., University of New Mexico, 2650 Yale Blvd SE MSC 11-6280. Albuquerque, NM 87106. Phone: (505) 925-2322. E-mail: mwmccool@unm.edu.

Contextual Factors of Cannabis Use

The operationalization of contextual factors of cannabis use in prior work has included categories related to social facilitation, peer acceptance, sexseeking, emotional pain (Beck et al., 2009), location of use, using companions (Spinella et al., 2019), place of purchase, route of administration (Parnes et al., 2018), environmental, emotional, and interpersonal contexts related to use (Gray et al., 2024). Recalling that cannabis use indicators such as quantity and frequency are moderately associated with consequences (Pearson, 2019), contextual factors are associated with cannabis use disorder (Beck et al., 2009), consequences, and protective behavioral strategy use (Gray et al., 2024) even when controlling for direct cannabis use indicators. However, much of the previous literature has limited the examination of contextual factors to a few behaviors (e.g., solitary use; Spinella et al., 2019), creating constructs out of specific behaviors (Beck et al., 2009), or using latent profile analyses to find patterns of contextual factors (Gray et al., 2024). Contextual factors of cannabis use are significantly correlated (Beck et al., 2009) which may lead to issues of multicollinearity when attempting to examine many individual contexts in a predictive model. Given the broad operational definition of cannabis use contextual factors and many distinct situations in which cannabis use can occur, we focused our study on the following contexts of cannabis use: social situations and settings of use, form of cannabis and route of administration, and the source of purchase.

Situational and Setting Contexts

Most of the research on social context of cannabis use has focused on solitary use vs. social use. Compared to social users, solitary cannabis users have reported higher levels of drinking to cope, higher levels of cannabis use, and greater endorsement of cannabis abuse/dependence (Spinella et al., 2019). Solitary cannabis use by adolescents has been shown to relate to cannabis use disorder symptoms during adolescence, but also prospectively predicts cannabis use disorder symptoms in young adulthood (Creswell et al., 2015). Solitary cannabis use frequency has been shown to mediate (i.e., account for) the effects of social anxiety on cannabis use and negative cannabis-related consequences (Buckner et al., 2016). Thus, solitary use of cannabis has been identified as a risk factor for negative cannabisrelated consequences. Beck et al., (2019) included settings of cannabis use (i.e., in a car, in a dorm room) as part of a social facilitation construct. Results assessing the relationship between social facilitation and DSM-IV cannabis use disorder criteria found that increased social facilitation was significantly associated with cannabis use disorder symptom severity. Overall, where and with whom individuals use cannabis are associated with cannabis use outcomes above and beyond direct use indicators.

Context of Cannabis Form and Route of Administration

With the rapid proliferation of legal cannabis markets, cannabis preparations have diversified to include a wide range of edible products and high-concentration products, which have unique routes of administration that are relevant to cannabis-related harms (Parnes et al., 2018). For example, oral ingestion of cannabis is associated with higher concentrations of 11-hydroxy-∆9 tetrahydrocannabinol (11-hydroxy-THC), which may be more potent than ∆9-tetrahydrocannabinol (THC) (Lemberger et al., 1973; Schwilke et al., 2009), and may lead to delayed onset of psychoactive effects, which leads to unintentional overintoxication (we avoid using the term overdose given that the primary intoxicating chemical in cannabis is non-toxic and non-lethal). High concentration products can be smoked with an assortment of essential equipment but can also be vaped in a concealable vape pen. Use of concentrates is associated with rapid and higher levels of intoxication compared to flower products (Bidwell et al., 2020).

Source of Purchase Context

An outer situational context of one's cannabis use includes how one obtains cannabis products. In the early days of recreational cannabis legalization in Los Angeles (i.e., 2016-2017), young adults who purchased products from cannabis dispensaries (compared to obtaining from family or friends) reported spending more money on cannabis, using more distinct cannabis products, using more frequently, using higher

quantities, using alone more often, and experienced higher negative cannabis-related consequences and cannabis use disorder symptoms (D'Amico et al., 2020).

Brief Machine Learning Overview

Machine learning approaches differ from traditional statistical approaches in several ways. First, traditional statistical approaches have focused on questions of inference, or using probabilities to test hypotheses describing how and why variables are related. Machine learning algorithms largely focus on answering questions related to prediction, or using existing data to find patterns that predict a precise outcome (Bzdok et al., 2018). While statistical models rely on parametric assumptions about the relationship between predictors and an outcome, machine learning algorithms do not and look for complex interactions to make the best prediction (Lantz, 2019; Witten & Frank, 2002). For example, multiple regressions use independent variables, as the predictors need to be independent from each other to not affect the standard errors of other predictors. As such, multicollinearity occurs when an independent variable is highly correlated with other independent variables resulting in unstable coefficients and problems with model convergence (Allen, 1997). Machine learning models such as random forests are less affected by correlated variables, as they do not attempt to isolate the effects of a single variable on an outcome when looking for complex interactions to make predictions. Though, multi-collinearity can slightly affect the selection of important variables (Strobl et al., 2008).

Machine learning models offer unique advantages in examining outcomes, specifically regarding their ability to make precise predictions. However, a trade-off exists such that improved prediction is balanced by a loss in explaining outcomes (inference) as no coefficients are provided examining direct relationships between predictors and outcomes. Machine learning algorithms have been used to examine cannabis-related outcomes such as consequences from use (Schwebel et al., 2022), cannabis use in daily life (Yu et al., 2023), and to examine the risk and protective factors of cannabis use (Henry et al., 2024).

The Present Study

Prior research has established relationships between constructs or latent profiles of cannabis contextual factors and cannabis protective behavioral strategy use, cannabis use consequences, and cannabis use disorder severity (Beck et al., 2009; Dyar et al., 2021; Gray et al., 2024; Parnes et al., 2018). However, grouping contextual factors together through variable or person-centered approaches limits the ability to identify specific contexts that may be of importance to predicting cannabis use outcomes. We aimed to extend prior research by using specific contextual indicators as predictors of cannabis use outcomes within three large samples of college student cannabis users. We sought to broadly characterize the social context of cannabis use among college students. We report descriptive statistics across each sample, and then used an exploratory modeling technique (random forest) to identify salient contextual predictors related to cannabis outcomes. Therefore, we examined contextual factors as separate indicators of cannabis protective behavioral strategies (Pedersen et al., 2017), negative cannabis-related consequences, and cannabis use disorder symptoms.

METHODS

Participants and Procedure

The Marijuana Outcomes Study Team (MOST) participants included college students recruited from the psychology department participant pools at 9 universities in 9 states throughout the United States who participated for research participation credit according to procedures approved by the institutional review boards at each participating university (for methodological details regarding MOST please see: Richards et al., 2021). Of 7,000 total participants, our analyses are focused on 2,077 who reported past month cannabis use. Data were collected between Fall 2016 and Spring 2017 such that at the time of data collection two states permitted recreational cannabis use (CO and WA), 3 states permitted medical cannabis use (NM, NY, and CA), and 4 states did not permit cannabis use (VA, TX, TN, and FL).

The Protective Strategies Study Team (PSST) participants included college students recruited using similar procedures from 10 universities in 10 states throughout the United States (for details regarding PSST please see: Pearson et al., 2019). Of 7,303 total participants, our analyses are focused on 2,222 who reported past month cannabis use. Data were collected between Spring 2017 and Fall 2017 such that at the time of data collection 3 states permitted recreational cannabis use (AK, CO, and WA), 1 state permitted medical cannabis use (NM), and 6 states did not permit cannabis use (ID, MO, MS, NE, VA, and WY).

The Addiction Research Team study (ART) participants included college students recruited using similar procedures from 10 universities in 8 states throughout the United States (for details regarding the method including participants and recruitment please see: Richards et al., 2022, 2023). Of 5,594 total participants, our analyses are focused on 1,397 who reported past month cannabis use. Data were collected between Spring 2020 and Fall 2020 such that at the time of data collection 4 states permitted recreational cannabis use (AK, CA, CO, and WA), 1 state permitted medical cannabis use (NM), and 3 states did not permit cannabis use (ID, VA, TX). Participants in all studies provided informed consent to participate.

In total, our analyses focused on 5700 participants (male = 2893 , female = 3702 , other gender identity $= 48$, missing $= 57$). The average age of the sample was 20.17 years $(SD = 3.36)$. Most of the participants identified as White $(White = 4110,$ American Indian/Alaska Native= 161, Asian = 568, Black/African American= 861, Native Hawaiian/Pacific Islander = 88 , and Other = 432) non-Hispanic ($n = 4487$).

Measures

Context of Cannabis Use. MOST investigators developed a broad assessment of contextual variables related to one's cannabis use to serve various purposes. This assessment characterizes the amount of money spent on cannabis; frequency, level, and length of intoxication; social and physical contexts of use; form of cannabis and route of administration; level of unplanned use; and source of cannabis (see Table 1 for the items, scales of measurement, and descriptive statistics for these items). Items in the context measure focused on proxies for direct use (e.g., money spent, subjective intoxication questions), social and setting places of use (e.g., with friends, at home), form of cannabis and route of administration (e.g., flower, concentrate, using a bong, vaporizer), and source of purchase (e.g., dispensary, black market). To focus our analyses on the predictive ability of contextual factors only, we excluded proxies of direct use. Most scale items asked participants to rate the percent of time they engaged in each contextual factor (0% - 100%). For example, participants were asked to report the percentage of time they used each form of cannabis, and totals had to equal 100%. Again, please see Table 1 for the specific items and scales of measurement regarding the context factors.

Table 1. Cannabis Use Contexts Across Datasets

	PSST MOST			ART		Total		
[Variable labels are underlined]	M	\emph{SD}	\boldsymbol{M}	SD	M	SD	\boldsymbol{M}	SD
Money Spent (Please estimate how much money	42.44	69.57	45.20	70.95	53.38	76.02	46.19	71.84
you have spent on marijuana in the past month $(\$)$.								
Typical Intoxication (On a typical marijuana use	60.77	24.41	61.64	25.14	64.09	23.13	61.84	24.41
day in the past 30 days, please indicate how high								
you get from using marijuana $(0 - 100\%)$.								
Peak Intoxication (Please indicate the highest you	71.81	26.78	73.44	26.96	75.07	23.94	73.15	26.21
have been from marijuana in the past month $(0 -$								
$100\%)$.								
Peak Frequency (What percentage of the time do	57.75	32.09	59.27	32.10	62.48	30.59	59.49	31.77
you get this high from using marijuana $(0 - 100\%)$?								
Length of Intoxication (On a typical marijuana use	3.83	13.86	3.27	5.09	2.87	2.02	3.39	8.99
day in the past 30 days, how long do you stay high								
from using marijuana (hours)?)								

÷,

 \overline{a}

Note. MOST = Marijuana Outcomes Study Team, PSST = Protective Strategies Study Team, ART = Addictions Research Team

Cannabis Protective Behavioral Strategies (PBS). We used the mean of the 17-item version (Pedersen et al., 2017) of the PBSM (Pedersen et al., 2016) to assess cannabis PBS use. Internal consistency was high in each sample $(\alpha = .903, .925,)$.902). The PBSM has been shown to be a robust protective factor associated with lower cannabis use (severity) and consequences (Pearson et al., 2017; Pedersen et al., 2018).

Negative Cannabis-Related Consequences. We used the sum of the 21-item version of the Marijuana Consequences Questionnaire (Simons et al., 2012) to measure negative cannabis-related consequences. Internal consistency was high in each sample $(a =$.859, .886, .879).

Cannabis use severity. We used the sum of the 8 item Cannabis Use Disorder Identification Test— Revised (CUDIT-R) (Adamson et al., 2010) to measure CUD symptoms. Internal consistency was adequate in each sample $(\alpha = .816, .833, .837)$.

Analysis Plan

We examined the context of use variables with means and standard deviations across the three datasets individually and joined as one dataset. Additionally, we wanted to examine potential predictive ability of contextual factors of use on cannabis use outcomes (i.e., cannabis PBS, negative cannabis-related consequences, and cannabis use disorder severity). We used machine learning, specifically random forests, to examine the potential for contextual factors to predict outcomes. Random forests are an extension of regression trees (Breiman, 2001). Regression trees use a nonparametric algorithm to create a split, or a point in a predictor that best separate the outcome variable (Strobl et al., 2009). In traditional regression trees, the output provides a single tree, or a visual representation of the algorithm's classification of the outcome. In random forests, hundreds of trees are created by randomly subsampling predictor variables at each split, and then averaging the predictions of each tree to find what variables are most important in predicting the outcome (Breiman, 2001). The same random forests procedures can also be used to impute missing data (Tang & Ishwaran, 2017).

First, we used the *missForest* (Stekhoven, 2022) package to impute all of the missing data via random forest imputation. Then, we separated the data into a training dataset (80% of the available data) that we used to run the initial random forest model and a testing dataset (20% of the available data) reserved to test the predictive ability of the model. Splitting the data in this way reduces the chances of the algorithm finding random variance and overfitting the model, as well as improves generalizability (Ho et al., 2020). We used the randomForest package (Liaw & Wiener, 2002) in $R(R)$ Core Team, 2023) to find the optimal number of random predictors (i.e., tuning) for the model to subsample at each split (mtry). Then, we ran a random forest model for each outcome variable (three models) with their respective tuning parameters with the training dataset. Finally, we used the random forest model to make predictions on the testing dataset. We report the mean absolute error (MAE; average distance between predicted and actual values), the mean squared error (average squared difference between predicted and actual values), the root mean squared error (root squared MSE), and the proportion of variance in the outcome explained by the model $(R²)$. Each model consisted of only contextual factors as predictors. The MAE and RMSE are dependent upon the scale (range) of the outcome variable, and therefore there are no general guidelines for what constitutes "acceptable" fit. However, lower values of the MAE and RMSE indicate a more accurate prediction.

RESULTS

The means and standard deviations of the percentages of endorsement across all contexts of use are reported in Table 1. In the results presented below, we report noticeable trends in all three datasets. We also include bivariate correlations between all contextual indicators and the three outcome variables to determine the directional relationship between the contextual factors and outcomes (Table 2).

Money Spent and Intoxication

Overall, participants reported they spent an average of \$46.19 on cannabis in the 30 days prior to study participation. The amount of money spent increased slightly between project MOST to project PSST and again from project PSST to project ART.

		MOST		PSST			ART		
	1	$\sqrt{2}$	$\boldsymbol{3}$	$\mathbf{1}$	$\overline{2}$	$\sqrt{3}$	1	$\overline{2}$	$\sqrt{3}$
1. CUDIT-R									
2. MACQ	$.629**$			$.609**$			$.649**$		
3. PBSM	$-424**$	$-0.363**$		$-0.397**$	$-0.291**$		$-0.489**$	$-0.363**$	
4. Source of Cannabis (dispensary1)	$.108**$	$.086**$	$-0.078**$	$.165**$	$.097**$	$-068**$	$.162**$	$.098**$	$-120**$
5 Source of Cannabis (dispensary2)	.067	.035	$-0.055*$.014	.000	-0.027	$.037\,$.046	$-0.055*$
6. Source of Cannabis (Black market)	$.355**$	$.292**$	$-318**$	$.362**$	$.257**$	$-0.287**$	$.266**$	$.198**$	$-236**$
7 Source of Cannabis (Did not	$-401**$	$-0.328**$	$.344**$	$-437**$	$-0.290**$	$.306**$	$-0.376**$	$-266**$	$.319**$
purchase)									
8. Money Spenta	$.483**$	$.408**$	$-448**$	$.526**$	$.352^{**}$	$-0.405**$	$.505**$	$.365**$	$-428**$
9. Typical Intox ^a	$.217**$	$.155***$	$-191**$	$.255**$	$.137**$	$-175**$	$.182**$	$.130**$	$-143**$
10. Peak Intox ^a	$.322**$	$.218**$	$-0.255**$	$.364^{\ast\ast}$	$.223**$	$-0.251**$	$.310**$	$.224**$	$-0.227**$
11. Peak Frequency ^a	.049	.010	$-069**$	$.067**$.001	$-0.070**$	$-0.056*$	$-0.053*$.021
12. Length of Intox ^a	.008	.003	$-0.044*$	$.072^{\ast\ast}$.026	$-0.056*$	$.028\,$.046	$\cdot.013$
13. Form of Cannabis (Plant)	.050	.010	.016	$.071**$	$.055*$	-0.013	$.070*$.023	$-079**$
14. Form of Cannabis (Edibles)	$-.110**$	$-085**$	$.099^{**}$	$-167**$	$-134**$	$.160**$	$-.211**$	$-154**$	$.148**$
15. Form of Cannabis (Concentrates)	.041	$.080**$	$-132**$	$.133**$	$.088**$	$-149**$	$.139**$	$.123**$	-0.026
16. Form of Cannabis (Other form)	.016	-0.004	.006	$-102**$	-037	-011	$-084**$	-0.37	-0.023
17. Route of Administration (joint)	-0.39	-009	.003	$\cdot.057^{*}$	-0.026	-006	$.004$	$-0.055*$.012
18. Route of Administration (joint tobacco)	$.094*$	$.067**$	-0.025	$.053^{\ast}$	$.049*$	-0.34	.045	-0.013	-0.044
19. Route of Administration (bong)	$.134**$	$.073**$	$-0.00**$	$.159**$	$.129**$	$-0.096**$	$.167**$	$.154**$	$-141**$
20. Route of Administration (bong tobacco)	$.107**$	$.084**$	-039	$.119**$	$.112^{\ast\ast}$	$-0.90**$	$.109^{**}$	$.103**$	$-0.097**$
21. Route of Administration (bowl)	-065	$-0.54*$	$.054*$.000	$-0.048*$	$.034\,$	-0.040	-0.014	$.027\,$
22. Route of Administration (bowl tobacco)	-0.018	-0.007	-008	-0.005	$.025\,$	-0.023	-0.016	.030	-0.002
23. Route of Administration (eaten)	$-161**$	$-0.072**$	$.072**$	$-181**$	$-136**$	$.175**$	$-0.210**$	$-162**$	$.128**$
24. Route of Administration (vaporizer)	$.035\,$	-0.36	$.002$	-0.015	-0.021	-0.023	.007	$.030\,$	$.034\,$
25. Setting of Use (At home)	$.211**$	$.141**$	$-152**$	$.199**$	$.123^{\ast\ast}$	$-121**$	$.206**$	$.140**$	$-184**$
26. Setting of Use (At friend's)	$-163**$	$-101**$	$.155**$	$-190**$	$-119**$	$.138**$	$-218**$	$-162**$	$.177**$
27. Setting of Use (At stranger's)	-0.035	.010	-0.010	.006	.006	$-087**$	-006	.021	.002
28. Setting of Use (outside)	-0.023	-0.36	-010	-0.014	-0.007	-0.025	$-0.059*$	-0.34	$.064*$
29. Setting of Use (car)	$.012\,$.042	-030	.014	$.030\,$.009	$.101**$	$.054*$	$-068*$
30. Setting of Use (party)	$-0.00*$	$-0.072**$	$.055*$	$-0.049*$	$-0.070**$	$.046*$	-051	-0.024	$.070**$
31. Setting of Use (Other place)	.014	-0.028	-009	-0.023	.005	-0.010	$-.091**$	-0.025	.030
32. Social Context of Use (Alone)	$.300**$	$.238^{**}$	$-0.261**$	$.332^{\ast\ast}$	$.224**$	$-0.272**$	$.278**$	$.185**$	$-0.297**$
33. Social Context of Use (With	$-304**$	$-217**$	$.265**$	$-0.281**$	$-182**$	$.257**$	$-0.249**$	$-166**$	$.287**$
friends)									
34. Social Context of Use (With family)	$.022\,$	-0.005	$-0.056*$	-0.015	-0.35	-017	.032	.016	-0.007
35. Social Context of Use (With strangers)	.074	$.067**$	-042	$.046*$	$.080^{\ast\ast}$	$-0.091**$	-0.025	.002	-0.014
36. Social Context of Use (With others)	.058	.019	-039	-0.09	-0.002	-0.005	-0.048	-0.30	-0.044
37. Unplanned Use ^a	$-348**$	$-0.200**$	$.206**$	$-0.280**$	$-141**$	$.125**$	$-175**$	$-105**$	$.067*$

Table 2. Raw correlations between cannabis use context variables and cannabis-related outcomes across each dataset

Note. * $p < .05$, ** $p < .01$, a Proxy of direct use or not a social context and removed from analyses.

Cannabis, A Publication of the Research Society on Marijuana

Participants were asked to rate on a $0 - 100$ scale how high they typically get when they use cannabis. In each of the three studies, participants reported percentages in the 60-65 range with an overall average of 61.835. Participants were also asked to rate their highest level of intoxication on the same scale. In all three studies, participants reported the highest level of intoxication in the 70s (Overall $M = 73.151$ and that they achieve this peak intoxication over half of the times they use cannabis (Overall $M = 59.487\%$). On days participants used cannabis in the past month they reported feeling high for an average of 3.392 hours.

Cannabis Form

In all three studies, participants reported using cannabis flower most of the time (Overall $M = 70.384\%$. However, a notable drop in cannabis flower use occurred between project MOST $(M = 78.646\%)$ and project ART $(M =$ 53.265%). The drop in the use of flower corresponded with a similar increase in the use of concentrates (MOST $M = 8.621\%$; ART $M =$ 26.073%).

Route of Administration

The most prominent route of administration in projects MOST and PSST was smoking a joint or blunt without tobacco. Participants in project ART endorsed using a joint or blunt without tobacco $(M = 22.205\%)$ and using a vaporizer (M) = 23.902%) at similar rates. Reported vaporizer use in project ART was much higher than reported vaporizer use in projects MOST $(M =$ 5.787%) and PSST $(M = 8.072\%)$.

Use Settings

Overall, using cannabis at home was the most endorsed setting. However, in project MOST and PSST, participants tended to use at home or at a friend's house at about the same frequencies. Compared to MOST and PSST, participants in project ART appeared to make a trade-off between using cannabis at their own house $(M = 53.180\%)$ and their friend's house $(24.540\%).$

Social Context of Use

Overall, participants reported mostly using cannabis with their friends in all three studies. One notable difference between the three studies is that participants in project ART reported using alone $(M = 30.418\%)$ more often than participants in MOST $(M = 15.884\%)$ and PSST $(M = 17.396\%).$

Source of Cannabis and Money Spent

Across all three projects, participants mostly endorsed not sourcing cannabis themselves. In projects MOST and PSST, the second most endorsed source was sourcing cannabis from a place other than a dispensary. In project ART, the second most endorsed source was obtaining cannabis from a dispensary. In all studies, below 5% of cannabis sourcing involved crossing state lines to purchase at a dispensary in another state.

Random Forest Models

First, we combined all three datasets and imputed missing values using the missForest package (Stekhoven, 2022). The number of missing values for the MACQ (MOST $= 1.97\%$; $PSST = 1.89\%, ART = 1.58\%$ and PBSM (MOST $= 2.27\%$; PSST = 1.75%, ART = 1.58%) was acceptable in all three datasets. Regarding the CUDIT-R, participants in project MOST were randomly assigned to complete one of four measures of cannabis use disorder symptoms, one of which was the CUDIT-R. Thus, missingness for the CUDIT-R in project MOST was high (69.57%). Missingness for the CUDIT-R in projects PSST and ART were acceptable (PSST = 1.80% , ART = 4.94%). The missForest package subsets the data into complete cases and variables with missing data. The package then runs a random forest algorithm based on the observed values to impute a value for missing data (Stekhoven, 2022).

After data imputation, we split the dataset into a training dataset and a testing dataset. For each outcome (PBSM, MACQ, CUDIT-R) we conducted a tuning model that examined the optimal number of variables randomly sampled at each split of the decision trees. The optimal number of variables randomly sampled for the CUDIT-R, PBSM, and MACQ models was 5. Finally, we conducted random forest models for all three variables using the selected number of splits, 500

Cannabis Use Context

decision trees, in the *randomForest* package (Liaw) & Wiener, 2002) in $R(R \text{ Core Team}, 2022)$. Below, we report the variable importance, or predictive utility of a variable across all the decision trees in a random forest, from the training dataset and model fit from using the training datasets on the testing datasets.

For the PBSM, our rank-ordered variable importance plot can be viewed in Figure 1. Using alone, using with friends, obtaining cannabis on the black market, using concentrate, and using at home were the most important variables in predicting the PBSM. However, when using the random forest model to predict values in the testing dataset, the random forest predictions had room for improvement (Table 3). On average, our model's predicted values deviated from the true values (MAE) by 0.74 units of the PBSM (range $1-6$). The squared differences between the predicted and actual values (MSE) was 0.94, and our model accounted for 16% of the variance in PBSM scores.

For the MACQ, our rank ordered variable importance plot can be viewed in Figure 2. Using alone, obtaining cannabis on the black market,

using with friends, primarily using a bong, and using at home were the most important predictors for the MACQ. We used the training model to predict MACQ scores in the portion of data set aside for predictions (Table 3). On average, our model's predicted values deviated from the true values (MAE) by 2.68 units of the MACQ (range $0 - 21$). The squared differences between the predicted and actual values (MSE) was 13.02. Overall, our random forest model of contextual factors accounted for 17% of the variance in the MACQ.

The rank ordered variable importance plot for our random forest model predicting the CUDIT-R can be viewed in Figure 3. The most important variables in predicting the CUDIT-R were using alone, obtaining cannabis on the black market, using with friends, using at home, and using a bong. On average, the model's predicted values deviated from the true values (MAE) by 3.12 units of the CUDIT-R sum (range $= 0 - 32$). The squared differences between the predicted and actual values was 18.79 and the model accounted for 38% of the variance in the CUDIT-R sum.

Figure 1. Plot of variable importance for the PBSM in order from least important (top) to most important (bottom).

Note. The $(+)$ and $(.)$ after each contextual variable indicates the directional relationship to the PBSM.

Note. The $(+)$ and $(.)$ after each contextual variable indicates the directional relationship to the MACQ.

Figure 3. Plot of variable importance for the CUDIT-R in order from least important (top) to most important (bottom).

Note. The $(+)$ and $(.)$ after each contextual variable indicates the directional relationship to the CUDIT-R.

Outcome	mtry	R^2	MAE	MSE	RMSE
Protective Behavioral	5	0.16	0.75	0.94	0.97
Strategies (PBSM)					
Negative Consequences	5	0.17	2.68	13.02	3.46
(MACQ)					
Cannabis Use Severity	5	0.46	2.89	16.59	4.07
$(CUDIT-R)$					

Table 3. Random Forest Fit Statistics

 $Note.$ PBSM= Protective Behavioral Strategies for Marijuana, MACQ = Marijuana Consequences Questionnaire, CUDIT-R = Cannabis Use Disorder Identification Test-Revised, mtry= the optimal number of random predictors (i.e., tuning) for the model to subsample at each split, MAE = Mean Absolute Error, MSE = Mean Squared Error, RMSE = Root Mean Squared Error.

DISCUSSION

Overall, the present study extends the previous literature, which has largely focused on using a few contextual variables to determine factor structures or latent profiles (Beck et al., 2009; Gray et al., 2024; Spinella et al., 2019), by providing descriptive statistics across a broad array of social contexts of cannabis use. Regarding the form of cannabis used, college students appear to predominately use flower cannabis, though the use of edibles and concentrates was not minimal. Bivariate correlations between cannabis form and cannabis outcomes (Table 2) indicated that edible usage was most consistently correlated (compared to other forms of cannabis) with less use disorder severity and consequences, and more PBS use. Though, the use of concentrates was significantly correlated with less PBS use. Of note regarding cannabis form, the use of concentrate was the most important cannabis form predictor in all three random forest models. This in part could be due to the greater exposure to THC when using concentrates versus flower (Bidwell et al., 2020).

Regarding different routes of administration, the use of a bong was most consistently correlated with increased consequences and disorder severity and decreased use of PBS, and similar to form results, eating cannabis appeared to be the most protective route of administration (Table 2). Additionally, using a bong tended to be the most important route of administration in the random forest models, other than for PBS, where using a joint was slightly more important. Bongs tend to be relatively indiscreet and would likely be owned by individuals that consume cannabis regularly, though more work is needed to determine why bong use specifically may be associated with worse outcomes. College students predominately consume using joints without tobacco and co-use with tobacco was not highly endorsed.

Regarding direct social contexts of use (who participants used with), participants tended to use with friends, and using with friends was the most consistent social context correlated with fewer consequences and disorder severity and more PBS use. In contrast, using alone was the most consistent social context correlated with negative cannabis outcomes, consistent with previous literature (Table 2; Buckner et al., 2016; Creswell et al., 2015). Using alone was the most important contextual factor in all of the random forest models. Considering that solitary use accounts for much of the relationship between social anxiety and poor cannabis use outcomes (Buckner et al ., 2016), it may be that solitary use is more associated with negative reinforcement, or using to remove unwanted emotional states. Given the fact that using alone was the most important predictor in all models, this may highlight the need for preventative and clinical treatments to focus on decreasing the amount of time individuals use cannabis alone. Regarding where individuals used, using at a friend's house was consistently correlated with positive cannabis outcomes, while using at home was associated with negative outcomes (Table 2). Additionally, using cannabis at home was the most important social setting in all three random forest models. Using at home and using alone are potentially conflated and our models cannot differentiate whether participants used at home alone or with friends. Future work should focus on examining social networks of individuals that often use at home and whether including others may have protective effects on cannabis outcomes.

Regarding source of cannabis purchase, the most protective source of purchase was not purchasing cannabis, and purchasing on the black market was the source most strongly correlated with negative outcomes (Table 2). In fact, sourcing cannabis on the black market was the most important source context in all of the random forest models. This likely indicates individuals that often purchase cannabis or go out of their way to source cannabis in places that do not have the same tax burden as a legalized market. One important note is that crossing state lines to source cannabis was not highly endorsed in any of the studies.

Lastly, the predictive models accounted for varying proportions of the cannabis outcomes' variances. Specifically, contextual factors accounted for 16% and 17% of the variance in PBS and cannabis use consequences respectively and 38% of variance in the CUDIT-R. Considering our models removed proxies of use (i.e., level of intoxication, money spent), our results indicate that contextual factors likely account for additional variance in cannabis-related outcomes beyond direct use. Additionally, the random forest models were able to predict outcomes relatively well. Recalling that the MAE is the average error of the model's prediction in the same scale as the outcome, the model's relative errors were all within 9%-13% of the outcome variables' range and may provide a benchmark for future studies using machine learning with cannabis contexts.

Limitations

Our study has several limitations. First, we created our contextual measurement tool, and said tool has not been validated for real-time use. Second, our models do not account for the legal status of cannabis in the participant's state of residence. As such, we do not know the status of how participants sourced cannabis. While the rates of crossing state lines to obtain cannabis were low in all three studies, the rates may change depending on the legal status of each state, and how far away the participants were from a dispensary. Lastly, while the studies were conducted over a span of 4 years, we do not make any assumptions regarding the trends of cannabis use as state-level legalization has become more widespread over time. This is especially relevant as the COVID-19 pandemic had not occurred

during PSST and MOST data collection but had already occurred during the ART data collection. Future work should focus on changes in these contextual trends and determine how potential changes may affect outcomes.

Future Directions

Much work needs to be done regarding the contexts surrounding cannabis use. Specifically, participants in our study tended to report using cannabis with friends. Under the assumption that friends using cannabis together are not using their own pipes, bongs, or joints, it is likely difficult to accurately measure the amount of cannabis consumed by everyone, even if the weight and potency are known prior to group consumption. For example, even if study participants are asked to report the potency and to pre-measure the weight of each joint/bowl in real time, there is no way to know what percentage of that weight in combusted THC that everyone in a group session is consuming. This predicament contrasts with alcohol use, where more accurate measurements can be assumed by standard drink conversions. It may be that measuring additional contexts such as subjective intoxication, money spent, and form of cannabis use can be appropriate proxies for precise dosage and weights.

The incorporation of assessing contexts of use could also provide pertinent information regarding environmental factors related to use. Implementing contextual measures from a theoretical framework could help improve existing models predicting cannabis use outcomes. That is, what are the effects of core predictors of cannabis use outcomes when incorporating environmental factors into existing models? Much of the modeling on cannabis use outcomes examine outcomes as functions of use, emotions, or urges. However, it is likely the predictors fluctuate between different contexts of use.

REFERENCES

Adamson, S. J., Kay-Lambkin, F. J., Baker, A. L., Lewin, T. J., Thornton, L., Kelly, B. J., & Sellman, J. D. (2010). An improved brief measure of cannabis misuse: The Cannabis Use Disorders Identification Test-Revised

(CUDIT-R). Drug and Alcohol Dependence, $110(1-2)$, $137-143$. https://doi.org/10.1016/j.drugalcdep.2010.02.0 17

- Bidwell, L. C., YorkWilliams, S. L., Mueller, R. L., Bryan, A. D., & Hutchison, K. E. (2018). Exploring cannabis concentrates on the legal market: User profiles, product strength, and health-related outcomes. Addictive Behaviors Reports, δ March), 102–106. https://doi.org/10.1016/j.abrep.2018.08.004
- Bravo, A. J., Pearson, M. R., Conner, B. T., & Parnes, J. E. (2017). Is 4/20 an event-specific marijuana holiday? A daily diary investigation of marijuana use and consequences among college students. Journal of Studies on Alcohol and Drugs. 78(1), 134–139. https://doi.org/10.15288/jsad.2017.78.134
- Bravo, A. J., Villarosa-Hurlocker, M. C., & Pearson, M. R. (2018). College Student Mental Health: An Evaluation of the DSM-5 Self-Rated Level 1 Cross-Cutting Symptom Measure. Psychological Assessment, 30(10), 1382–1389.

https://doi.org/10.1037/pas0000628

- Breiman, L. (2001). Random Forests. Machine *Learning*, $45(1)$, $5-32$. https://doi.org/10.1023/A:1010933404324
- Buckner, J. D., Ecker, A. H., & Dean, K. E. (2016). Solitary cannabis use frequency mediates the relationship between social anxiety and cannabis use and related problems. American Journal on Addictions, 25(2), 99–104. https://doi.org/10.1111/ajad.12339
- Cloutier, R. M., Calhoun, B. H., & Linden-Carmichae, A. N. (2022). Associations of mode of administration on cannabis consumption and subjective intoxication in daily life. Psychology of Addictive Behaviors, 36(1), 67– 77. https://doi.org/10.1037/adb0000726
- Creswell, K. G., Chung, T., Clark, D. B., & Martin, C. S. (2015). Solitary cannabis use in adolescence as a correlate and predictor of cannabis problems. Drug and Alcohol Dependence, 156(2015), 120–125. https://doi.org/10.1016/j.drugalcdep.2015.08.0 27
- D'Amico, E. J., Rodriguez, A., Dunbar, M. S., Firth, C. L., Tucker, J. S., Seelam, R., Pedersen, E. R., & Davis, J. P. (2020). Sources of cannabis among young adults and associations with cannabis-related outcomes.

International Journal of Drug Policy, 86, 102971.

https://doi.org/10.1016/j.drugpo.2020.102971

Gray, K. M., Watson, N. L., & Christie, D. K. (2009). Challenges in quantifying marijuana use. American Journal on Addictions, 18(2), 178–179.

https://doi.org/10.1080/10550490902772579

- Ho, S. Y., Phua, K., Wong, L., & Bin Goh, W. W. (2020). Extensions of the external validation for checking learned model interpretability and generalizability. Patterns, 1(8), 100129. https://doi.org/10.1016/j.patter.2020.100129
- Lemberger, L., Martz, R., & Rodda, B. (1973). Comparative pharmacology of Δ9 tetrahydrocannabinol and its metabolite, 11 OH Δ9 tetrahydrocannabinol. Journal of Clinical Investigation, $52(10)$, $2411-2417$. https://doi.org/10.1172/JCI107431
- Liaw, A., & Wiener, M. (2002). Classification and regression by random Forest. R News, $2(3)$, 18– 22. https://cran.rproject.org/web/packages/randomForest/
- Parnes, J. E., Bravo, A. J., Conner, B. T., & Pearson, M. R. (2018). A burning problem: Cannabis lessons learned from Colorado. Addiction Research and Theory, $26(1)$, $3-10$. https://doi.org/10.1080/16066359.2017.131541 Ω
- Pearson, M. R., Bravo, A. J., & Protective Strategies Study Team. (2019). Marijuana protective behavioral strategies and marijuana refusal self-efficacy: Independent and interactive effects on marijuana-related outcomes. Psychology of Addictive Behaviors, $33(4)$, $412-419$. https://doi.org/10.1037/adb0000445

Pearson, M. R., & Henson, J. M. (2013). Unplanned drinking and alcohol-related problems: A preliminary test of the model of unplanned drinking behavior. Psychology of Addictive Behaviors, $27(3)$, 584–595. https://doi.org/10.1037/a0030901

Pearson, M. R., Liese, B. S., Dvorak, R. D., Anthenien, A. M., Bravo, A. J., Conner, B. T., Correia, C. J., Dvorak, R. D., Egerton, G. A., Hustad, J. T. P., Kholodkov, T., King, K. M., Liese, B. S., Messina, B. G., Murphy, J. G., Neighbors, C., Nguyen, X. T., Parnes, J. E., Pearson, M. R., … Read, J. P. (2017). College student marijuana involvement: Perceptions, use, and consequences across 11 college campuses. Addictive Behaviors, 66, 83–89. https://doi.org/10.1016/j.addbeh.2016.10.019

- Pedersen, E. R., Huang, W., Dvorak, R. D., Prince, M. A., Hummer, J. F., & (The Marijuana Outcomes Study Team) (2017). The Protective Behavioral Strategies for Marijuana Scale: Further examination using item response theory. Psychology of Addictive Behaviors, $31(5)$, 548–559. https://doi.org/10.1037/adb0000271
- Pedersen, E. R., Hummer, J. F., Rinker, D. V., Traylor, Z. K., & Neighbors, C. Measuring Protective Behavioral Strategies for Marijuana Use Among Young Adults (2016). Journal of Studies on Alcohol and Drugs, $77(3)$, $441-50$.
- https://doi.org/10.15288/jsad.2016.77.441
- Pedersen, E. R., Villarosa-Hurlocker, M. C., & Prince, M. A. (2018). Use of Protective Behavioral Strategies among Young Adult Veteran Marijuana Users. Cannabis, 1(1), 14– 27.

https://doi.org/10.26828/cannabis.2018.01.002

- Prince, M. A., Conner, B. T., & Pearson, M. R. (2018). Quantifying Cannabis: A field study of marijuana quantity estimation. Psychology of Addictive Behaviors, 32(4), 426–433. https://doi.org/10.1037/adb0000370
- R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-Project.org/
- Richards, D. K., Pearson, M. R., & Field, C. A. (2021). A Comprehensive Examination of Alcohol-Related Motivations Among College Students: Unique Relations of Drinking Motives and Motivations for Drinking Responsibly. Experimental and Clinical Psychopharmacology.

https://doi.org/10.1037/pha0000526

- Richards, D. K., Schwebel, F. J., Field, C. A., & Pearson, M. R. (2023). The Associations of Basic Psychological Need Satisfaction and Need Frustration with Cannabis-Related Outcomes in a Multi-Site Sample of College Students. Journal of Psychoactive Drugs, $00(00)$, 1–10. https://doi.org/10.1080/02791072.2023.219160 5
- Richards, D. K., Schwebel, F. J., Sotelo, M., Pearson, M. R., & Marijuana Outcomes Study Team. (2021). Self-Reported Symptoms of

Cannabis Use Disorder: Psychometric testing and validation. Experimental and Clinical $Psychopharmacology, \t29(2), \t157-165.$ https://doi.org/10.1037/pha0000455

- Schubart, C. D., Boks, M. P. M., Breetvelt, E. J., van Gastel, W. A., Groenwold, R. H. H., Ophoff, R. A., Sommer, I. E. C., & Kahn, R. S. (2011). Association between cannabis and psychiatric hospitalization. Acta Psychiatrica Scandinavica, 123(5), 368–375. https://doi.org/10.1111/j.1600- 0447.2010.01640.x
- Schwebel, F. J., Richards, D. K., Pfund, R. A., Joseph, V. W., & Pearson, M. R. (2022). Using Decision Trees to Identify Salient Predictors of Cannabis-Related Outcomes. Journal of Psychoactive Drugs, 1–10. https://doi.org/10.1080/02791072.2021.201408 1
- Schwilke, E. W., Schwope, D. M., Karschner, E. L., Lowe, R. H., Darwin, W. D., Kelly, D. L., Goodwin, R. S., Gorelick, D. A., & Huestis, M. A. (2009). Δ9-tetrahydrocannabinol (THC), 11 hydroxy-THC, and 11-nor-9-carboxy-THC plasma pharmacokinetics during and after continuous high-dose oral THC. Clinical Chemistry, 55(12), 2180–2189. https://doi.org/10.1373/clinchem.2008.122119
- Simons, J. S., Dvorak, R. D., Merrill, J. E., & Read, J. P. (2012). Dimensions and severity of marijuana consequences: Development and validation of the Marijuana Consequences Questionnaire (MACQ). Addictive Behaviors, $37(5)$, 613–621.

https://doi.org/10.1016/j.addbeh.2012.01.008

- Spinella, T. C., Stewart, S. H., & Barrett, S. P. (2019). Context matters: Characteristics of solitary versus social cannabis use. Drug and Alcohol Review, 38(3), 316–320. https://doi.org/10.1111/dar.12912
- Stekhoven, D. J. (2022). *MissForest:* Nonparametric Missing Value Imputation using Random Forest. R package version 1.5. https://cran.rproject.org/web/packages/missForest/index.ht ml
- Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random Forests. Psychological Methods,

 $14(4)$, $323-348$. https://doi.org/10.1037/a0016973 Tang, F., & Ishwaran, H. (2017). Random forest missing data algorithms. Statistical Analysis and Data Mining: The ASA Data Science $Journal,$ $10(6),$ $363-377.$ https://doi.org/10.1002/sam.11348

Funding and Acknowledgements: Matison McCool was supported in part by the National Institute on Alcohol Abuse and Alcoholism, T32AA018108 (PI: Witkiewitz).

Conflict of interest statement: We have no conflict of interest to declare.

We would like to acknowledge the efforts of Sarah L. Simons with conducting literature searches and contributing to an early version of this manuscript. Data were collected by three research teams: Marijuana Outcomes Study Team (MOST), Protective Strategies Study Team (PSST), and the Addictions Research Team (ART).

*MOST includes the following investigators (in alphabetical order): Amber M. Anthenien, University of Houston; Adrian J. Bravo, University of New Mexico; Bradley T. Conner, Colorado State University; Christopher J. Correia, Auburn University; Robert D. Dvorak, University of Central Florida; Gregory A. Egerton, University at Buffalo; John T. P. Hustad, Pennsylvania State University College of Medicine; Tatyana Kholodkov, University of Wyoming; Kevin M. King, University of Washington; Bruce S. Liese, University of Kansas; Bryan G. Messina, Auburn University; James G. Murphy, The University of Memphis; Clayton Neighbors, University of Houston; Xuan-Thanh Nguyen, University of California, Los Angeles; Jamie E. Parnes, Colorado State University; Matthew R. Pearson, University of New Mexico; Eric R. Pedersen, RAND; Mark A. Prince, Colorado State University; Sharon A. Radomski, University at Buffalo; Lara A. Ray, University of California, Los Angeles; Jennifer P. Read, University at Buffalo.

**PSST includes the following investigators: Matthew R. Pearson, University of New Mexico (Coordinating PI); Adrian J. Bravo, University of New Mexico (Co-PI); Mark A. Prince, Colorado State University (site PI); Michael B. Madson, University of Southern Mississippi (site PI); James M. Henson, Old Dominion University (site PI); Alison Looby, University of Wyoming (site PI); Vivian M. Gonzalez, University of Alaska-Anchorage (site PI); Amber M. Henslee, Missouri Science & Technology (site PI); Carrie Cuttler, Washington State University (site PI), Maria M. Wong, Idaho State University (site PI), Dennis E. McChargue, University of Nebraska-Lincoln (site PI).

***ART includes the following investigators: Matthew R. Pearson, University of New Mexico (Coordinating PI); Adrian J. Bravo, William & Mary (site PI); Bradley T. Conner, Colorado State University – Fort Collins (site PI); Carrie Cuttler, Washington State University (site PI); Craig A. Field, University of Texas at El Paso (site PI); Vivian Gonzalez, University of Alaska - Anchorage (site PI); James M. Henson, Old Dominion University (site PI); Jon M. Houck, Mind Research Network; Kevin M. King, University of Washington (site PI); Benjamin O. Ladd, Washington State University (site PI); Kevin S. Montes, California State University – Dominguez Hills (site PI); Mark A. Prince, Colorado State University – Fort Collins (site PI); Maria M. Wong, Idaho State University (site PI).

Copyright: © 2024 Authors et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License,](https://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, distribution, and reproduction, provided the original author and source are credited, the original sources is not modified, and the source is not used for commercial purposes.

