

Music Groove: Stimuli Affecting Cognitive Ability

UC Berkeley School of Information

W241.3 Spring 2019: Experiments and Causality: Final Project

Danielle Adler, Osmar Coronel, Dan Kent

A dark blue diagonal gradient bar that starts from the bottom left and extends towards the top right, covering the lower half of the slide.

Research Question: Do Different Stimuli Impact Cognitive Ability?

Personal Life Motivation:

- Some people have same song on repeat when writing essays
- Others like very loud music to code
- These same people may like classical music to read
- Can we say anything general about music preference?

Research and Use Case Motivation:

- U.S. Government - used different songs as a form of torture
- Journal of Consumer Research - high level of noise hurts creativity
- The Journal of the Acoustical Society of America - ambient, natural sounds could improve productivity and moods

Hypothesis: Unpleasant Music will Cause Worse Performance

Treatment Stimuli:

- Barney and Friends Theme Song
(Treatment; unpleasant song)



- White Noise (Control; neutral noise)



Cognitive Function Mechanism:

- Randomized Sudoku Board: goal is to play for 5 minutes or until complete



- Pairwise analysis completed where each participant played two games, one with each treatment with a randomized order

Experimental Design: Combination of Qualtrics and Mturk

Step 1: Recruitment via Mturk Masters (paid \$2.15; max 35 minutes)



Step 2: Gather Demographic Metrics on Qualtrics (knowledge of Sudoku, morning / night person, location, native language, minutes spent playing games per week, education-level)



Step 3 (1st Game): Randomize order of song treatment and play Sudoku; listen for specific number during song to check for compliance

Step 4 (2nd Game): Play Sudoku with the second song option

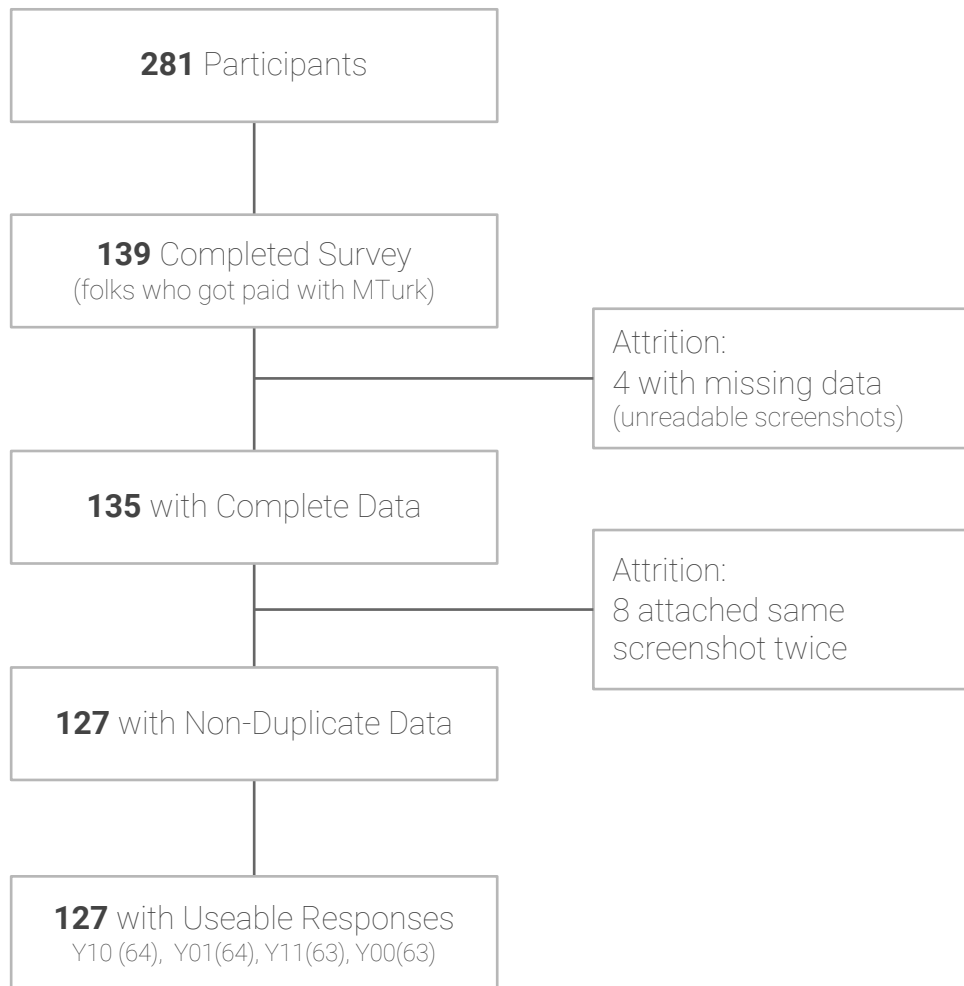


Step 5: Use the Qualtrics-generated code to input into MTurk to validate response

Experimental Design: Flow Chart

Attrition: Larger issue with missing data, and duplication of first puzzle so second puzzle proper outcomes were not visible

Compliance: Did not consider people with missing data to be noncompliant as they had to see all treatment to complete the survey



Experimental Design: Randomization and Experimentation Overview

ROXO Methodology - Randomization leading to a treatment or control:



E = set **e**quivalent groups

O = **o**bservational information (demographics, initial questions, response back into MTurk)

R = **r**andomized order of treatment and control

X = treatment

-- = control

Experimental Design: Potential Outcomes / Outcome Measures

Potential Outcome Table

	Which Experiment (Puzzle Number)	
	Time 0	Time 1
Which Version (Specific YouTube Video)	Control First (White Noise): Y00	Treatment Second (Barney): Y11
	Treatment First: (Barney): Y10	Control Second (White Noise): Y01

Experimental Outcome Measures

Completion: Within-subject outcome where one puzzle is completed and compared to another which is not

of White Spaces: Within-subject outcome where # of spaces left is compared between both games

Completion of the First Game: Between subject outcome where completion of the first game is compared

Covariate Balance Check

All covariates show balance besides one outlier

	Treatment → Control Barney (Y10) → White Noise (Y01)	Control → Treatment White Noise (Y00) → Barney (Y11)
Number of Respondents with Usable Data	64	63
Treatment	1	0
Average Age	35.50	36.57
Average of Morning (1), Night (0)	0.594	0.508
Average Logic Game Expertise (out of 5)	3.08 / 5	2.86 / 5
Average Time Playing Games per Week	62.28 min	32.10 min
Average Sudoku Skills (out of 5)	2.56 / 5	2.40 / 5

Analysis of Within–Subject Results

Table 10: Within-Subject Outcome Results

	<i>Dependent variable:</i>	
	White_spaces Model of White Spaces	Completion Model of Game Completions
	(1)	(2)
Treat	0.868 (0.951)	−0.039 (0.046)
Puzzle_Num	−0.710 (0.951)	−0.008 (0.046)
Constant	−0.079 (0.657)	1.024*** (0.030)
ID Fixed Effects	Yes	Yes
P-value	0.363	0.393
Observations	254	254
R ²	0.846	0.866
Adjusted R ²	0.688	0.728
Residual Std. Error (df = 125)	5.351	0.259
<i>Note:</i> *p<0.05; **p<0.01; ***p<0.001		

White Spaces Results:

- The treatment (Barney music) increases the # of white spaces by 0.868 so just under one additional white space, which is significant.
- We observe a 36.3% chance that the difference in white spaces between the subject's two games happened randomly

Game Completion Results:

- The treatment (Barney music) decreases the likelihood of a completed game by 3.94 percentage points which is not significant.
- We observe a 39.3% chance that the difference in completed games between the subject's two games happened randomly

Analysis of Between Subject Results

Table 12: Between Subject Outcome Result

	<i>Dependent variable:</i>	
	Completion	
	First Game Completion: Basic Model	First Game Completion: Full Model
	(1)	(2)
Treat	-0.038 (0.089)	-0.038 (0.090)
Age_2		0.026 (0.519)
Expertise_logic_sudoku_2		1.029*** (0.218)
Expertise_logic_games_2		-0.628 (0.408)
Morning		-0.022 (0.089)
Degree_Interval		-0.163 (0.195)
Constant	0.460*** (0.064)	0.435 (0.286)
ID Fixed Effects	No	No
P-value	0.668	0.673
Observations	127	127
R ²	0.001	0.154
Adjusted R ²	-0.006	0.111
Residual Std. Error	0.500 (df = 125)	0.470 (df = 120)

Note:

*p<0.05; **p<0.01; ***p<0.001

- Treatment (Barney music) decreases the completion rate by 3.84 percentage points, which is not statistically or practically significant
- We observe a ~67% chance that the difference in first game completion happened randomly
- The treatment coefficient is almost identical between the basic and full model showing that the additional coefficients do not make a difference in the treatment coefficient
- In this situation, our constant represents the mean completion rate when someone begins with the control (white noise) music.

Conclusions

- **Music Does Not Statistically Significantly Affect Cognitive Ability** - Forcing participants to listen to Barney music does not affect their ability to complete Sudoku puzzles (i.e. cognitive functions) quickly
- **Pairwise Analysis Works Well** - When we are comparing within the same population, we do not have to worry about nuanced interpretations as we are clustering for individuals
- **Data Collection is Hard** - Initially, we tried to administer each Sudoku game but we realized facilitation bias, time, and diverse population would all be compromised
- **Pilots are Necessary** - The first time we ran our MTurk pilot, we did not have validation procedures in place and could not tell which respondents completed correctly
- **Some MTurk Participants try Hard** - We expected a low completion rate given that we had no specific incentive for completion, but over $\frac{1}{3}$ completed both puzzles
- **Group Question** - Do you agree with our labeling of attrition and compliance?

Questions?

