Preregistration Report

Study Information

1. Title
   1.1. Activation of basic-level categories in the human brain based on different cue types.

2. Authors
   2.1. Jacob Dudek
   2.2. Elias Harjes
   2.3. Joel Haupt
   2.4. Gerrit Bartels

3. Description
   3.1. Participants take part in an online experiment that was implemented using _magpie, a minimal architecture for the generation of portable (browser-based) interactive experiments. During the experiment, the subject will be hearing auditory cues (either a spoken word or an environmental sound). After the auditory cue, an image will appear in the center of the screen. The task then will be to determine for each picture whether or not the sound belongs to the same basic-level category as depicted in the image.
   3.2. We replicate experiment 1A of "What makes words special? Words as unmotivated cues" (Pierce Edmiston, Gary Lupyan 2015). The replication is part of a university project in order to acquire skills in conducting and reporting experiments.
   Furthermore, the general idea of replicating this study is to ensure that the results of the experiment are reliable and credible.

4. Hypotheses
   4.1. We hypothesize that environmental sounds as auditory cues cause longer reaction times than spoken category labels when trying to assess and match their basic-level category to the image that is being displayed. We expect that environmental sounds activate a more specific category instance linking the sound to its likely source, although participants are instructed to only treat it as its basic-level category. This would mean environmental sounds are less effective cues than spoken category labels.
   4.2. We also hypothesize that congruent environmental sounds exhibit a faster reaction time than incongruent ones because there might be less mental effort involved in assigning congruent sounds to their corresponding basic-level category (For a detailed explanation on congruency see chapter 14).

Design Plan

5. Study type
   5.1. Confirmatory Experiment - We are conducting an online experiment with a randomized order of trials.

6. Blinding
   6.1. The participants aren't aware of the experimental manipulations, as well as the general purpose of the study.

7. Study design
   7.1. The experiment uses a 2 (cue type: label, sound) x 2 (match: same, different) x 2 (congruency: congruent, incongruent) factorial within-subjects design. Where *same* refers to trials in which auditory cue and image match with respect to their basic-level category and *different* refers to trials in which they do not. Congruency is a variable that only applies to matching trials, where *congruent* describes all trials in which the instance shown in the image matches the likely source of the environmental sound that was played. For *incongruent* trials, the opposite is the case. Each trial starts with a 250 ms fixation cross that is displayed in the center of the screen, followed by an auditory cue. After a 1000 ms delay, an image is displayed in the center of the screen and stays visible until a decision is made by the participant. Participants should respond with a forced binary choice (Yes, No) indicating their decision via a keypress on their keyboard (q = Yes, p = No). Participants are advised to answer as accurately as possible while trying to minimize the reaction time.
   The experiment consists of 6 practice trials and 144 main trials. We decided to reduce the number of main trials compared to the original study (384), in order to obtain enough participants, since we cannot offer any kind of compensation for our subjects. Furthermore, participants from our pilot study, that had the same amount of trials, reported that the number of trials should not be any larger, as this would lead to increased boredom and exhaustion. This in turn would have led to a bias in reaction times. In the practice trials, the participants receive feedback via a popup window telling them whether their given answer was correct or not. In case of an incorrect answer, the popup window also gives a brief explanation why. We purposely excluded the feedback function during the main trials to prevent a possible training effect that might make certain differences in reaction times less visible.
   We use 24 color photograph images that can be divided into 6 basic-level categories: *bird, car, dog, instrument, phone and typing*. Each basic-level category comprises 4 images where two of them represent a different instance of the respective category (e.g. two images of a *Chihuahua* and two images of a *Rottweiler* for the basic-level category "dog"). For each category, we have one congruent and one incongruent environmental sound[1] and one spoken word for each basic-level category ('*bird', 'car', 'dog', 'instrument', 'phone', 'typing')* as an auditory cue.

---

[1] e.g. <Chihuahua bark> congruent for the Chihuahua images, incongruent for the Rottweiler images
<Rottweiler bark> congruent for the Rottweiler images, incongruent for the Chihuahua images
each representing the likely sound source for their respective instance of the basic-level category "dog"

The peak amplitude of all sounds, spoken words and environmental sounds, is normalized to -10db and the duration of each environmental sound is set to 1s, using the free audio editing tool audacity.
To control for cue variability we use a female and male synthetic voice for each category name. This way we level out any potential influence that is caused by the higher or lower pitch of the respective voice.

8. Randomization
   8.1. We randomize the order of the trials while keeping the amount of match and mismatch trials, as well as environmental sounds and labels even.

Sampling Plan

9. Existing data
   9.1. Registration prior to any human observation of the data: As of the date of submission, the data exist but have not yet been quantified, constructed, observed, or reported by anyone - including individuals that are not associated with the proposed study.

10. Data collection procedures
   10.1. Participants will mainly consist of family and friends, as well as fellow students. We can't offer any compensation because this study is only performed for educational purposes at our university. The recruiting time will take 7 days since we are on a tight schedule to meet the deadline of the course. We require from our participants that they have a computer or laptop with speakers or headphones and a stable internet connection. We collect our data using _magpie software and netlify as a hosting platform.

11. Sample size
   11.1. A study with a sample size that is too small may produce inconclusive results because the amount of data that has been collected is insufficient to perform reliable inference. Therefore, we want to acquire a target sample size of at least 43 participants, the same amount that was recruited in the original study.

12. Stopping rule
   12.1. We are going to stop the data collection process after 7 days (24.07.2020). In this time period, we want at least as many participants as our sample size suggests. Receiving more data than this would not pose a problem.

Variables

13. Manipulated variables
    13.1. We are manipulating the cue type of each trial, the congruency and whether or not the auditory cue matched with the shown image.
    Each manipulated variable has two levels: cue type (label vs. sound), congruency (congruent vs. incongruent) and match (match vs. mismatch).
14. Measured variables
    14.1. **RT**: The main outcome variable will be the reaction time of each participant that indicates for each trial how fast a participant gave a response. Other variables are:
    14.2. **congruency**: a variable that only applies to matching trials, where *congruent* describes all trials in which the instance shown in the image matches the likely source of the environmental sound that was played. For *incongruent* trials, the opposite is the case.
    14.3. **match**: whether or not the image category matched the sound category
    14.4. **subject_id**: unique anonymous identifier for each participant
    14.5. **cue type**: auditory cue or spoken label
    14.6. **correctness**: whether or not the participant gave the correct answer
    14.7. **image/sound category**: basic-level category of the respective image/sound
    14.8. **image/sound instance**: an instance of the respective image/sound

Analysis Plan

15. Statistical models
    15.1. We will fit two Bayesian generalized linear models using the brms package to infer our results. The manipulated, categorical independent variables are the sound type and congruence, whereas the dependent variable is Reaction Time (RT).
        15.1.1. The first model will predict RT based on the cue type including random by subject intercepts and slopes, random by category intercepts and slopes, as well as random by instance intercepts (Random Effects).
        15.1.2. The second model will predict RT based on congruency including random by subject intercepts and slopes, random by category intercepts and slopes, as well as random by instance intercepts.

16. Transformations
    16.1. In the data preprocessing step we will filter all incorrect and mismatching trials out for the first model. For the second model, we will additionally filter out all label trials.
    16.2. Furthermore, we will subtract 2250ms from the RT of each trial, due to a custom delay that we added in magpie and the duration of the displaying time of the fixation cross.
    16.3. Measurements such as reaction times are bound to zero (you can't be quicker than 0 ms), i.e. measures cannot vary a lot toward the left tail of its distribution. There is much more variation toward the right tail. We will compute the natural logarithm of our reaction time values as they are normally distributed after having been log-transformed.

For each model, the respective categorical predictor (sound_type or congruency) will be dummy coded with label / congruent being the respective reference level.

17. Inference criteria

17.1. We will make use of Credible Interval based inference in order to find out if we obtained significant differences in RT for the respective levels of the predictor. Our results are significant if the CrI's of the respective slope coefficients do not include 0. Moreover, we are going to evaluate the random effects we included by looking at their CrI's to judge whether or not our models lend credence to the random effects having played a significant role in the data generating process. Furthermore, we are going to check if the complex model structure (using random effects) of each model is justified using the LOO-Package by comparing them to simpler models (leaving out the random effects one by one).

18. Data exclusion

18.1. We exclude every participant that responds slower than 1,5s after the picture was displayed or faster than 250ms. We found these to be fitting measures to exclude accidental key presses and participants that are not focused on the task or being distracted.

19. Missing data

19.1. In principle it should not be possible that we receive submissions with missing data, however, we do commit to excluding all the trials of the respective participant if this problem should arise.

20. Exploratory analysis

20.1. We want to explore whether the assumed effect from hypothesis one is still present if we exclude all incongruent trials or whether the hypothesized difference in reaction times is solely caused by them.

20.2. Furthermore, we want to explore whether the assumed effect from hypothesis one is also present when only considering mismatching trials.

Other

21. References

    21.1. We replicate experiment 1A of "What makes words special? Words as unmotivated cues" (Pierce Edmiston, Gary Lupyan 2015).
    21.2. _magpie: https://magpie-ea.github.io/magpie-site/
    21.3. Netlify: An all-in-one platform for automating modern web projects. https://www.netlify.com/
    21.4. brms: Paul-Christian Bürkner (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software, 80(1), 1-28. doi:10.18637/jss.v080.i01 & Paul-Christian Bürkner (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. The R Journal, 10(1), 395-411. doi:10.32614/RJ-2018-017
    21.5. loo: Vehtari A, Gabry J, Magnusson M, Yao Y, Gelman A (2019). "loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models." R package version 2.2.0, <URL: https://mc-stan.org/loo>.
    21.6. Rstudio: R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
    21.7. tidyverse: Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686
    21.8. Audacity: Audacity® software is copyright © 1999-2020 Audacity Team. The name Audacity® is a registered trademark of Dominic Mazzoni.
    21.9. Sounds (all sounds were normalized and cut to 1s in length):
        21.9.1. CC: "ringtone.wav" by davidferoli
        21.9.2. CC: "5_cello_E2.wav" by flcellogrl
        21.9.3. CC: "Aggressive Guard Dogs" by Oneirophile
        21.9.4. CC: Bird Whistling, Robin, Single, 13.wav by InspectorJ
    21.10. Synthetic voices: https://www.text2voice.org/
    21.11. All the images that were we used stem from: https://www.pexels.com/