Study Information

Title: Perception and Identification of Randomness

Authors: Linus kleine Kruthaup, Cosima Oprotkowitz, Mara Rehmer, Katharina Trant

Description: Perception of randomness, the ability to perceive and discriminate structured versus unstructured events, is an important ability for survival and involved in many day-to-day activities and thus also important to study.

In this replication study, discrimination and identification of random versus non-random stimuli are tested between-participants in two experimental groups. For an exploratory analysis, we are looking into a possible effect of language on the identification of randomness, as the experiment is conducted in German and in English.

Hypothesis: "The probability of correctly identifying stimuli from R [random sources] and N [non-random sources] coincides with the ease of distinguishing between the two sources." (Zhao & Hahn, 2014)

Design Plan

Study type: Experiment

Blinding: The participants are not informed that there are two different experimental groups and each participant is only assigned to one group. The experiment will be conducted via the internet, so no direct contact between experimenters and participants will take place.

Study design: The experiment is a 2x51 mixed-factorial design. The first factor is *condition*, with the two values *discrimination* and *identification*. The second factor is *switch rate*, which has 51 levels. Each participant contributes data points for only one *condition* (between-participants), but multiple data points (ideally 6, for some 10, see below) for *switch rate* (within-subject, repeated measures).

Before the main part of the experiment, we added ten practice trials (also with *random switch rates*) to let the participants get used to the task. The original paper does not state whether they conducted practice trials, but after first trials, we received feedback that the participants would have wished for practice trials.

The overall design is the same as in the replicated experiment, but we made some changes to the colors of the stimuli materials and added practice trials as well as the instructions in

German. We additionally lowered the number of stimuli presentation per *switch rate* from 10 to 6, because we received negative feedback after first trials, that the participants had difficulties concentrating during the long version of the experiment. The data from the participants who had 10 trials per switch rate are nevertheless included in the final analysis, as we did not get feedback that the task was not understood correctly.

A more detailed description can be found in the corresponding Experimental Design plan.

Randomization: The participants are randomly assigned to either the *discrimination* or the *identification* condition by a coin-flip generator.

Sampling Plan

Existing data: As of the date of the submission of this preregistration, we have collected and analyzed data from a small pilot study and collected data from the main experiment, due to feedback-induced changes in the design. (Four data sets of the latter will be included in the final analysis, even though these participants did not have practice trials and conducted the experiment with ten data points per *switch rate,* as we received feedback that these participants did not have difficulties in generally understanding the task.)

Data collection procedures: Participants will be drafted through social media and direct messages (e-mails and text messages). Participation is voluntary and will not be compensated. After sending out the invitations on the 1st of August, we will wait 12 days until closing the data collection. Participants are only eligible if they are at least 18 years old and, although this may seem redundant, if they have full or corrected vision. A participant is only allowed to participate once.

Sample size: We will try to recruit as many participants as possible.

Sample size rationale: Since time is critical due to a deadline, our pool of reachable participants is limited and we do not offer any compensation for participation, we cannot state a minimum number of participants.

Stopping rule: We stop data collection on 23:59 of the 17th of August, which is 17 days after sending out the invitations, due to the project deadline. (This was originally planned for 12 days, but due to technical problems and feedback, we extended the data collection period.)

Variables

Manipulated variables: Firstly, we manipulate the condition, as some participants are assigned to one and other participants to the other condition. We then manipulate the *switch rate*, that is the density of randomness, of the given stimuli picture. The *switch rate* has 51 different levels, each of which is shown ten times per participant. This *switch rate* lies between 0 and 1. The closer to 0.5 the *switch rate*, the more does the color assignment happen like one would expect from a random source. The manipulated variable *switch rate* is the same for *discrimination* and *identification*.
More on this in the 'Experimental design'.

Measured variables and indices: We will measure the reaction time at each task (for data exclusion, see below) and we will measure whether the tasks were answered correctly or not. Then we will calculate the average accuracy, that is the proportion of correctly answered tasks, at every *switch rate* for the *discrimination* and the *identification* condition. We will measure this by first calculating the average accuracy at every *switch rate* for each participant and then calculate grand means by averaging across the participants in the respective condition.

Analysis Plan

Statistical models: We will have a Bayesian logistic regression model, where *correctness* is the dependent variable and *condition* (group) and *switch rate* are independent variables. We will conduct the analysis using the programming language R and the ´brms´ package. We will use weakly informative priors, namely a student_t(1, 0, 2) for all class b parameters and the formula we will use for our model is *correctness ~ condition * switch rate + (1 | submission_id).* The script "Pilot-Analysis" from our pilot contains our analysis as planned.

Inference criteria: We will use 95% credible intervals (CI) from the posterior estimates. We will only be confident in the estimates if their 95% CI does not include zero.

Data exclusion: We will exclude data points with reaction times slower than 8000ms, as we cannot guarantee that the answer is related to the respective stimulus when that much time has passed. We will not exclude data points with fast reaction times, as, before a participant can answer, there is already a delay of 1500ms.
The practice trials will not be included in the analysis.

Missing data: Should a data set not be recorded completely and data points be missing, we will just calculate the respective participant's accuracy with the remaining values as described above (`Measured variables and indices´).

Exploratory analysis: We further plan to look for relationships between the language the participants stated as their main language and the results of the *identification* part. (This of course only applies to those in the *identification* group.) We only do this if we have sufficiently enough participants for both languages.

Other

This study is a replication of Experiment 1 in "Perception and identification of random events" by Zhao and Hahn (2014) (https://doi.org/10.1037/a0036816). Changes we made to their experimental design and analysis are either stated here or in our 'Experimental design'.