Preregistration report

Study information

Title:

The effect of atypical exemplars on the manual response in a categorization task - a replication study

Authors:

Luisa Drescher (ldrescher), Hannah Gauditz (hgauditz), Lilian Menzner (lmenzner),

Lilith Okonnek (lokonnek)

Description:

As final project for our course ‘Experimental Psychology Lab’, offered by Prof. Dr. Michael Franke at the University of Osnabrueck in the summer term of 2020, we will perform a replication attempt and conduct our own experiment. Our study is based upon parts of the study ‘Graded motor responses in the time course of categorizing atypical exemplars’ by Dale et al., published in ‘Memory & Cognition’ in 2007. The focus of our study lies on the first experiment where animal stimuli are presented in lexical form on a screen and the participants select the fitting category for the animals by moving their mouse cursor to the appropriate corner. Typical and atypical representatives of the categories will be presented with their correct category and a competing one. During the classification many parameters will be measured, most important the correctness of the classification and the curvature of the path the cursor is following. The study uses graded motor response to analyse the cognitive processes involved in categorizing typical and atypical exemplars.

Hypothesis:

Before inspecting any data we assume the following hypotheses.

We believe the participants' mouse trajectories gravitate to the alternative (incorrect) category more when presented with an atypical exemplar than when presented with a typical instance. Furthermore, we expect more typical category members to be recognized on average more quickly than atypical exemplars. Regarding the correctness of categorization, we anticipate the category members which are more typical instances to be recognized more accurate than atypical exemplars. Tests on the reaction time and correctness of categorization will be analyzed using a two sample t-test and we will assume zero difference in both typicality categories as our null hypotheses.

When facing atypical exemplars, the participants' mouse movements should show evidence of competition between the two categories, most noticeable in a bias toward the competing category. This hypothesis will be tested by using a two sample t-test while assuming the difference of the *x*-coordinates of typical and atypical exemplars being zero as null hypothesis. A conservative t-test is used here, where divergence is defined as consecutive statistically significant divergences. The number of divergences is established by bootstrapping. In addition to the aforementioned t-test, we will be using repeated measures ANOVA. Bimodality of AUC will be examined in order to see whether it explains the divergence of atypical trajectories.

We theorize to find results that resemble the results that were obtained in the the original study by Dale et al.

Design Plan

Study type:

In this web based replication experiment, each participant will face control and experimental trials in a random order. They will either be confronted with a typical instance (control condition) or an atypical instance (experimental condition).

Blinding:

All participants will be presented with both control and experimental trials without having any knowledge about the categorization of the presented case. The researchers have no direct contact with the partakers apart from a uniform recruitment message, as the experiment will be constructed via internet.

Study design:

In this within-subject experiment, the one factor design with two levels is defined by the two clusters the presented stimuli might belong to. The displayed animals count either as typical or atypical representatives of their categories. While the goldfish (typical instance) for example belongs to fish, which mirrors the observable features of the animal, the bat (atypical instance) showcases characteristic features of birds while belonging to mammals. The participants have the task to select one of two animal categories the representatives belong to. The categories for atypical exemplars are in a competing manner, the choice for the bat for example is between a bird and a mammal while the categories for typical exemplars like the goldfish are not in competition . Follow the link below to find a more detailed description of the study in the file ‘experimental design’.

Randomization:

All 19 prepared items, six atypical to their category and 13 typical, will be presented to the participants in a randomized fashion.

Sampling Plan

Existing Data:

A pilot study was run to gather data which will not be included in the final analysis. This was done in order to test for errors in the implementation of the experiment and to obtain an overview of the data that could be expected in the final experiment.

Some mistakes were found after the initial phase of the test run, causing most of the data to be unserviceable due to one missing parameter. A sample of datasets (n=3) was usable for the construction of the statistical analysis script.

Data collection procedure:

The participants will be contacted via email and social media and receive a uniform message by the researchers. It will include an online link and requirements to participate. The sample will include multiple ages, nationalities and education levels. Participation is voluntary and will not be compensated. To take part in the study, the partakers must be able to understand English and execute the experiment on a device with mouse or a mousepad, touchscreens are inadmissible for data collection.

The online link will be active for five days, participation is possible in the timespan of 20th of July until 24th of July.

Sampling size:

For the study there will as many people as possible recruited.

Stopping rule:

Due to the limited time of this project, the data collection will be stopped on the 24th of July at midnight.

Variables

Manipulated Variables:

The variable that the researchers are able to manipulate is the category the animal stimulus belongs to. Either the participants face a typical representative of a category, here the dominant features agree with the correct class, or an atypical instance. In that case, the animal shows features of the competing category which poses an alternative to the correct class in the decision.

Measured Variables:

We will measure the correctness of the categorization (binary), the overall time spent per trial between stimulus onset and selection of a category, the time between stimulus onset and first movement, the movement duration, the distance traveled by the cursor (in pixels) and the trajectory of the cursor path. The last measure can be further divided. One part is presented by averaging of a full trajectory by normalizing it into timesteps and projecting it to a coordinate system where (0,0) marks the origin. Discrepancies in the cursor paths from a previously computed averaged tract is quantifiable by its x-coordinates in the set timesteps. The other aspect of the cursor path is measurable by a space normalized examination. In addition to the fixed starting point at (0,0), the endpoint is settled as (1,1). On the journey from (0,0) to (1,1), the real-time positions are clustered into time intervals for a better analysis of the path.

Indices:

No indices used.

Analysis Plan

Statistical models:

All analysis is done on the correct responses. Incorrect responses are removed. For the analysis of the time-normalized data, two-sample t-tests for the difference between *x*-coordinate of each of the 101 time-steps to zero between the atypical and typical trials are conducted. To provide more statistically powerful results, the right- and left-ward responses are pooled. In addition, a more conservative test is performed. A *reliable divergence* as some *d* consecutive time-steps with significant *p*-values is established by performing a variation of bootstrapping of the 101 time steps. 10,000 simulations are performed with the same mean and standard deviation of *x* at each timestep for atypical and typical trials, simulated for each participant. The frequencies of significantly different consecutive time steps are recorded. A *d* s.t. $p < 0.01$ (cf. the probability that this number of consecutive timesteps occurs randomly) is picked in order to discern whether the divergence occurred by chance or not. Once the sequence size is fixed, the t-tests are applied on each timestep between atypical and typical trials. If *d* consecutive timesteps are found, the trajectories are considered divergent. Here, *reliable divergence* is defined as at least seven consecutive time-steps that have significant *p*-values with significance at 0.05. Afterwards, pooled bins from the time-normalised data are computed to conduct a 2x3 repeated measures ANOVA (viz. typical and atypical by the bins 1-33, 34-67, 68-101). A planned comparison between trial types is then conducted within each bin to reveal portions of the data exhibiting divergence. For the space-normalized data, a 2x3 (type by space-normalized bins) repeated measures ANOVA and planned comparisons are performed as described above.

A comparison between the accuracy of atypical and typical trials is made by using a two-sample t-test against the null hypothesis of the difference being 0. Accuracy of atypical and typical trials here indicates how often participants answered correctly Further, the performance of two-sample t-tests between atypical and typical trial types is conducted for movement duration, total categorization time,

total distance traveled and movement initiation latency. Item-based repeated measures ANOVA on these same measures is examined with exclusion of the accuracy measure.

The nature of curvature across trajectories is examined. In particular, the focus is on exploring its bimodality. If AUC exhibited bimodality, the curvature of trajectories in the atypical category towards the competitor can be explained by two distinct types of behaviour, namely a movement towards the competitor followed by a course-correction, or a relatively straight path towards the answer. The z-scores for all areas are computed and subjected to the distributional analysis. A Kolmogorov-Smirnov test is performed on the difference in distribution. Bimodality coefficients are computed for each trial type. It was investigated if the coefficients were within the unimodal range of $b < 0.555$.

Transformations:

Analysis of the data is conducted in two ways. First, all trajectories are normalized into 101 discrete time-steps, translated s.t. they all start at (0, 0). This is called the time-normalized dataset. Second, the real-time information of the trajectories is recorded by the x,y coordinates as the mouse cursor travels from 0 to 1, more precise from. the origin (0,0) to the final answer (1,1). Additionally, the mouse movement initiation time is calculated as well as the movement duration, the total categorization response time, and the distance traveled in pixels. Further, the area under the curve, viz. the area in pixels between the trajectory line and a straight line between the start and finish, is computed.

Inference criteria:

For our evaluation of significance we will be using the standard $p < 0.05$ criteria in our tests.

Data exclusion:

Participants who noted that they used a smartphone or tablet for the study will be excluded because no cursor movement can be collected from these devices. For our analysis only correctly categorized trials will be included.

If a participant denotes severe technical difficulties, that impact the measured variables, their dataset will be excluded from our analysis as well. Further, a Rosner Test is used to remove any outliers in reaction time (using the median with threshold of 3.5).

Missing Data:

In case a participant did not execute the entirety of the experiment, the collected data will still be used in the analysis.

Exploratory:

In addition to our replication of the experiment by Dale et al. we will also explore our dataset and examine if there occurs a difference in performance between left- and right-handed participants. Although the partakers are instructed to use their right hand the participants will be asked to report their ordinary dominant hand in the post trial query. We examine the effect of the participants dominant hand on both the proportion of correct responses and the total reaction time. Two-sample t-tests will be executed to test for an effect. As our nullhypothese we will assume a difference in reaction time and correctness.

Other

Materials: Link to word stimuli (table)

Statistical analysis script: Link to analysis (Rmd file)

Experimental design: Link to experimental design (pdf file)

Data of sample study: Link to results from pilot study (csv file)