Preregistration

Preregistration: Transitive inference in children

Petra Borovská[1], Jacob Bellmund[2], Artem Karew[2]

[1] University of Leipzig

[2] Max Planck Institute for Human Cognitive and Brain Sciences

16. July 2020

Study Information

Title	Influence of Sleep on Transitive Inference in Children Preregistration: Transitive inference in children
Description	The main aim of this study is to assess an influence of sleep or longer resting period on transtive inference ability in children in age 8-9 year old. The transitive inference task is an ability to generalized from existing knowledge forming a novel combinations and thus providing an insight to basic cognitive skills, children possess (Zeithamova, 2012; Ellenbogen, 2007). Since couple of decades an open discussion about when exactly are children capable of passing the transitive inference task prevails (Bryant & Trabasso, 1971) and this question remains also valid for adults, whose performance on this task is not above chance immediately after training compared to participants that where tested after a delay (Ellenbogen, 2007). This

effect is related to internal brain processes involving a memory integration mechanism triggered by hippocampal-medial prefrontal circuit. As a result of hippocampal replay during sleep or longer delay, it is hypothesized that by recombination of previously learned knowledge a novel connection, i.e. inference, occur. We expect to see similar effect in children of age 8 and 9, assuming that by this age, they are already making inference similarly to adults. However, it is also possible that they will not pass the task, thus revealing that the processes are not yet fully developed. Similarly it is still open which kind of strategy children or even adults using when making the inference. It is speculated that the judgement on more distant pairs are done faster, when the whole hierarchy is encoded as single representation, opposite to retrieval based mechanism, since more distant pairs recquire more time to retrieve more complex information in comparison to less distant pairs (Zeithamova, 2012). It remains unclear if children generally use different strategy or the strategy is affected by influence of the sleep or delay. The goal is to behaviorally assess the two strategy approach by using response time measure.

Hypotheses

1. Performance on inference pairs will be significantly higher in delay condition vs. non-delay.
2. Performance on 2 degree pairs increase after delay vs. non-delay condition.
3. Response time decreases after delay vs. non-delay condition.
4. Response time is higher in more distant pairs vs. less distant pairs.
5. Premise pairs correlate with inference pairs.

Design Plan

Study type **Experiment**. A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.

Blinding For studies that involve human subjects, they will not know the treatment group to which they have been assigned.

Study design We have a mixed design with 2 x 2 factors of one being between (delay by non-delay) and one within-subject (premise pair by inference pair) with peformance score as dependent measure. Additionally we use similar design for contrasting one degree inference pairs and two degree inference pairs with assumption that premise pair performance will be similar in both groups. Further we use the same model for second dependent measure, response times. Lastly, we look at linear regression model of inference pairs as function of middle premise pair performance and possibly conduct an exploratory analysis on detailed type of pairs in case the model will be significant.

The study is divided into training and testing, both organized into blocks. Training contain 5 mandatory blocks with 10 trials each, followed by futher blocks where score is counted on middle pairs (B > C, C > D, D > E) to assess appropriatelly if the hiearachy was properly learned, as the first and last pairs are easily learned, because A is always correct and F is always incorrect. If the participant pass the block above 75% correct in two consecutive blocks, the training is over. And testing is followed depending on the group assignment. In the delay condition, test is submitted after 24 hours. In the non-delay condition the test follows after 10 minutes break from training. The test includes 5 blocks, each consisting of 18 trials.

Each stimulus is a representation of an item in the hierarchy and each participant receive his own order of the hierarchy, which is individual and unique, meaning that 2 participants cannot have the same hierarchy. So for example *A* is represented by *monster1* and *B* is represented by *monster2*. Correct response is then defined by the default hierarchy order, when A is always higher than B, B than C, etc. The hierarchy is 6-item long, from A to F.

The study is conducted on computer in online manner. The stimuli are presented in random order, such that none of the items of each pair shown on the screen are followed by the same item of the next stimuli pair (e.g. when A > B presented, nor A nor B will follow on the next trial). The side of the stimuli pair is counterbalanced, such that each stimuli pair is presented twice in each block, the higher item of the pair always on the left or on the right side to prevent any side preferences.

Randomization

We will use simple randomization, each participant will be randomly assigned to the same-sized group - delay or non-delay, prior recruiting. Randomization will be done in R, using *sample()* function. For the details please see R markdown document attached.

Sampling Plan

For the data collection, database of Max Planck Institute for Human Cognitive and Brain Sciences will be used for recruitment of participants. We plan to collect sample of size 40, assuming that the recruitment sample will be in reality higher due to dropouts, etc.

Existing data

Registration prior to creation of data. As of the date of submission of this research plan for preregistration, the data have not yet been collected, created, or realized.

Note: the data used for analysis has not been collected, however adult pilot has been done and used for analysis preparation.

Explanation of existing data

Apart from pilot data, that won't be included in analysis, no further existing data will be used.

Data collection procedures

Participants will be recruited from the Max Planck database as mentioned above and paid a voucher in amount of 10 Euro. Target group are children of age 8 and 9 without any serious medical history. The consent will be collected prior of testing in form of online confirmation from parents and children. Participants who does not finish the whole procedure will be excluded. Similarly participants that exceed 30 training blocks will be excluded.

Sample size

Our target is 40 children, 20 in each group, however we will recruite at least 50 assuming that not all complete the task.

Sample size rationale

In order to estimate recquired sample size, power analysis using G*Power program was used. Calculation was based on main previous research paper from Ellenbogen (2007). Because not all descriptive data were provided in the paper, calculation was based on F statistics information, following a study by Lakens (2013), when calculating the effect size.

Parameters:

F tests: ANOVA: Repeated measures, within-between interaction

A priori: Compute required sample size

Effect size: 0.37

Alpha err prob: 0.05

Power: 0.80

Number of groups: 2

Number of measurements: 2

Corr among rep measures: 0

Nonsphericity correction: 1

Resulting total sample size was calculated on 32.

Additionally R power analysis was used to verify the result. The recquired sample size based on R power analysis slightly exceeds 40 participants given the power of 0.80.

Please see attached R script.

Stopping rule

The data collection will be stopped when 40 participants complete the procedure.

Variables

Dependent measures.

Performance: percent correct, if selected picture will be higher in defined hierarchy, e.g. A ? B, the correct response is A.

Response times: calculated from stimulus presentation until key response given.

Independent measures:

Group: delay and non-delay

Delay = testing is conducted after 24 hours from training.

Non-delay = testing is conducted after 10 minutes break from testing.

Pair Type: premise and inference

Premise = all pairs not involving inference judgments: A>B, B>C, C>D, D>E, E>F, the total score is 50 in all blocks of testing.

Inference = all pairs involving inference judgments and not involving first and last pair: B>D, C>E, B>E, the total score is 30

One Degree Pair = all pairs with the distance of one member apart: B>D, C>E, the total score is 20 in all blocks of testing.

Two Degree Pair = all pairs with the distance of two member apart: B>E, the total score is 10 in all blocks of testing.

Manipulated variables

We manipulated in which group participants are assign to, delay vs. non-delay and which pair type was presented: premise, inference, one degree, two degree.

Measured variables

We will use two types of dependent measures, percent correct, calculated from correct responses 1 or 0 summed across all trials and blocks and expressed in proportions and response times calculated from stimulus presentation until a key response (see above).

Indices

We will calculate sum of correct responses in each pair category that will be expressed in proportion and averaged when collapsing a category, e.g. one degree and two degree are firstly calculated separately as proportions and then averaged when including as one category - iference pairs.

Analysis Plan

We plan to conduct mixed ANOVA and linear regression. Firstly, the data sets will be prepared to for appropirate models, e.g. long format of the data for ANOVA. We will conduct an assumption checks for both types of test. We will provide descriptive statistics including mean, median, sd, standar errors and confidence intervals. Similarly effect sizes will be calculated.

Statistical models

We plan to conduct mixed ANOVA, involving at least 3 models.

1. With performance as outcome variable and with one between factor *group* with two levels - delay and non-delay and one within-subject factor *pair type* with two levels - premise and inference.
2. With performance as outcome variable and with one between factor *group* with two levels - delay and non-delay and one within-subject factor *pair type* with two levels - one degree and second degree inference.
3. Model with response time as outcome variable and with one between factor *group* with two levels - delay and non-delay and one within-subject factor *pair type* with two levels - one degree and second degree inference.

Transformations

No data transformation will be conducted with performance measure, but it's still remain open how to tackle response times.

Inference criteria

We will use the standard $p<.05$ criteria for determining if the ANOVA and the post hoc test suggest that the results are significantly different from those expected if the null hypothesis were correct. We expect to see significant interaction between group and pair type, where inference will have significantly higher performance in delay condition than in non-delay condition, where we expect performance on the chance level. Further we will provide a standard error as a measure of how well our sample represents the population. Effect sizes for performance between groups will be provided in proportions. Effect sizes for response times will be reported with confidence intervals.

Data exclusion

We will assess the outliers by measuring a cook distance and will expore an individual cases. If the outliers will represent general population deviation and won't be considered as errors, e.g. missing responses, they will be included in the analysis.

Missing data

If the subjects don't complete training or testing, they will be excluded from analysis. If participants fail to pass the training until reaching 75% criterion within 30 blocks, they will be also excluded.

Exploratory analyses (optional)	Enter your response here.

Other

Other (Optional)	Enter your response here.

References