



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY

H Y D E R A B A D

BRSM Project

Multifaceted information-seeking motives in children

Date: 07/05/2024

by

Manender Yadav (2023701020)
manender.yadav@research.iiit.ac.in

Anmol Vashishtha (2023201079)
anmol.vashishtha@students.iiit.ac.in

Under the supervision of

Dr. Vinoo Alluri (Assistant Professor)
Dr. Vishnu Sreekumar (Assistant Professor)
Teaching Assistants

Contents

1	Scope and Intention for Topic Selection	1
2	Introduction	1
3	New Proposed Hypothesis	2
4	Dataset and Methods	2
5	Results and Analysis	5
6	Discussion and conclusion	12

1 Scope and Intention for Topic Selection

In this project centered on understanding and implementing various statistical techniques learned throughout the course, we will try to first replicate a few experiments. After thoroughly understanding the original study, we plan to propose some new hypothesis based on the study. Then we will be testing those hypothesis using the statistical tests that we have learned. Through experimentation with practical implementation of statistical techniques, this project aims to advance our understanding of Statistical Methods in Behavioural Research.

This study, '**Multifaceted information-seeking motives in children (Molinaro et al.)**', explores information-seeking behavior in children (preschool to early adolescence). It not only talks about the motives driving information-seeking behavior but also has a methodologically rigorous approach to check these behaviors. Replicating this study will help us to focus on behavioral research methodologies and statistical analysis. It provides a practical application of these concepts in developmental psychology.

2 Introduction

Infants seek information through gestures and vocalization and once they become able to speak, they seek more and more information. It serves multiple, sometimes competing goals. Hence, deciding which information to seek out can be a difficult task.

Let's consider a scenario involving a child who suspects their friend received more playing time during a game than they did. The child notices that during a sports game, their friend seemed to spend more time on the field than they did. This creates uncertainty about whether they were treated fairly in terms of playing time. Directly asking the friend, how much time they got to play will reduce their uncertainty about the fairness of the situation. However, the child may also be hesitant to directly confront their friend about the perceived injustice for fear of creating animosity or tension in their friendship. Additionally, the child may weigh whether they believe they can rectify the potential unfairness. If they trust their coach to address any discrepancies in playing time allocation, they may be more inclined to seek information from them rather than directly confronting their friend. In this scenario, the child must navigate between these different motives for seeking information: reducing uncertainty, avoiding conflict, and assessing their ability to address the perceived injustice. The decision on whether to inquire about playing time will likely depend on how much weight the child assigns to each of these motives and their perception of the situation's fairness.

A study has already proposed a theory which outlines three primary reasons responsible for seeking information[4]. According to this study, individuals first assess what the information might uncover, followed by estimating its anticipated effects on their emotions, thoughts and potential actions. They also consider how it might enhance their understanding of the world around them and to achieve their goals, avoiding negative consequences.

Several studies associate affective evaluation of information with activity in dopamine-rich mid-brain (ventral tegmental area and substantia nigra), whereas, the other two, instrumental utility of information and its ability to reduce uncertainty have been associated with activity in the frontal and prefrontal cortex [1, 2, 3]. Since, sub-cortical regions mature earlier than frontal and prefrontal regions of the brain, the authors hypothesized that the weights assigned to affect might become fully apparent earlier in the development, while weights related to the other two, instrumental utility and uncertainty reduction becomes visible slightly later in the child's development and may continue to increase throughout their childhood.

In this current paper, the study was conducted on young children and the three infor-

mation seeking methods were tested using some experiments. Reported results were re-implemented by us to some extent and we found matching results. However, this study briefly mentions that there were no effects on the results due to gender, but doesn't share any relevant analysis or results proving the same.

3 New Proposed Hypothesis

Based on the original study, we propose 4 new hypothesis, which we plan to test further in the project:

1. There is no difference in comprehension scores between male and female children. (**Manender yadav**)

2. There is no difference in comprehension scores between genders across different age groups of children. (**Manender yadav**)

Reasoning: Our hypothesis centers on the premise that children exhibit uniform task comprehension levels, irrespective of gender distinctions. Also, we tested this across the age groups to ensure uniformity for the further tests.

3. Individual differences in perceived competence influence children's propensity to seek information, especially under conditions of uncertainty. (**Manender yadav**)

Reasoning: After all the tests, we were sure that all the children had understood the task. We also wanted to see, if Children's propensity to seek information under uncertainty is influenced or not, by their perceived competence in the task.

4. There is no difference in the parameters namely uncertainty, agency and expected gain between genders. (**Anmol Vashishtha**)

Reasoning: Here, we wanted to verify if gender plays any role in the development of the three key parameters of information seeking. This could indirectly hint at the different brain region development based on gender.

4 Dataset and Methods

Over seven experiments involving 521 children (aged 4–12) were conducted in the original study. Data is collected and stored in a tabular format. A sensitivity analysis revealed that the final sample sizes can detect an effect size of Cohen's $\omega = 0.20$ on a χ^2 -test with the desired power and one degree of freedom. This suggests that the study was sufficiently powered to detect small to medium effects. However, very small yet significant effects may still go undetected.

Experiment 1 involved around 200 children in a 1:1 gender ratio. Dataset contains participant details along with variables such as comprehension task scores, decision-making choices

and reaction times. Comprehension task score were recorded by another set of experiments which were specifically designed for the stated purpose.

Experiment 2 is the same as experiment 1 and is used for replication analysis.

Experiment 3,4 and 5 are specially designed to make sure the three motives for information seeking are not affecting each other. i.e in each experiments the variables for the remaining two motives were kept constant.

Experiment 6 is similar to the first experiment, its just the participants were adults in this. And experiment 7 is its replication.

Hypothesis 1:

Null Hypothesis (H0): There is no significant difference in comprehension scores between males and females.

Alternative Hypothesis (H0): There is significant difference in comprehension scores between males and females.

Here data from experiment 1 was used as it had the comprehension scores of the children along with their age and gender. For task comprehension experiment (Fig.1), Participants answered one question at a time by selecting one of the buttons displayed above each option. If they answered correctly, they were presented with the next one. If not, they were prompted to try and answer the question again. Participants could attempt each question three times at most to avoid being excluded from the dataset.

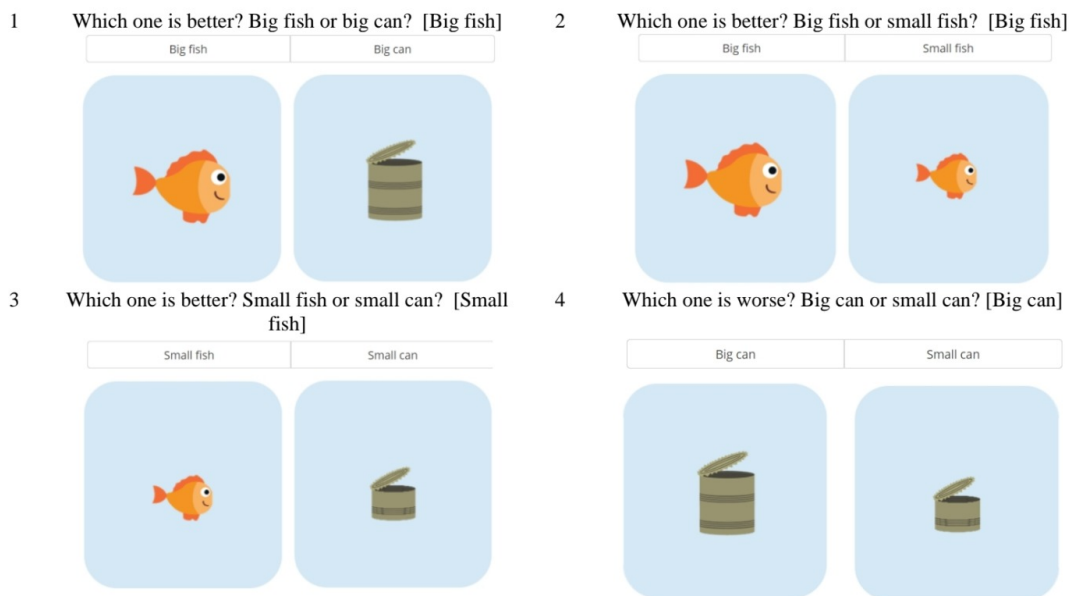


Figure 1: Sample questions for the comprehension task.

The data is divided in two groups, i.e. male and female. Descriptive statistics of the "comprehension score" variable were extracted from the data. Normality of the data is tested first visually by plotting a box-plot and Q-Q plot. Later Shapiro-Wilk test was employed. For comparing the scores of the two groups, we could go for independent samples T test or Mann-Whitney U test. Based on normality testing results, we used Mann-Whitney U test, it is a non-parametric test used to compare the distributions of two independent groups. It assesses whether the distributions of scores in the two groups are the same or if one group tends to have higher scores than the other. This test does not assume normality and is

robust to differences in variance between groups.

Hypothesis 2:

Null Hypothesis (H0): There is no difference in comprehension scores between genders across different age groups.

Alternate Hypothesis (H1): There is difference in comprehension scores between genders across different age groups.

Here we proposed this hypothesis to understand if there are any effects of gender in different age groups. For this, we again visualize the data to check for the normality. We plotted the box plots and Q-Q plots. We also performed Shapiro-Wilk test for normality. Followed by Mann-Whitney U test.

Hypothesis 3:

Null Hypothesis (H0): There is no difference based on gender in median comprehension scores between two groups.

Alternate Hypothesis (H1): There is a difference based on gender in median comprehension scores between the two groups.

Here, group 1 is children who did not check the other side pond with seaweed (less uncertainty reduction). And group 2 is children who did check the other side pond with seaweed (more uncertainty reduction).

After the comprehension task experiments, we wanted to check if Children's propensity to seek information under uncertainty is influenced or not, by their perceived competence in the task. The original study found that children's tendency to seek information under conditions of uncertainty increased with age, which might reflect not just cognitive development but also an increasing sense of competence and self-efficacy as they grow older. So we came up with a new hypothesis. It suggests that individual differences in perceived competence could influence their propensity to seek information, especially under conditions of uncertainty. Here, perceived competence is basically how capable children feel they are doing the tasks, This adds an internal psychological dimension to the information-seeking behaviors. Relevant experiment was selected which would give us the desired data. And then data of our use was extracted from the whole dataset.

While the original study provides a developmental overview of how, certain information-seeking behaviors emerge and intensify, the new hypothesis introduces a layer of psychological self-assessment (perceived competence). Which offers a different angle to analyze the existing data. This approach could provide insights into how self-perception influences cognitive behaviors.

Hypothesis 4:

Null Hypothesis (H0): There is no difference in the parameters namely uncertainty, agency and expected gain between genders.

Alternate Hypothesis (H1): There is difference in the parameters namely uncertainty, agency and expected gain between genders.

Based on the dataset we proposed this hypothesis to inquire if there were any significant differences in development of certain regions of the brain between male and females. To test this we replicate the methodology highlighted in the original paper. The model and

techniques used for this hypothesis are beyond the scope of this courses, however we felt that this particular analysis was instrumental testing the correlation between ΔEV (expected value), Δ agency, Δ uncertainty and age of children in years

Logistic Mixed-Effects Modeling:

- Logistic mixed-effects regression models were likely used to analyze the data and predict children's information-seeking choices on each trial.
- Factors such as ΔEV (expected value), Δ agency, and Δ uncertainty were included as fixed effects, along with interactions with age to assess developmental trajectories.
- Age in years was included as a fixed effect to examine how children's age influenced their information-seeking motives.
- Covariates such as instruction comprehension score, EV comparisons task score, and correct fishing choices were controlled for in the models.

Model Interpretation and Analysis:

- The models were used to determine the relative impact of different motives (EV, agency, uncertainty) on children's information-seeking choices.
- Interactions between age and each motive were examined to understand how the importance of these motives changed with age.
- Additional models may have been run to explore triple interactions between age and pairs of information-seeking motives to further analyze developmental patterns.

Johnson-Neyman technique:

- Johnson-Neyman technique was used to estimate the age at which the effects of uncertainty on confidence in information-seeking became significant.
- By applying this technique, we analyzed the specific age range at which uncertainty had a significant impact on children's confidence in their information-seeking decisions.

5 Results and Analysis

Hypothesis 1:

Both the groups have a median score as 1 (0-1) and minimum score as 0.62, which is sufficiently above the required threshold (Fig.2). We can infer that both the groups have comprehended the tasks fairly well. However, if there is any differences amongst the gender, is tested using the mentioned statistical tests (see sec. Dataset and Methods). The Q-Q plot visualization (Fig.3) here shows us that the data is not normal. If the data points fall close to a straight diagonal line, then the distribution of the comprehension scores is similar to a normal distribution. In this plot, the data points appear to deviate from a straight diagonal line. This suggests that the distribution of the comprehension scores is likely not normal.

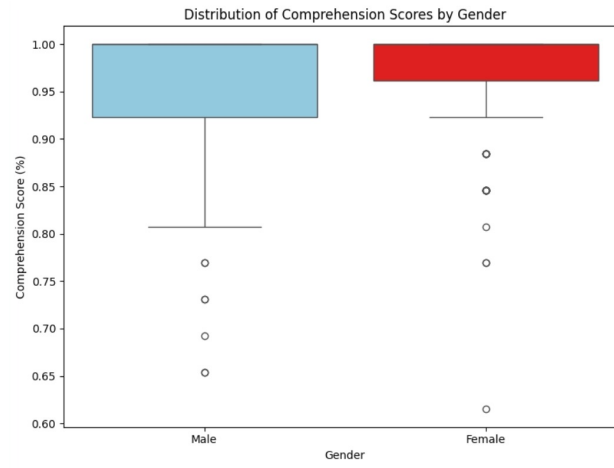


Figure 2: Box plot visualization of comprehension scores of male and female.

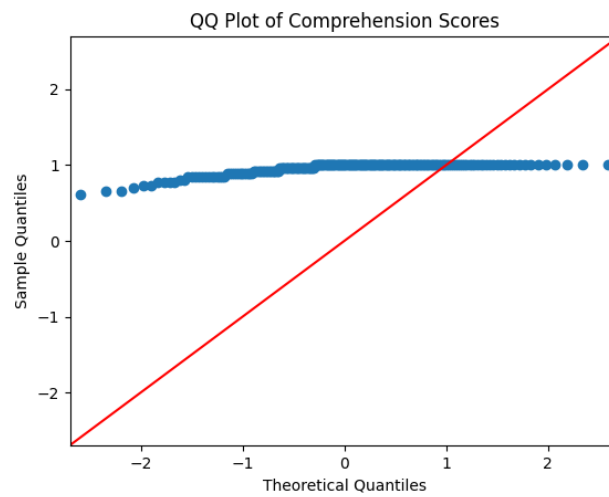


Figure 3: Q-Q plot visualization of comprehension scores.

Shapiro-Wilk test for male and female resulted in a p-value of $2.3677078093281567e-13$ and $1.701075430960309e-14$ respectively. Since the p values are below 0.05, we reject the assumption of normality. As the data is not normal, we can not apply independent T test, we go for the Mann-Whitney U test (non-parametric test). It assesses whether the distributions of scores in the two groups are the same or if one group tends to have higher scores than the other. The null hypothesis for this test is that there is no difference in scores between the two groups. As we can see, the P-value is 0.297, which is larger than 0.05, so we fail to reject the null hypothesis.

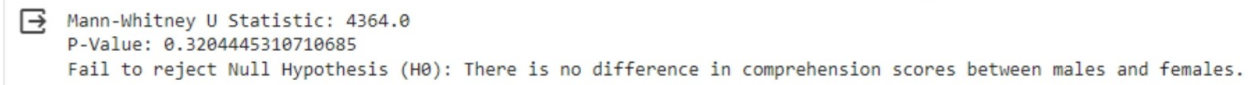


Figure 4: Mann-Whitney U test for Hypothesis 1

Hypothesis 2:

For this hypothesis also, we first visualized the data using a box plot with which we could get a rough idea about the median scores, and it can be said that the scores increase with age and in a group there is not any significant difference between male and female (Fig.5). After this, to test the normality, we visualized the data using Q-Q plot. And we can confirm

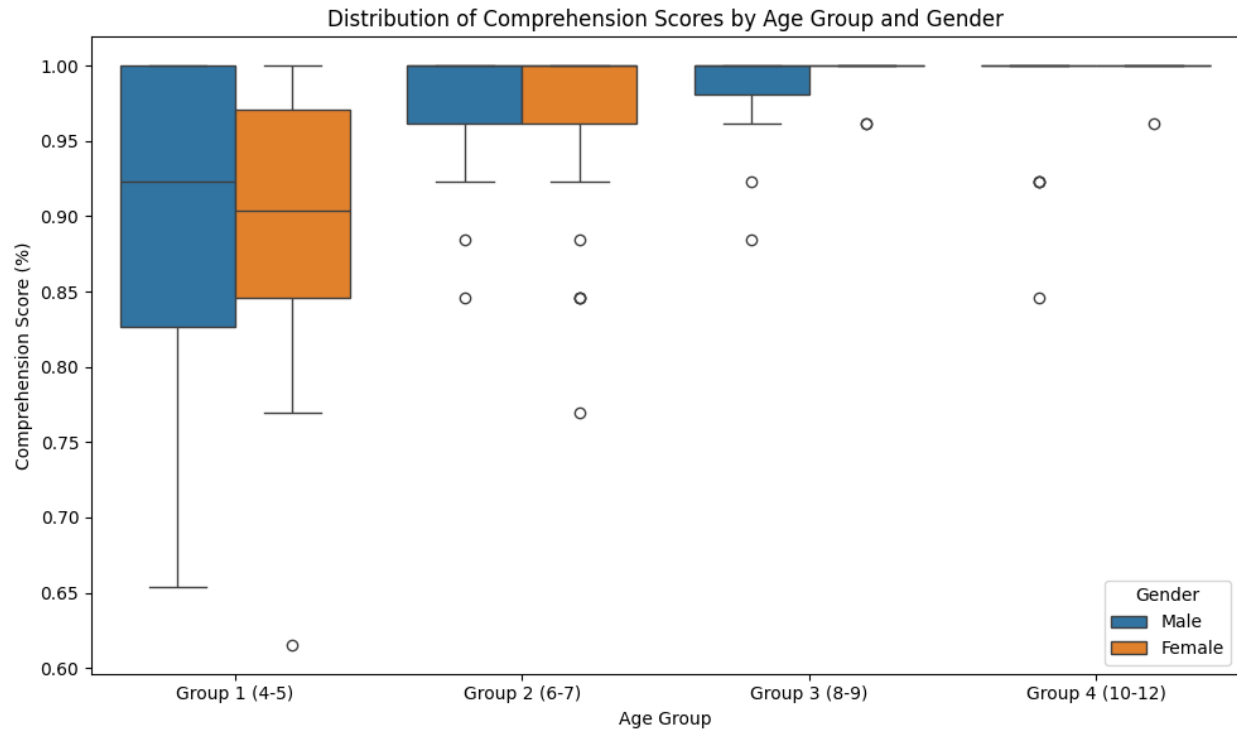
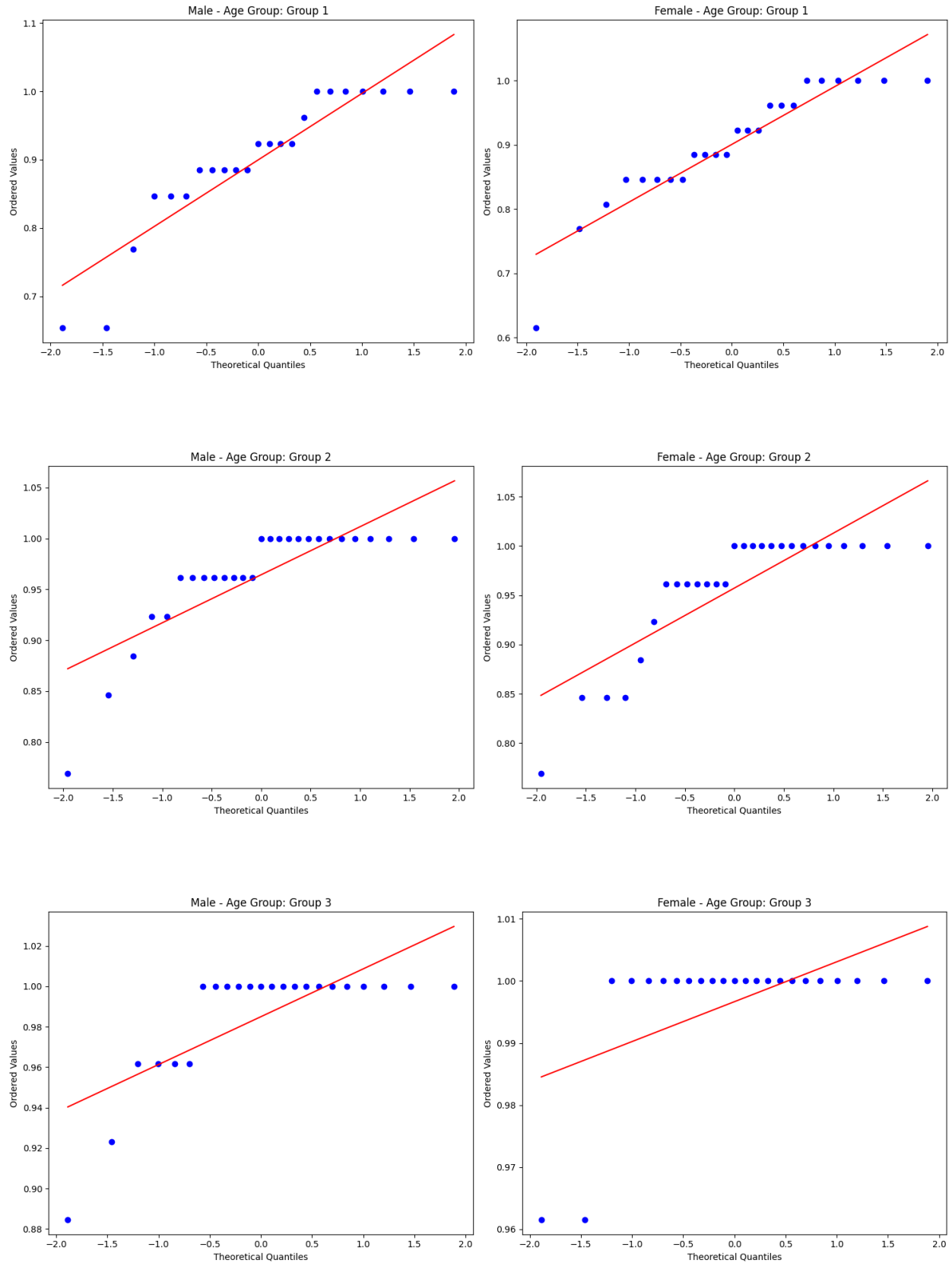


Figure 5: ox plot visualization of comprehension scores of male and female across different age groups.

visually that the data is not normal (Fig6). Based on the above results, it is evident that the data is not normal, hence we go for the non parametric Mann-Whitney U test for our hypothesis testing. We checked the effects of gender in different age groups, here again we



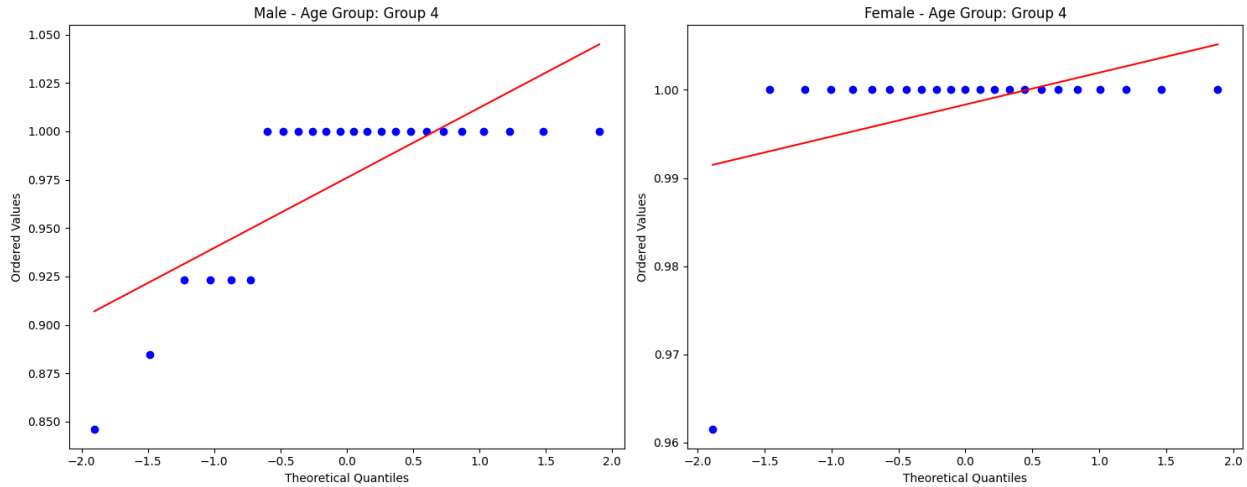


Figure 6: Q-Q plots of male and female data for different age groups.

did the Mann-whitney U test as the groups are not normal and also not equal. The obtained results are in Fig. 7.

```

]
[
Age Group: Group 1 (4-5)
Mann-Whitney U Statistic: 312.0
P-Value: 0.8254782427993288
Fail to reject Null Hypothesis (H0): There is no difference in median comprehension scores between genders.

Age Group: Group 2 (6-7)
Mann-Whitney U Statistic: 374.0
P-Value: 0.6602701286391891
Fail to reject Null Hypothesis (H0): There is no difference in median comprehension scores between genders.

Age Group: Group 3 (8-9)
Mann-Whitney U Statistic: 216.5
P-Value: 0.11328205141840612
Fail to reject Null Hypothesis (H0): There is no difference in median comprehension scores between genders.

Age Group: Group 4 (10-12)
Mann-Whitney U Statistic: 204.0
P-Value: 0.11451577218461767
Fail to reject Null Hypothesis (H0): There is no difference in median comprehension scores between genders.

```

Figure 7: Mann-Whitney U test results for Hypothesis 2

Hypothesis 3:

For this hypothesis also, firstly we visualized the data using histograms to get an idea of the nature of the data variables (Fig.8). After this, we tried the statistical tests to check for normality using Shapiro Wilk test. We found that the data is not normal as p-value for normality testing for group 1 and group 2 is less than 0.05, and hence we go for the Mann-Whitney test.

With this test, we see that the p value is more than 0.05, this shows that we fail to reject null hypothesis. In conclusion we can say that there is not enough evidence to back our rationale.

.

.

.

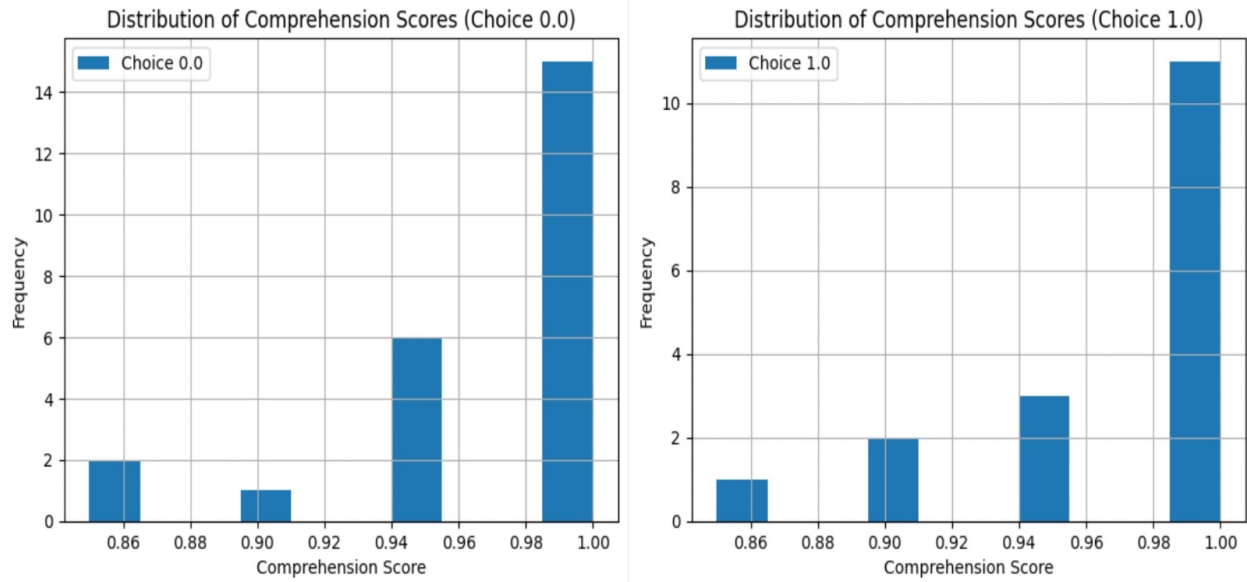


Figure 8: Histogram plot for distribution of comprehension score of group 1 (Choice 0) and group 2 (Choice 1)

Shapiro-Wilk test for normality:

```
Normality test for info_choice 0.0: p-value = 3.846083018288482e-06
Normality test for info_choice 1.0: p-value = 8.29836935736239e-05
Normality assumption violated. Consider a non-parametric test like Mann-Whitney U test.
```

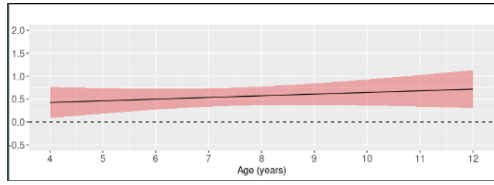
Mann-Whitney U test:

```
➡ Mann-Whitney U Test:
Test Statistic (U): 202.5
p-value: 0.9753599628872213
Fail to reject null hypothesis. Evidence is inconclusive about the medians.
```

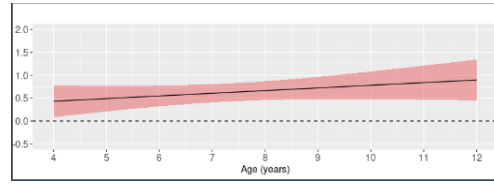
Figure 9: Results for hypothesis 3 (Shapiro-Wilk test and Mann-Whitney U test)

Hypothesis 4:

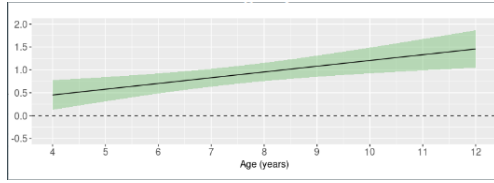
- We begin by fitting the logistic mixed model applying ridge regularization on the dataset (segregated based on gender). We fit this model individually on the parameters ΔEV (expected value), Δagency , and $\Delta \text{uncertainty}$ against age (in years).
- Fitting this model provides with β coefficients for each parameter which signify the correlation.
- Using the slopes obtained from the mixed model, we use Johnson Neyman technique to estimate the age at which the effects of uncertainty on confidence in information-seeking became significant.



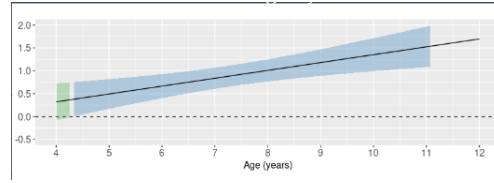
(a) β Coeffs Δagency for male children



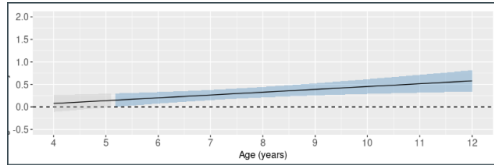
(b) β Coeffs Δagency for female children



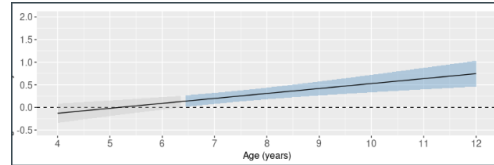
(c) β Coeffs ΔEV for male children



(d) β Coeffs ΔEV for female children



(e) β Coeffs $\Delta \text{Uncertainty}$ for male children



(f) β Coeffs $\Delta \text{Uncertainty}$ for female children

Figure 10: Estimated fixed effects of EV, agency, and uncertainty on information-seeking choices as a function of age

- Based on the above plots we observe that there is no difference plots of ΔEV and ΔAgency , however for $\Delta \text{Uncertainty}$ we observe slight deviation between the average age of male and female children in terms of developing information seeking behaviour for reducing uncertainty. However, the effect is negligible and we fail to reject our null hypothesis.
- We provide original plot of original study for reference

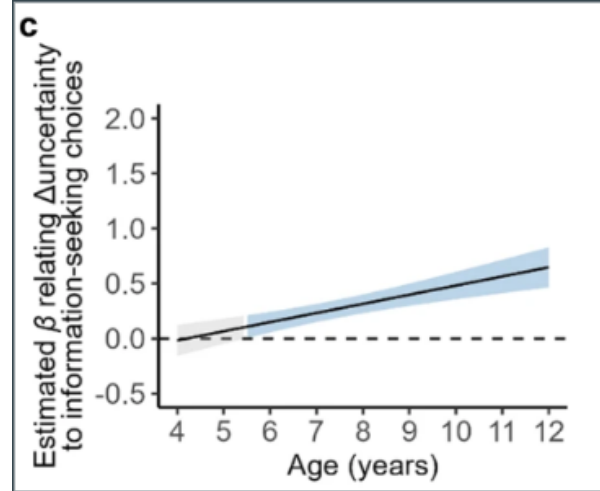


Figure 11: Estimated fixed effects of Uncertainty

6 Discussion and conclusion

We proposed to inquire about differences in brain development and information seeking motives between male and female children from pre-school(4-5) to adolescence(10-12).

- Male and female children follow similar trajectories in the context ΔEV (expected value), $\Delta agency$, and $\Delta uncertainty$. The comprehension scores across different ages also resonates the same.
- Through various hypothesis' and statistical analysis we found that there is no statistically significant difference between the information seeking motives between male and female children.

References

- [1] Caroline J Charpentier, Ethan S Bromberg-Martin, and Tali Sharot. "Valuation of knowledge and ignorance in mesolimbic reward circuitry". In: *Proceedings of the National Academy of Sciences* 115.31 (2018), E7255–E7264.
- [2] Kiyohito Iigaya et al. "The value of what's to come: Neural mechanisms coupling prediction error and the utility of anticipation". In: *Science advances* 6.25 (2020), eaba3828.
- [3] Kenji Kobayashi and Ming Hsu. "Common neural code for reward and information value". In: *Proceedings of the National Academy of Sciences* 116.26 (2019), pp. 13061–13066.
- [4] Tali Sharot and Cass R Sunstein. "How people decide what they want to know". In: *Nature Human Behaviour* 4.1 (2020), pp. 14–19.