Schoolchildren's Transitive Reasoning with the Spatial Relation 'is left/right of': A Replication and Extension Study

Amit Mettapalli

Neel Patel

Sharona Yang

Tauren Brown

Tayeba Chowdhury

Undergraduate Student Mentor: Anant Sherwal

Graduate Student Mentor: Brooke Staveland

Cognitive Science and Psychology ULAB

University of California, Berkeley

Abstract

The mental model theory (MMT) proposes that reasoners mentally construct iconic representations of the information they have processed (Johnson-Laird, 2020). This study focuses on the mental model theory to explore the effect of working memory on reasoning, how mental models are represented internally, and how these features of reasoning vary across developmental stages. We referred to Demiddele et al.'s (2019) methods as the foundation of our analysis of the mental model theory in schoolchildren. The main finding of this study was that when note-taking is available, participants spontaneously draw iconic representations of the information consistent with MMT. Their performances varied with and without notes, with participants generally scoring more points when they took notes. In our replication, we attempted to recreate the results of the original study through an in-depth analysis of the statistical models utilized that assessed the implications and the application of the mental model theory in schoolchildren. In our extension, we discovered that when taking into consideration the number of notes taken by the children working through the various problem types, the results yielded another perspective with regards to the mental model theory, essentially showing that note-taking for some problems was not beneficial. The trends also aligned with the original study when considering the groups that were allowed to take notes compared to those that weren't. The findings of this study will allow future researchers to build upon the mental model theory and apply it in other areas of study.

Introduction

There are two main different kinds of theories of how we perform relational reasoning: Syntactic and semantic. Syntactic refers to symbols such as greater than or less than. These theories originated with Jerry Fodor, who described relational reasoning as symbolic reasoning (Pessin, 1995). Semantic reasoning, on the other hand, refers to the inherent meaning of an imagined arrangement (Katz et al., 1963). According to the mental model theory, reasoners are able to mentally construct representations with the information they have processed (López-Astorga, 2021). Johnson-Laird's "Mental model theory says that reasoners mentally construct representations that are iconic to the information they have processed" (Johnson-Laird, 2020). An example of the mental model is the premise "the hat is left of the shirt". To visualize the locations of each item, the abbreviations of each word would be written down relative to each other. In this phrase, for example, it would be written as H S. Such mental models require the exercise of working memory, and as the complexity of these models increases, so does the use of working memory (Demiddele et al., 2020). The reasoners do not memorize these premises but instead, add them to their mental models in increments (Todorovikj et al., 2019). For example, if one were to add on to the hat and shirt from the aforementioned phrase, stating that "the jacket is to the right of the shirt" can be elaborated as H S J. This can be even furthered by having the phrase "the cap is to the right of the shirt" be H S J C or H S C J. Reasoners that consider both tend to spend more of their working memory than those who only consider one (Richardson et al., 1994).

To gain a better understanding of the mental model theory, Demmiddele et al.'s paper from 2020 - regarding schoolchildren's transitive reasoning with spatial relations - provided a perspective of the MMT in children. There were three types of problems that assessed different ways school children approached problems with spatial relations. M1 problems were single-model, meaning that they presented relational rules that constrained the order of clothing to only one arrangement. MMv (Multi-Model, valid) problems were multi-model, meaning that they presented relational rules that constrained order such that more than one arrangement was possible, and contained a valid answer to the question because, in all possible arrangements, the relation of the objects in the question remained the same. MMnv (Multi-Model, not valid) problems were multi-model problems with no valid conclusion, meaning that they presented relational rules that constrained order such that more than one arrangement was possible, and contained no valid answer to the question because, out of these multiple possible arrangements, the relation of the objects in question was ambiguous and could not be guaranteed. This paper found that note-taking led to higher scores overall in M1 and MMv problems. Note-taking led to lower scores in MMnv problems.

This finding then sparked another question: based on the mental model theory, is there a significant correlation between the performances being higher without notes than with notes for the MMnv problems? To approach this question, we first replicated the experiment. Our replication consists of recreating the results of an in-depth analysis of the statistical models utilized that assessed the implications and the application of the mental model theory in schoolchildren (Demiddele et al., 2020). In our extension, we sought to answer the question: based on the mental model theory, is there a correlation between the performances being higher without notes than with notes for the MMnv problems? This study answers a variety of questions regarding the way schoolchildren approach problem-solving with phrases that have spatial relations and whether or not note-taking was beneficial to the students' performance. While the original study demonstrated that the group that took notes on the MMnv problems did worse on average than the group that did not take notes, here we will extend these findings by looking at whether or not there was a dose effect in the number of notes taken on the students' performance. In other words, did any amount of note-taking lower the score by a certain amount, or did students do worse the more notes they took? We hypothesize the latter because the mental model theory would be more applicable when solving problems without a valid conclusion than with the other problem types, so there would be a correlation between the results being higher without notes than with notes for those problems (Schneider, 2001).

Materials and Methods

Experiment 1: Replication

For our replication, we retrieved the Rstudio CSV data of the original study from a link provided by ULAB to an Open Science Framework (osf.io). We successfully loaded their project data into Rstudio from a folder we downloaded from osi.io and downloaded the necessary dependencies. The problem accuracy was modeled using a linear mixed-effects framework with note taking and problem type as fixed effects. The sex and socioeconomic status were covariates, and the results of the student were modeled in our experiment.

The files ‘processedDataExp1.csv’ and ‘processedDataExp2.csv’ contain the data of the participants in Experiment 1 and Experiment 2 of the original study, respectively. These files contained the results of how the students performed. The files ‘ffData1.csv’ and ‘ffData2.csv’ contain the notes the students took, their school, class, gender, number of notes taken, and other relevant information regarding the participants. All of the data were loaded onto RStudio. The plots that included the results of the replication were created using the ‘yarrr’ package in RStudio. We altered the colors of the violin plot (Figure 6) and also the box plot (Figure 1) using ggplot2 and color codes, which we referred to from the official R colors sheet.

Experiment 2: Extension

For our extension, we processed participant results of the MMnv reasoning test (problems with no valid conclusion) by extracting the MMnv data from the full test score CSV file. The number of notes taken by each student was also provided with the score. We hypothesized that there was a correlation between the performances being slightly higher without notes than with notes, so to visualize the accuracy of our hypothesis, we made a linear regression plot in R. The x-variable of the graph was stored as how many notes were taken. The y-variable of the graph was stored as the score awarded to each student. The score was out of four points, with four being the highest amount of points awarded. Then, with the data that we stored, we first plotted the graph. We then enhanced the plot to add colors to the plots as well as the labels of the axes, by adding on to the plot() command. Then, we determined the slope of the linear regression line by using the lm() command on both of the variables. With the slope and intercept now found, we then added those values to the plot() code to generate the finalized linear regression graph. The package that made it possible to fit the linear models was 'lm'.

Results

Replication

Similar to the original study's findings with children from grade 3, we found that grade 4 children's scores on M1_v and MM_v were slightly higher when notes were allowed, while their scores on MM_mv problems showed lower percentages when allowed notes. In grade 6 children, the children allowed notes had larger increases in scores than their grade 4 counterparts when compared to having no notes for M1_v and MM_v problems, but had lower scores on MM_mv problems when allowed notes compared to those without. Both of these trends in scores align with the original study results.

Figure 1. Experiment 1 mean scores in percentages by notes group, grade, and problem type.

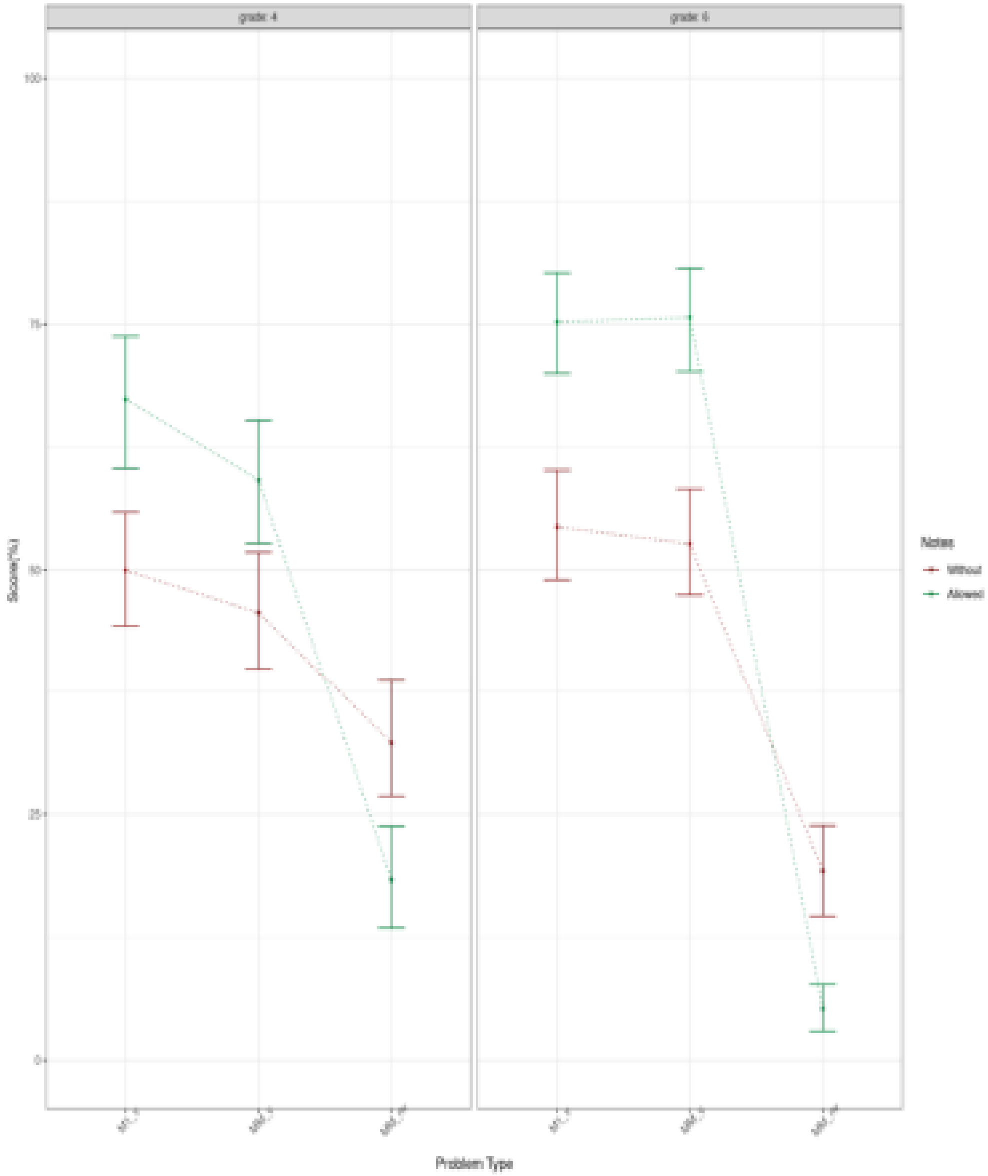


Table 1. Experiment 1 mean scores in percentages by notes group, grade and problem type.

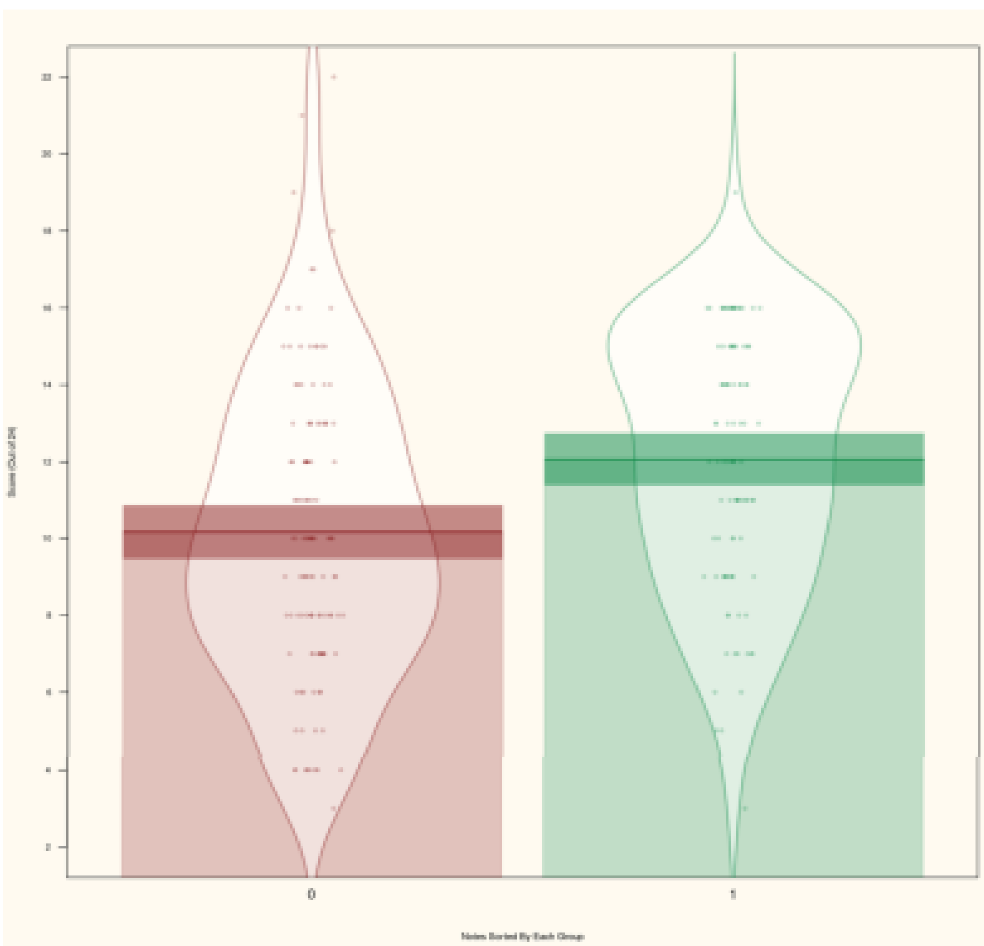
Problem type	Grade	Estimate Standard error Z	p
M1_v	6	1.05 0.19 5.39	< .001
MM_v	6	1.15 0.20 5.91	< .001
MM_nv	6	−1.58 0.28 −5.62	< .001
M1_v MM_v	4 4	0.79 0.21 3.72 0.59 0.21 2.82	< .001 .005
MM_nv	4	−0.82 0.23 −3.57	< .001

Table 2. Differential effect of note-taking across grade for each problem type.

Problem type	Estimate Standard error Z	p
M1_v	0.26 0.29 0.92	.356
MM_v	0.57 0.29 1.99	.046
MM_nv	−0.76 0.36 −2.09	.037

When viewed across groups allowed, and not allowed to take notes, scores for the children in both grades were on average higher than those not allowed to take notes. This corresponds to the findings in the original study as well, solidifying the result of the note-taking increasing scores. The results in both cases were very similar in this regard. The standard errors, z scores, along with overall margins were consistent with the replication and the initial study.

Figure 6. Total scores in Exp 1 by notes groups



We replicated Figure 6 from Demiddele et al.'s (2019) study, which was a violin plot that sorted the scores of both age groups. These are separated into two groups: notes taken and no notes taken. The findings demonstrate that the mean score is higher when notes are taken to solve the problems. On the left side of the graph, which grouped the scores without any notes being taken, the mean score is slightly over 10 out of 24 possible points. On the right side of the graph, the mean score is approximately 12 out of 24 possible points.

Extension

When the number of notes taken was taken into account for the children solving MM_nv problems, a general trend that emerged was that children that took more notes generally had lower scores on the MM_nv problems compared to children who took fewer or no notes at all.

Figure 1. Number of Notes Taken vs. Score for the MM_nv Problems

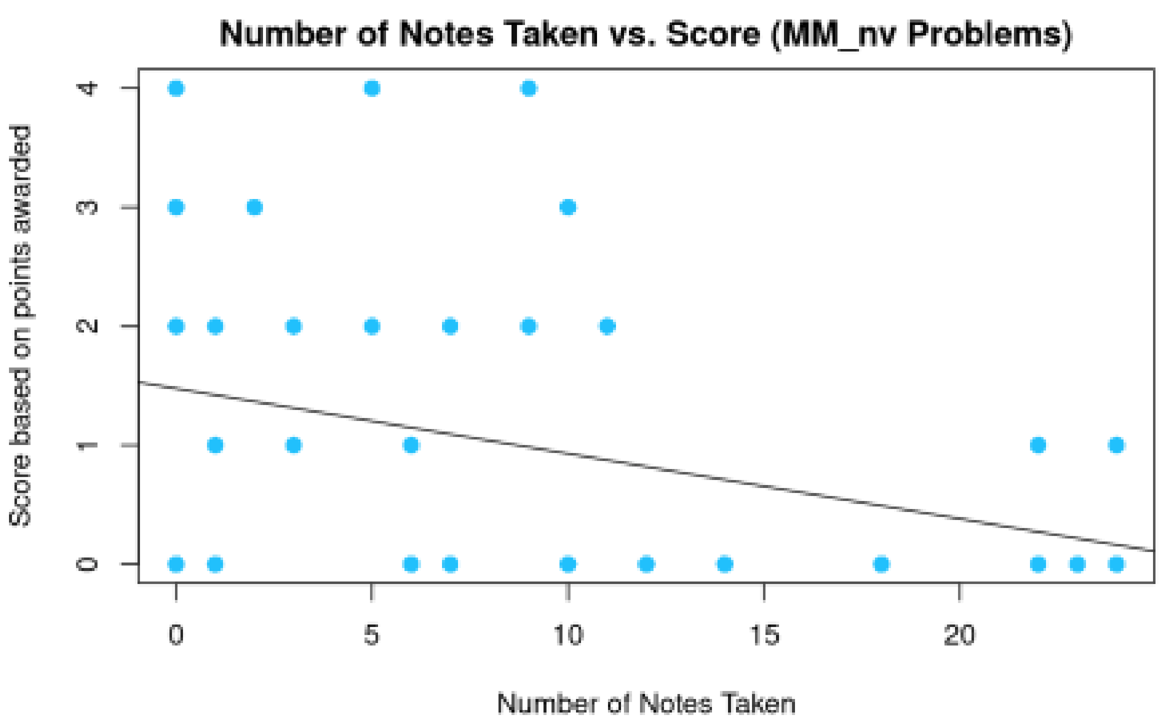


Figure 1 is a linear regression plot that seeks to answer the main question of the extension: is there a significant correlation between student scores and note-taking within the group that took notes on the MMnv problems? Our results show there is a significant negative correlation on the graph that demonstrates that as more notes are taken for the MMnv problems, the likelihood of scoring higher decreases. The y-intercept of this graph is 1.47594. This value indicates what the average score would be if zero notes were taken for any MM_nv problem. The slope of the regression line of this graph is -0.05477, the p-value was 0.03, with a beta coefficient of 0.001 ± 0.0005, and a t-value of 2.161. The data demonstrates a slight negative correlation, the slope being -0.05477 indicates that this effect may be negligible. While the slope on its own does not demonstrate if the effect is meaningful, the effect size is significant. The beta value being -0.05 means that the score drops by approximately 12% of the total possible score as the number of notes taken progress.

Discussion

Replication

The replication was done by compiling a line graph and violin plot in R Studio based on the dataset given to us. The results in figure 1, show us that in the third question, notes had a negative effect on performance because the notes would confuse the kids on the question that has no valid conclusion. In figure 2, we can see that the mean score for the group with notes was higher than that for the group without notes. Overall, participants strongly prefer concluding only one possible solution even when multiple are possible, and reasoning performance does not invariably improve with working memory alleviation.

In Figure 1, the group that did not have notes did better on the third question (MM_nv) for both the 4th and 6th graders. This was also the only question in which the 4th graders outperformed the 6th graders. In this question, the correct answer was always that there was “no valid conclusion”, so the usage of notes could have over-complicated the problem for the students which caused them to do worse than those who had notes. The students would have used these notes to think of a right answer, even when there is none. Additionally, the 6th graders could have done worse for the same reason.

In figure 2, we see that those who had notes scored better on average than those who didn’t. This result was expected as the students who are given the ability to take notes should get a higher score, because they have to remember less information that is presented to them. In absolute terms, the difference in scores between those who had notes and those who didn't may not seem too large, but there is around a 12 percent increase in scores for the kids who had notes. 12 percent is a sizable difference, but we expected a higher increase when the students were given the opportunity to take notes, around 20 percent. We believe that this raises the question that the questions were too simple, and the children did not need the notes to do well. We believe that further research could be done, where the questions are much harder and more complex to see if having notes will have a bigger impact.

This study allowed us to view the difference in transitive reasoning and working memory between two different age groups, and also between groups who were allowed to take notes and those who were not allowed to take notes. The majority of the results pointed to the older students having higher scores, and this can be attributed to the fact that part of transitive reasoning comes from the prefrontal cortex, and this part of the brain grows as we grow older (Koscik & Tranel, 2012).

Extension

In Figure 1, we can see that there is a negative correlation between the number of notes taken and the score. This finding is significant because it shows a different perspective of the mental model theory. It shows the usage of notes does not alleviate the stress of working memory, and can not be relied on to consistently improve one’s score. The M1 and MMv problems have students scoring significantly higher, compared to the scores of the MM_nv problems. This has more to do with the simplicity of the problems and how easily the kids were able to interpret the problems, regardless of note-taking. The MM_nv problem types required a different way of thinking, and this is because of the process of problem-solving. Taking more notes may have ended up confusing the students, thus causing them to answer the problem incorrectly.

Phrases depend on problem type; the problems asking simple arrangement of items allow one to map it out in a coherent fashion and thus answer the problem properly. This was not necessarily the case with the MM_nv problems because taking notes using a visual model was not beneficial in this problem set.

MM_nv problems do not have a conclusive answer; the answer for those problems is always "None of the above." In this case, the mental model theory can be applied as it causes one to mentally visualize the applications of the phrases with spatial relations to reach an answer.

The extension and replication showed two different results, signifying that there needs to be more work done on this topic to come to a valid conclusion. Further research should be done in which the difficulty of the questions is increased, and where there are different types of problems.

Conclusion

This study sought to more deeply understand working memory's effect on reasoning ability by allowing children to minimize working memory load by taking notes while they performed relational reasoning tasks. Seeing as this note-taking did not help them solve problems involving insufficient logical constraints, we conclude that there is more to good reasoning than working memory. Our replication yielded results consistent with the original paper, and our extension confirmed a suspected trend in the MMnv problem that taking more notes leads to lower scores than taking few or no notes at all. A possible explanation for this result is that taking notes solidifies the children's mental model, while when held in their mind their mental model is either dynamic or multiplicitous (Markovitz et al., 2002). Future researchers interested in validating or disproving these results may want to find a better proxy to more accurately simulate working memory as an isolated variable. For example, researchers might give the participants physical props they can move around instead of leaving them to make static drawings.

References

Demiddele, K., Heyman, T., & Schaeken, W. (2020). Schoolchildren's transitive reasoning with the spatial relation 'is left/right of'. *Thinking & Reasoning*. https://doi.org/10.1080/13546783.2020.1843536

Girotto, V., Mazzocco, A., & Tasso, A. (1997). The effect of premise order in conditional reasoning: A test of the mental model theory. *Cognition*, *63*(1), 1-28.

Johnson-Laird, P. N. (2020). A proof of the consistency of the mental model theory of sentential Reasoning.

Katz, J. J., & Fodor, J. A. (1963). The Structure of a Semantic Theory. *Language*, *39*(2), 170–210. https://doi.org/10.2307/411200

Koscik, T. R., & Tranel, D. (2012). The human ventromedial prefrontal cortex is critical for transitive inference. *Journal of cognitive neuroscience*, *24*(5), 1191–1204. https://doi.org/10.1162/jocn_a_00203

López-Astorga, M. (2021). Compound nouns do not always fulfill the conjunction thesis: An analysis using the mental model theory. *Lenguaje*, *49*(2), 484-496.

Markovits, H., & Barrouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review*, *22*(1), 5-36.

Pessin, A. (1995). Mentalese Syntax: Between a Rock and Two Hard Places. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *78*(1), 33–53. http://www.jstor.org/stable/4320579

Richardson, G. P., Andersen, D. F., Maxwell, T. A., & Stewart, T. R. (1994). Foundations of mental model research. In *Proceedings of the 1994 International System Dynamics Conference* (pp. 181-192). EF Wolstenholme.

Schneider, R. (2001). Toward a cognitive theory of literary character: The dynamics of mental-model construction. *Style*, *35*(4), 607-639.

Todorovikj, S., Friemann, P., & Ragni, M. (2019). A Cross-Domain Theory of Mental Models. In *Pacific Rim Knowledge Acquisition Workshop* (pp. 184-194). Springer, Cham.