Schoolchildren's Spatial Reasoning

Kevin Demiddele[1] (kevin.demiddele@kuleuven.be), Tom Heyman[1,2] (t.d.p.heyman@fsw.leidenuniv.nl), Walter Schaeken[1] (walter.schaeken@kuleuven.be)

[1] Laboratory of Experimental Psychology, KU Leuven, Tiensestraat 102, 3000 Leuven, Belgium

[2] Leiden University

Abstract

We examine schoolchildren's reasoning with spatial relations, such as 'is to the left of'. Our aims are to obtain a more precise account of the effect of working memory on reasoning, a more detailed understanding of the internal representation of mental models and a developmental perspective. We discuss two experiments in which 348 children, between eight and twelve years old, needed to verify conclusions for 24 reasoning problems describing the spatial relations between pieces of clothing. In both experiments, children in the experimental condition were allowed to take notes by means of paper and pencil. In both experiments we find that the participants spontaneously draw iconic representations of the items' spatial ordering, have a strong preference for only considering one possible state of affairs even when more are relevant, and that an explanation in terms of working memory capacity alone cannot fully explain the data.

Keywords: spatial reasoning; mental models; working memory; developmental psychology

Introduction

Arguably, the most widely accepted theory of how people reason with spatial relations is mental model theory (cf. Johnson-Laird, 1983; 2006). It says that reasoners mentally construct representations that are iconic to the information they have processed. Consider, for example, the following premise.

The hat is to the left of the shirt

This can be represented by the mental model below.

H S

Keeping such models in mind taxes working memory and the more complex they are, the more will be demanded in terms of working memory. Reasoners will not keep the premises in mind, but, according to the principle of economicity (Cf. Manktelow, 1999), incrementally add the information in the premises to their mental model. Consider, for example, the following premise being added to the information.

The dress is to the right of the hat

Now at least two different representations can be constructed.

H S D *and* H D S

Both are consistent with the premises. Anyone realizing this and representing these possibilities mentally, needs to invest more working memory effort than someone only representing one of both. In fact, it has convincingly been shown (Goodwin & Johnson-Laird, 2005; Jahn, Knauff, & Johnson-Laird, 2007; Ragni, Fangmeier, Webber, & Knauff, 2007) that reasoners by default only do construct one of both, the explanation being that this is done because of parsimonious use of working memory. In general, working memory capacity is known to limit reasoning capability (Cf. Bara, Bucciarelli, & Johnson-Laird, 1995; Gilhooly, Logie, Wetherick, and Wynn, 1992; Klauer, 1997).

What is currently lacking, is a precise account of the effect of working memory, more detailed understanding of the internal representation of mental models and a developmental perspective. To obtain the developmental perspective, we have chosen children in different age groups as participants for our experiments. For more insight in their internal representation and a more specific understanding of the effect of working memory capacity, we allow them to use paper and pencil in the experimental condition. The assumption that reasoners who are given paper and pencil will write down something that matches their mental representations to some extent, has been made earlier by, for example, Van Der Henst, Yang, and Philip (2002) and Bucciarelli and Johnson-Laird (1999). On the basis of several illustrative problems in mathematics and physics, Larkin and Simon (1987) argue convincingly that external and internal spatial representations can be viewed as equivalent. Larkin and Simon (1987) state that the creation of a mental image employs inference processes similar to those that make information explicit in the course of drawing a diagram (for a similar argument see also Evans, 2000). Huttenlocher (1968) made a similar argument with three-term series problems and claimed that what people are physically doing when they manipulate objects on a scale is isomorphic to the mental processes involved when they reason without objects. Moreover, taking notes by means of paper and pencil can be an auxiliary to remember things and thus, alleviate the weight on working memory, as explained by Bauer and Johnson-Laird (1993).
In order to maximize this effect of the working memory aid, it also makes sense to try this kind of experiment with participants that are known to have limited working memory, i.e. children (see Gathercole, Pickering, Ambridge, & Wearing, 2004). At the same time, we should be ascertained that participants have the ability to conduct the type of reasoning exercises we want to present them. Our participants

ranged from eight- to twelve-year olds. This is old enough to deal with transitive inferences (see Andrews & Halford, 1998).

If children's reasoning performance is impaired because of limited working memory, and we provide them with a mechanism, i.e. taking notes, to overcome this limitation, we would expect the result to be a significant improvement of reasoning performance. Moreover, based on the working memory capacity account, we would expect a clear effect for the cases with multiple possibilities, as these put the heaviest load on working memory. That is, unless the bias to construct only one single model is based on more than working memory limitation alone.

Data of both Experiment 1 and 2 are available at OSF.[1]

Experiment 1

Method

Participant Info We tested 216 children. There were 106 boys and 110 girls; 120 sixth-graders (M = 11.48 years; SD = 0.28) and 96 fourth-graders (M = 9.47 years; SD = 0.28). All data was collected at schools in Flanders. All children's data was anonymized before processing. The experiment was approved by the social and societal ethics committee of KU Leuven (G-2017 11 970) and all participating children had an informed consent signed by their parents. Among the fourth-graders, 55 children were assigned to the control condition and 41 to the experimental condition. Among the sixth-graders, 65 and 55 children were in the control and experimental condition, respectively.

Procedure The experiment leader collected the data per class. Each class as a whole was assigned to either the experimental or control condition. During the data collection, she followed a script to ensure that the same instructions were given in each class. Before handing the exercises to the children, she introduced the topic by collectively solving a real life example of the type of problem they were to encounter in the exercises. She showed that a sweater was to the right of a pair of trousers and a cap was to the right of the sweater. Then she agreed with the children that the cap was also to the right of the trousers. Once the introduction was over, they each received their exercise sheets and could start working individually.

Material All children received 24 reasoning problems. These consisted of premises describing relative positions of pieces of clothing, on the basis of which they had to draw a conclusion that could be chosen from multiple choice options. These indicated that what could be concluded from the premises was either that the questioned object was left of the other object, right of it, or no conclusion was possible.

There were three problem types. The first problem type, single model problems (M1), consisted of premises that describe an unambiguous arrangement of items. Hence, for all of these problems it was possible to infer what the relation between the question items was. Here is an example of a single model problem, translated to English from the Dutch original. In Dutch all nouns were singular.

The trousers are to the left of the hat.
The skirt is to the left of the trousers.

Where are the skirt and the hat relative to each other?

The correct, unambiguous mental model that can be constructed for these premises is

S T H

The second problem type, multiple-model problems with a valid conclusion (MMv), consisted of premises that describe a situation consistent with two different arrangements of items, but posed a question on items that had the same relation in both of these representations. Thus, for these problems, too, the correct left-right relation between the question items could be inferred. It is important to realize that these problems could be answered correctly by constructing only one of the two possible representations and judging the conclusion based on that one model, possibly without even realizing that multiple possible representations are involved. Add "the dress is to the left of the trousers" as third premise to the example above. The result is a description of a situation where the dress can be either left or right of the skirt, but this does not matter for the question at hand, as in both possible representations, the skirt is to the left of the hat.

D S T H and S D T H

Finally, the third problem type were multiple-model problems with no valid conclusion (MMnv). In these problems, premises also described a situation consistent with two different arrangements of items, but now a question was posed on items that had a different relation in both of these representations. No valid conclusion could be inferred from these premises. To obtain such a problem, again add "the dress is to the left of the trousers" as third premise to the problem above, obtaining the same two possible mental models, but now change the question to "where are the skirt and the dress relative to each other?". As this relation is different in each of the two models, a valid conclusion is not possible.

Each child solved eight M1 problems, eight MMv problems and eight MMnv problems. In the experimental condition, there was blank space (7cm x 16cm) below the question and answer options, where the children could write or draw what

[1] osf.io/ukep3/?view_only=93de6a189e0f4e2f838f37dcabf5fe32

they wanted. In the control condition, there was no blank space and children were not allowed to take notes, which was checked by the experiment leader.

Results and Discussion

No data were excluded. We performed a multilevel analysis, predicting scores from type, notes and grade including all their interactions, with individual participants as random intercept. The R^2 of the resulting model was .38 (based on the theoretical variance) and .33 (based on the observation-level variance via the delta method of Nakagawa, Johnson, & Schielzeth, 2017). If we only look at the fixed effects, the R^2 estimates were .29 and .26, respectively. Critically, there was a significant three-way interaction between notes, grade, and problem type ($\chi^2(2) = 12.81$, $p = .002$), but the pattern of results was not one we had expected a priori (see Figure 1).

Taking notes proved to be beneficial for only two of the three problem types. Both fourth- and sixth-graders

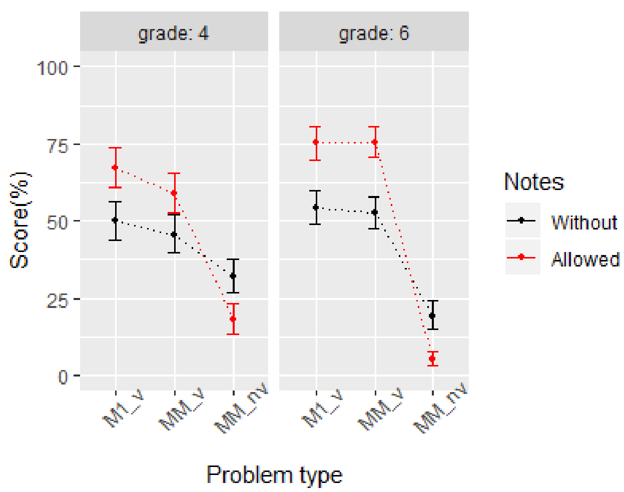

Figure 1: Experiment 1 mean scores in percentages by notes group, grade and problem type.

performed significantly better on single-model problems and multiple-model problems with valid conclusion when given the opportunity to take notes. Surprisingly, they performed worse on multiple-model problems without valid conclusion when allowed to take notes, the expectation being a significant improvement especially for these problems. Moreover, compared to fourth-graders, sixth-graders showed a stronger beneficial effect of note taking on multiple-model problems with valid conclusion, but also a stronger adverse effect of note taking on multiple-model problems without valid conclusion (see Table 1 and Table 2).

Table 1: Effect of note taking for each problem type by grade combination

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	4	0.79	0.21	3.72	< .001
MM_v	4	0.59	0.21	2.82	.005
MM_nv	4	-0.82	0.23	-3.57	< .001

Table 2: Differential effect of note taking between grade (6 minus 4) 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

Our interpretation is that most children always constructed just one model. This is in accordance with the claim (in Goodwin & Johnson-Laird, 2005; Jahn et al., 2007; and Ragni et al., 2007) that, not only children, but all reasoners in general by default build a single, simple, and typical mental model but neglect other possible models. The children who could take notes were better at this than the ones that could not, but for the MMnv problems it was not the right strategy, which resulted in a worse score for that problem type when note taking was allowed. In other words, taking notes seemed to improve performance on the easier problems, whether by facilitating the synthetization of information, enhancing children's focus and motivation, and/or decreasing the working memory load. At the same time, it appeared to prompt them to disregard multiple possibilities. This was especially true for older children, presumably because note taking skills improve with age.

One of our research goals was getting a better understanding of the internal representation of our reasoners. For this reason we provided them only with blank space and no instructions about how they could use it, apart from stating that they could use it in any way they wanted. This way, we wanted to see what kind of notes they would make spontaneously. It turned out that they spontaneously drew analogical representations, with drawings, words or letters representing the items, arranged on paper in the same way as described in the premises. This arguably shows that they also mentally used the same type of representations for solving these problems. Their notes can be interpreted as an external representation of their mental models. The ones that started out by drawing the items often switched to more economical representations as the experiment proceeded. This was likely because they grew tired of drawing unnecessary details but at the same time can have been functional as for mental reasoning activity visual imagery can impede reasoning (Cf. Knauff & Johnson-Laird, 2002).

Our specific interest was in whether and how they would represent multiple possibilities. Given the results, it will not come as a big surprise that not many children did represent multiple possibilities. Of the 97 children that were in the notes condition, we only identified one who drew representations of multiple possibilities. This sixth-grader drew multiple models three times for MMnv problems and correctly answered the corresponding questions.

Experiment 2

Experiment 1 demonstrated that the children showed a strong preference for only considering one possible state of affairs. When they could take notes, they improved at this strategy, even in the cases where it was the wrong strategy. A question that naturally arises from this result, is how we can stimulate them to consider multiple possibilities. And when we succeed at this, it would be interesting to see whether the children in the notes condition can also use their notes to their benefit for the MMnv problems.

In Experiment 2 we tried a very simple approach to reach this goal: we added a multiple possibilities example to the example that the instructor presented before the experiment. This was a relatively minimal adaptation, but now at least there was an example case in which it was shown that, e.g., the trousers could be either left or right of the sweater given certain information, whereas the example in Experiment 1 only concerned a case with a definite answer.

Method

Participant Info We tested 132 children. There were 62 boys and 70 girls, 63 sixth-graders (M = 11.42 years, SD = 0.25) and 69 third-graders (M = 8.29 years, SD = 0.25). All data was again collected at schools in Flanders, different from those of Experiment 1. All children's data was anonymized before processing. The ethical approval and informed consent were identical to those of Experiment 1.

Procedure and Material The procedure and material were identical to that of Experiment 1, with the already explained exception that the instructor now also showed the children an example with multiple possibilities.

Results and Discussion

The same analysis as in experiment 1 was performed. The R^2 of the entire model was .43 (based on the theoretical variance) and .38 (based on the observation-level variance). If we only look at the fixed effects, the R^2 estimates were .35 and .31, respectively. Again, there was a significant three-way interaction between notes, grade, and problem type ($\chi^2(2)$ = 13.28, p = .001, see Figure 2).

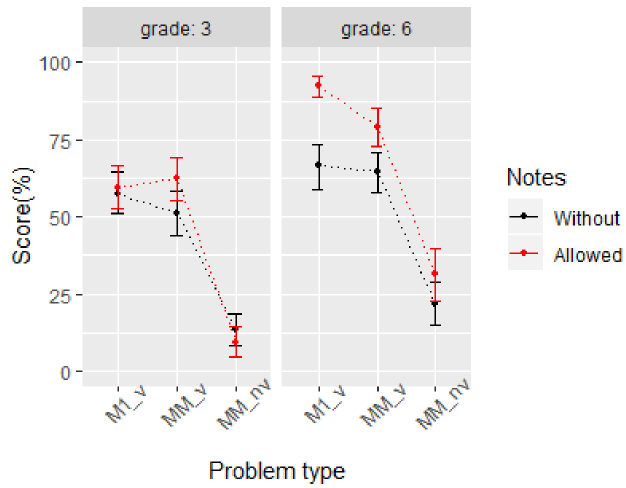

Figure 2: Experiment 2 mean scores in percentages by notes group, grade and problem type

It seems that the explicit mentioning of the cases with multiple possibilities had its effect. For Grade 6 participants, taking notes resulted in significantly better results for the problems where one model was sufficient, as in Experiment 1. But whereas in Experiment 1 those taking notes performed significantly worse at the MMnv problems, this effect was no longer present. In fact, there was a slight benefit of taking notes for those problems as well now. For the Grade 3 participants, there was only a significant difference between the notes and no notes conditions for the MMv problems. Maybe the mechanism of taking notes to represent mental models was too difficult for them, requiring some level of meta-cognition that they do not yet master. See Table 3 and Table 4 for more details on the effect of note taking. The improved scores for Experiment 2 are not just random noise. When combining the data of the sixth-graders from both experiments, we see a significant interaction of experiment with problem type ($\chi^2(2)$ = 9.14, p = . 010). The contrasts in Table 5 shows that sixth-graders in Experiment 2 score significantly higher for each problem type, compared to those in Experiment 1.

The notes they made were very similar to those in Experiment 1: analogical representations of the spatially ordered items. Four out of 67 participants in the notes condition, all sixth-graders, drew representations of multiple possibilities.

Table 3: Effect of note taking for each problem type by grade combination in Experiment 2

Problem type	Grade	Estimate	Standard error	Z	p
M1_v	6	1.95	0.33	5.92	< .001
MM_v	6	0.81	0.27	2.97	.003
MM_nv	6	0.54	0.27	1.98	.048
M1_v	3	0.10	0.24	0.42	.675
MM_v	3	0.52	0.24	2.15	.031
MM_nv	3	-0.47	0.32	-1.46	.144

Table 4: Differential effect of note taking between grade (6 minus 4) for each problem type in Experiment 2

Problem type	Estimate	Standard error	Z	p
M1_v	1.85	0.41	4.52	< .001
MM_v	0.29	0.36	0.79	.430
MM_nv	1.02	0.42	2.39	.017

Table 5: Differential effect of experiment for each problem type

Experiment	Problem type	Estimate	Standard error	Z	p
1 - 2	M1_v	-1.05	0.16	-6.52	< .001
1 - 2	MM_v	-0.62	0.15	-4.06	< .001
1 - 2	MM_nv	-1.16	0.17	-6.72	< .001

General Discussion

The first research goal we defined, was to specify the effect of working memory on relational reasoning. For problems that could be solved by means of a single mental model, we found what was expected: alleviating the weight on working memory results in better performance. For problems where considering multiple models is required, the situation is more complicated. The commonly accepted explanation of why people do not vary upon their preferred mental model, is the principle of parsimony, which is explained in terms of limited working memory capacity (Goodwin & Johnson-Laird, 2005; Jahn et al., 2007; Markovits & Barrouillet, 2002). In light of our results, it is worth reconsidering this explanation. By allowing our participants to use paper and pencil, we provided them with a mechanism to overcome the limitations of working memory. This had the expected result for the reasoning problems in which varying the preferred model was not required: the children who could take notes performed significantly better. The best explanation for this indeed seems to be in terms of working memory capacity. Now, if the reason why people do not vary their preferred model is limited working memory capacity, we would especially expect an improvement for those problems in which varying their preferred model is required. However, we saw a reversed effect (Experiment 1) or only a slight improvement (Experiment 2). This suggests that working memory capacity in itself cannot be the sole motivation why reasoners choose to refrain from constructing multiple models.

A first plausible alternative explanation that comes to mind is insufficient inhibitory control. Once one mental representation, consistent with the premises, is constructed, inhibitory control is required to not halt calculations and instead look for further possibilities

A second alternative explanation, both compatible with an account in terms of working memory and in terms of inhibitory control, concerns rational thinking dispositions. “Consideration of alternatives”, mentioned in Markovits and Barrouillet (2002), seems particularly apt for capturing the effect we are after, with “actively open-minded thinking” a more general candidate that could carry the load (Cf. Baron, 1985, 1993; Stanovich & West, 1998; Toplak, West, & Stanovich, 2014).

Evidently, these alternatives are presently mere suggestions that will require further research before strong claims can be made. More specifically, measuring these variables and seeing how well they predict the score on MMnv problems will be needed.

The second research goal, a more detailed understanding of reasoners’ internal representation, can be answered with a much clearer picture. Our young participants spontaneously drew iconic representations of the spatially organized items. It seems plausible that their internal strategy for solving the reasoning problems is also based on such representations, rather than on a logical strategy involving, for example, understanding of the transitivity of the relation “is to the left of”. Making abstraction of whether the iconic representations were drawn with images, words or letters, we can interpret them as mental models and interpret our results as supporting mental model theory.

Finally, our third research goal was obtaining a developmental perspective. It is clear that children get better at reasoning with one representation as they grow older. Less clear is whether their understanding of multiple models improves a lot at the ages we tested. In Experiment 1 the sixth-graders scored worse than the fourth-graders on the MMnv problems. The sixth-graders in Experiment 2 scored a bit better, but still only at chance level. So the extent to which they understand multiple possibilities is not that clear. One hypothesis is that producing multiple models themselves is still too challenging for them, although they may be capable of understanding the concept when it is explained to them.

What is clear, is that our third-, fourth- and sixth-graders spontaneously use iconic representations and that a strong bias for ignoring multiple possibilities is present.

One limitation of this study was the third answering option. The third multiple-choice option was 'none of the above'. This is the correct answer if you interpret the question as 'what can you validly conclude from these premises?'. As a descriptive statement, however, it is incorrect: the answer is either left or right, even if one cannot know which is the case based on the information in the premises. In retrospect, this should have been stated more clearly. However, the children were clearly instructed that this was the option to choose if they thought there were multiple possibilities or if they had the feeling there was no correct answer. This also relates to a possible aversion to answer 'no valid conclusion', as described for syllogistic reasoning by Ragni, Dames, Brand and Riesterer (2019). As a second limitation, it could be argued that the iconic nature of the notes taken was partly due to demand characteristics. Showing them examples with pieces of clothing and then leaving ample blank space for notes could have given the signal that a solution in terms of drawn items was expected. A third limitation was varying the age of the youngest group between Experiment 1 and 2. For better comparison between the experiments, it would have been more appropriate to stick to grade 4 in Experiment 2 as well.

Conclusion

We ran two reasoning experiments in which schoolchildren of two different age groups each solved 24 reasoning exercises. The control groups just had to choose the correct conclusion from three different options, while the experimental groups additionally had some blank space where they could make helpful notes before choosing their answer. In the first experiment, we observed a strong tendency to construct only a single model, resulting in a much lower score for the multiple-model problems with no valid conclusion, for which taking into account multiple possibilities is required to arrive at the correct answer. Understanding these multiple possibilities proved to be rather challenging for our young participants, even for the oldest ones and with the help of notes and some concise explanation. Taking notes was useful to improve their single-model strategy, which explains why in the notes condition accuracy was higher for the M1 and MMv problems, but significantly lower for the MMnv problems. Likewise, sixth-graders had lower scores than fourth-graders on these problems, because they were better at applying the preferred model strategy. In an attempt to overcome this preferred model bias, we explicitly showed them a multiple-model example with no valid conclusion in the second experiment, thus subtly explaining the preferred model mistake. This had some beneficial effect, but understanding of the multiple-model cases was still surprisingly low.

Based on these results, we argued that the reason why people do not vary their preferred model, is not only because of working memory considerations. Taking notes should be a substantial help on the working memory front, but did not yield very beneficial results when it comes to varying the preferred model. So parsimonious use of working memory cannot be the only reason why participants tend to be satisfied with one mental model, even in cases where multiple ones are possible. Our results suggest that producing representations of multiple possibilities is by no means evident until the age of twelve, and that the principle of parsimony for mental model construction in itself is not sufficient to explain it.

References

Andrews, G., & Halford, G. S. (1998). Children's ability to make transitive inferences: The importance of premise integration and structural complexity. *Cognitive Development*, 13(4), 479–513

Bara, B. G., Bucciarelli, M., & Johnson-Laird, P. N. (1995). Development of syllogistic reasoning. *The American Journal of Psychology*, 108(2), 157–193.

Baron, J. (1985). *Rationality and intelligence.* Cambridge: Cambridge University press.

Baron, J. (1993). Why teach thinking?—An essay. *Applied Psychology: An International Review*, 42, 191–214.

Bauer, M. I., & Johnson-Laird, P. N. (1993). How diagrams can improve reasoning. *Psychological Science*, 4(6), 372–378.

Bucciarelli, M., & Johnson-Laird, P. N. (1999). Strategies in syllogistic reasoning. *Cognitive Science*, 23(3), 247–303

Evans, J.St.B.T (2000). What could and could not be a strategy in reasoning? In. Schaeken,W., De-Vooght, G., Vandierendonck, A., & d'Ydewalle, G. (Eds). *Deductive reasoning and strategies.* Mahwah, NJ, US: Lawrence Erlbaum Associates, (pp. 1-22).

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40(2), 177–190.

Gilhooly, K., Logie, R., Wetherick, N., & Wynn, V. (1992). Working memory and strategies in syllogistic reasoning tasks. *International Journal of Psychology*, 27(3-4), 148–148.

Goodwin, G. P., & Johnson-Laird, P. N. (2005). Reasoning about relations. *Psychological Review*, 112(2), 468–493.

Handley, S. J., Capon, A., Beveridge, M., Dennis, I., & Evans, J. S. B. (2004). Working memory, inhibitory control and the development of children's reasoning. *Thinking & Reasoning*, 10(2), 175–195.

Huttenlocher, J. (1968). Constructing spatial images: a strategy in reasoning. *Psychological Review*, 75, 550-560.

Jahn, G., Knauff, M., & Johnson-Laird, P. N. (2007). Preferred mental models in reasoning about spatial relations. *Memory & Cognition*, 35(8), 2075–2087.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge: Harvard university press.

Johnson-Laird, P. N. (2006). *How we reason.* Oxford: Oxford University Press.

Klauer, K. C. (1997). Working memory involvement in propositional and spatial reasoning. *Thinking & Reasoning*, 3(1), 9–47.

Knauff, M., & Johnson-Laird, P. N. (2002). Visual imagery can impede reasoning. *Memory & Cognition*, 30(3), 363–371.

Larkin, J.H., & Simon, H.A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65-100.

Manktelow, K. (1999). *Reasoning and thinking*. Psychology Press, Hove (UK).

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

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination r2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society*, Interface, 14(134). Retrieved from http://search.proquest.com/docview/1938853354/

Ragni, M., Dames, H., Brand, D., & Riesterer, N. (2019). When does a reasoner respond: Nothing follows? In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society* (pp. 2640–2646). Montreal, QB: Cognitive Science Society

Ragni, M., Fangmeier, T., Webber, L., & Knauff, M. (2007). Preferred mental models: How and why they are so important in human reasoning with spatial relations. In *Lecture Notes in Computer Science* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 4387, pp. 175–190).

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, 127(2), 161–188.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology*, 50(4), 1037–1048.

Van Der Henst, J.-B., Yang, Y., & Philip, J.-L. P. N. (2002). Strategies in sentential reasoning. *Cognitive Science*, 26, 425–468.