Children’s Strategy Use in Playing Strategic Games

Maartje E. J. Raijmakers[1], Sara van Es[1], Marian Counihan[1]

[1]Department of Psychology, University of Amsterdam, M.E.J.Raijmakers@uva.nl

Abstract. Strategic games require to reason about other peoples and one’s own beliefs or intentions. Although they have clear commonalities with psychological tests of theory of mind, they are not clearly related to these tests for children between 9 and 10 years old [6]. We study children’s (5 – 12 years of age) individual differences in playing a strategic game by analyzing the strategies that they apply in a zero, first, and second-order reasoning tasks. For the zero-order task, there were two subgroups with different accuracy. For the first-order task subgroups apply different suboptimal strategies or an optimal strategy. For the second-order task only different suboptimal strategies were present. Strategy use for all tasks was related to age. For the 5 and 6 years old children strategy-use was related to working memory, and not to theory of mind, after correction for age, verbal ability and general IQ.

Keywords: Strategic games, child development, reasoning, theory of mind, strategy analysis.

1 Introduction

Strategic games require to reason about other peoples and one’s own beliefs or intentions. Hedden and Zang [8] designed a matrix game to distinguish the use of first-order and second-order theory of mind in adults. First-order reasoning involves a proposition of the form: “The other person (A) plays X” and second-order reasoning involves a proposition of the form: “A knows that I will play X, so A will play Y”. That is, second-order reasoning involves 2 propositions that are embedded. Hedden and Zang suggested that for optimal play in a strategic game one needs a theory of mind.

Since theory of mind is still developing into childhood [20], a limited theory of mind is expected to be a factor in children’s ability to play strategic games. The development of theory of mind is most extensively tested with false-beliefs tasks, appearance-reality tasks, and deception tasks [23,15,5,7]. These tasks all require first-order reasoning, that is, they involve a proposition “A beliefs X”. In development, theory of mind, as measured with these tasks, is strongly related to executive functions, especially a combination of working memory, inhibitory control and planning, independent of age, verbal abilities, and intelligence (e.g., [1,2]).

Second-order false-belief tests require reasoning with reference to what another person beliefs about your own intentions, that is, it involves two propositions “A believes that I believe X” that are embedded, as in the strategic game that requires second-order reasoning. Second-order reasoning is mostly studied with stories from which one has to infer a person's belief [14,15]. Second-order reasoning involves more information, more complex worded sentences, and puts more demand on working memory. Success of second-order reasoning emerges around 5 and 6 years of age, but differs substantially between tasks and studies. Second-order reasoning abilities are related to inhibition, planning, and working memory, but not in all studies independently from verbal abilities and general intelligence [13].

Flobbe et al. [6] adapted the task by Hedden en Zang [8] such that the strategic game is understandable and appealing to children. They showed that 55% of the 8 to 10 years old children perform first-order reasoning (at least 5 out of 6 items correct) and these children can show second-order reasoning above chance level. However, the game results were not related to two theory of mind tasks, a false belief task and a sentence-comprehension task. As they concluded, successful first and second-order theory of mind in 8 to 10 years old children depends crucially on the domain in which it must be applied.

In summary, we could state that, looking at the structure of the tasks, first-order and second-order reasoning in theory of mind requires the same type of reasoning as in the strategic games. However, the ability to play the strategic game appears not to be related to other theory of mind tasks in 8 to 10 years old children. In development, abilities measured by typical theory of mind tasks are related to executive functions, inhibition, working memory, and planning, but verbal abilities, general intelligence and age partially contribute to this relation. As yet, the relation between playing strategic games and executive functions is not known.

1.1 Individual differences

In Flobbe et al. [6] children make more mistakes on first-order and second-order reasoning tasks than adults, but they found also considerable variation within the group of children. The source of this inter-individual variation did not become clear. Individuals can differ in game playing in multiple ways. They can play different strategies and/or they can differ in the number of mistakes in applying one and the same strategy. By inspecting sum scores of test items, it is difficult to disentangle these different sources of variation in the data. Only the use of different strategies would indicate different insight into the games, for example first-order and second-order reasoning. The aim of the present research is to study different strategies in playing strategic games and the way these strategies are related to age, executive functions, and a standard false-belief task. For multiple cognitive domains, children appear to acquire increasingly complex reasoning strategies [18].

To this end, we studied first and second-order reasoning in playing Flobbe et al.'s strategic game in 5 to 12 years old children. Our approach is novel in analyzing the reasoning performance. We designed the reasoning task such that different items in the task distinguish between expected strategies in an optimal way. The strategies that could be expected are firstly, optimal strategies where children optimize absolute gain. Secondly, it can be expected that, as Flobbe et al. found, that some children optimize relative gain instead of absolute gain. Thirdly, young children might not master first or second-order reasoning and hence it is possible that zero-order reasoning is applied to the first-order task and first-order reasoning (that is only played if first-order reasoning is mastered) is applied to the second-order task. Finally, children could also have a position bias.

We applied the statistical technique of latent class analysis (LCA) to model the strategies from the accuracy data. LCA (McCutcheon, 1987) provides a statistically reliable method to detect strategies from response patterns [10,19]. Hence, by the application of LCA one can establish which and how many strategies are actually applied. It is not required to fully define the expected strategies beforehand. After revealing the strategies for 5 to 12 years old children, for the 5 and 6 years old children we relate the use of strategies to age, IQ, Verbal Ability, working memory, and theory of mind. These children are expected to show the most variation in theory of mind and also the executive functioning that we measure.

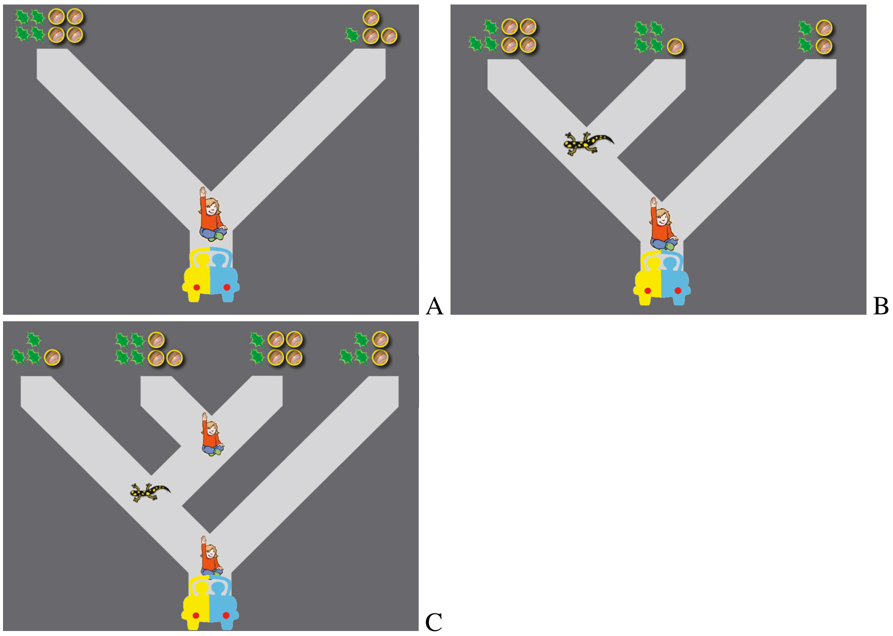

Fig. 1. Three items from the traveling game, which is based on Flobbe et al. (2008). 1A: a zero-order reasoning item, 1B: a first-order reasoning item; 1C: a second-order reasoning item. The child travels together with a lizard in a car. She/he has to acquire as many marbles as possible, but the lizard will try to gain as many leafs as possible. At each cross of the road either the child or the lizard (as is indicated) can decide where they go together, left or right. The player plays the child and the computer plays the lizard. The lizard has an optimal playing strategy, assuming the child uses an optimal strategy.

2.1 Overview

The main research question is whether inter-individual differences in playing a strategic game are due to using different strategies or to playing the same strategy with different accuracy. The next question is whether strategy use is related to the developmental notion of theory of mind, as was suggested by [8], or to other cognitive abilities, specifically working memory after correction for age, verbal abilities and general IQ.

The reasoning task we applied is a traveling game as in Flobbe et al. [6], but the appearance is somewhat different (Figure 1). There exist three types of items, which were presented in three tasks: items that require zero-order (1A), first-order (1B), and second-order reasoning (1C). In the task, the child travels together with a lizard in a car. She/he has to acquire as many marbles as possible, but the lizard will try to gain as many leafs as possible. At each cross of the road either the child or the lizard can decide where to go, left or right. The player plays the child and the computer plays the lizard, which has an optimal playing strategy, assuming an optimal strategy of the child.

2 Method

Participants were 129 children in the age range of 5 to 12 years: 23 5-years old, 26 6-years old, 16 7-years old, 14 8-years old, 15-9 years old, 10 10-years old, 18 11-years old, 7 12 years old children. Children were tested at a middle-class primary school in Amsterdam, The Netherlands

2.1 Materials

Traveling game: The strategic game is briefly explained in Figure 1. The test consisted of three tasks: a task with 2 example and 9 zero-order test items, a task with 3 example and 15 first-order test items and a task with 3 example and 9 second-order test items. All items are listed in Appendix A.

For the 5 and 6 years old children we used the following battery of cognitive tests:
IQ test: The Raven's Progressive Matrices for fluid intelligence, part A, B, and C for which we calculate a sum score.
Verbal ability test: A Dutch test for sentence comprehension, TAK (Taaltoets voor Alle Kinderen; [21]).
Working Memory test: the digit span forward and backward task.
Theory of Mind test: For the false belief test, the participants heard two second-order false belief stories, accompanied by drawings by the hand of Flobbe. The first story was the 'Birthday Puppy Story' reported in [20], a standard second-order false belief task. The second story, the 'Chocolate Bar Story', was a second-order adaptation of a first-order story by Hogrefe and Wimmer [9]. Both stories had first- and second-order questions. These test were exactly the same tests as were used in [6] experiment 1.

2.2 Procedure

Children were tested in two sessions on two different days if they completed all tests (5 and 6 years old children), otherwise they were only tested in one session. The first day they played the traveling game, the second day they completed the cognitive tests battery. The traveling game was explained and tested on a computer. Children started with the example zero-order items. The first item was used to explain the game. The child played the second item. Children responded by clicking an arrow on the road they wanted to go. In the example items children saw the animated car moving on the screen and they were presented on the screen the resulting marbles for her- or himself and the lizard was presented the leafs. If the second example item was made incorrect, the first item appeared again and the game was explained a second time. In this way, we also tested whether (the youngest) children could count. After the example items the child made the test items. Now, the child saw the animated car but did not get any direct feedback. Only after 3 items the cumulative gain was presented on the screen as a bag full of marbles.

The first-order items were explained to the children with 3 example items. Children were explained the task from the first item. Then, with the second item, which required first-order reasoning, the experimenter used instructional scaffolding to direct the child towards an optimal choice [16]. The child made the third example item, which could also be solved by a suboptimal strategy, by her-/himselves. After the example items the children made 15 test items. Again, only during the example items feedback occurred on the screen. After a choice on the test items only the first part of the animation was shown. After three items, the cumulative gain was shown on the screen.

After the first-order task, a total score for the first-order items was calculated. Only the children with 12 (out of 15) items correct, continued with the second-order items. The procedure for the second-order items was equivalent to the first-order items, again with three example items.

3 Results

All 129 children completed the experiment. However, only 55 children (43%) passed the first-order reasoning task and completed the second-order reasoning task. Their mean age is 9.78 years (sd = 1.96). Mean scores for the three tasks were above change level for the zero-order task ($t(128) = 40.1$, $p < .001$) and the first-order task ($t(128) = 10.6$, $p < .001$), but not for the second-order task ($t(54) = 1.6$, $p = .06$). Table 1 (last column) shows the scores for each task.

Strategy analysis was conducted for the three tasks separately. For each task latent class models with different number of classes were fitted to the accuracy scores of the items. The best fitting, most parsimonious model (according to the BIC, Schwartz, 1978) was selected for each task.

See table 2 for the resulting most parsimonious, best fitting models. Two strategies were found for zero-order reasoning (N = 129): The first strategy (12% of the participants) is suboptimal and has lower probability correct ($p = .62$) for the items for which the largest sum of leafs and marbles was not the optimal choice. Mean score for participants following this strategy is .66, which is above change level (sd = 12, $t(15) = 5.1$, $p < .001$; see Table 1). The second strategy (88%) is an optimal strategy.

The mean score for participants following this strategy is .98 (sd = .05). Strategies for the zero-order task were related to age (p = .002; Figure 2).

Table 1. Mean scores for the three reasoning tasks per strategy

Task	S1	S2	S3	S4	All
0-order	0.66 (0.12)	0.98 (0.05)			0.94 (0.12)
1st-order	0.58 (0.16)	0.54 (0.05)	0.51 (0.08)	0.94 (0.06)	0.70 (0.22)
2nd-order	0.42 (0.16)	0.60 (0.17)			0.54 (0.19)

Note. Columns S1 - S4 denote the mean (sd) proportion correct for the different tasks for participants responding according to strategies S1 – S4 respectively. Column All shows the mean (sd) proportion correct for all participants. Note that cell S2 for 0-order relates to different subjects than cell S2 for 1st-order task, etc. The strategies are listed in the same order as in Fig. 1.

Four strategies were found for the first-order reasoning task (N = 129): The first strategy (39%) cannot be distinguished from guessing. The second strategy (19%) is a zero-order strategy. The third strategy (5%) is to go right, which avoids a choice by the lizard. Although the third group is very small (6 children) the class does contribute to a better, parsimonious description of the data. The fourth strategy (37%) has a high probability of responding optimally for all items. The children in this group have an optimal strategy. Table 1 shows the mean scores per strategy. Strategies for the first-order task were related to age (p < .001; Figure 2).

Table 2. Resulting models from latent class analysis.

	prior	conditional probabilities			
zero-order items		type 1	type 2		
bias to sum	17%	.62	.75		
optimal	83%	.98	.99		
first-order items		type 1	type 2	type 3	
guess	39%	.62	.47	.79	
0-order	19%	.96	.04	.96	
go right	4%	.03	.97	.03	
optimal	38%	.94	.94	.94	
second-order items		type 1	type 2	type 3	type 4
guess	44%	.5	.5	.5	.5
first-order	56%	.5	1	.5	.5

Note. The estimated parameters from the best fitting, most parsimonious models of the accuracy scores of each task. The priors show the prior probability of belonging to that class (in percentages). The conditional probabilities are the probabilities of responding correctly for the corresponding item type given the strategy. For the zero-order task items 1, 4, 7, 9 are type 1, items 3, 5, 6 are type 2. For the first-order task items 1, 4, 6, 9, 15 are type 1, items 2, 5, 7, 8, 10, 12, 13 are type 2, items 3, 14 are type 3. For the second order task items 4, 7, 9 are type 1, items 2, 3 are type 2, item 6 is type 3, item 8 is type 4.

For the second-order reasoning task (N = 55) two strategies were found. The first strategy (33%) is not distinguishable from guessing. The second strategy (67%) resembles most a first-order strategy where the final choice by the child is neglected in the decision. Table 1 shows the mean scores per strategy. The mean score for participants following a kind of first-order strategy is .60, which is above chance level (sd = .17, t(36) = 3.5, p < .001). Strategies for the second-order task were related to age (p = .005). Figure 2 shows per task the distribution of strategies (in percentages) for each age group. The relation between strategy-use and age is apparent from this figure.

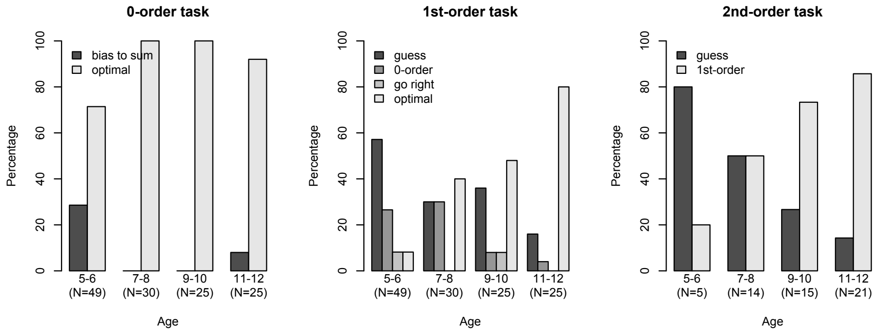

Fig 2. The distribution of strategies (in percentages) for each age group. Below the bar the number of participants within the age group is depicted. A) zero=order task, B) first-order task, C) second-order task.

Table 3. Summary data cognitive tests

Task	5 years		6 years	
	mean	sd	mean	sd
ToM	4.80	1.96	5.92	1.44
ToM1	2.65	1.18	2.92	0.74
ToM2	2.15	1.14	3.00	1.02
DS	5.90	2.05	7.42	1.58
RPM	13.30	4.50	15.88	4.23
Tak	21.85	3.69	24.77	2.39

Note. ToM is the Theory of Mind test, the sum of scores to first-order (ToM1) and second-order (ToM2) questions. DS is the Digit Span test, backwards and forwards DS summed. RPM is the Raven's Progressive Matrices, part A, B, C. Tak is the Sentence Comprehension test.

The cognitive abilities were tested only for the 5 and 6 years old children (mean age = 6.0 years, sd = 0.6). Summary statistics are shown in Table 3. For all measures we have a considerable variation, which is important to detect a relationship with strategy use. After correcting for age and verbal ability, there was a correlation between theory of mind and working memory (r = .32, p = .02) and theory of mind and IQ (r = .16, p = .005). For the zero-order task, in this age group, only age had a unique relation to strategy use and not to the other abilities that were measured (logistic regression: coeff. = .1, p = .047). For the first order task, in a logistic regression analysis with all cognitive abilities and age as predictors, only working memory had a unique relation to strategy use (coeff. = .59, p = .017) and not the other abilities or age.

4 Conclusion

Strategy analysis of playing a strategic game gives interesting insights into children's reasoning. For the zero-order tasks all children play with the same strategy, but with different accuracies. For the first- and second-order tasks children play with different strategies. The subgroups of children with different accuracies and strategies for the zero-, first and second order task were revealed by careful construction of items and with latent class analysis. On average children have high scores on the zero-order task, but nevertheless they show two types of performances. A subgroup of the children is making more mistakes and is distracted by a large amount of total gains (the sum of marbles and leafs), which results in a suboptimal choice for type 1 items. This strategy is more frequent in younger children. For the first-order task, there is one group of children with an optimal strategy. The other children have different ways of being suboptimal: guessing, location bias or zero-order reasoning. Zero-order reasoning means that the choice of the lizard is not taken into account.

For the 5 and 6 years of age children, strategy use is not related to theory of mind (in addition to age and other abilities), as was suggested in the literature, but only related to working memory. Although the age range is small, the variation in theory of mind scores is quite large. It can be questioned whether the strategic-game tasks and the theory of mind tasks have something specific in common at all. The fact that we do find a specific relation between the strategic games and working-memory task indicates that the reliability of the strategic games are large enough to find relationships with other cognitive abilities.

Finally, for the second-order task, we find one subgroup who's choices could not be distinguished from guessing. The other group seems to apply a kind of first-order strategy, combined with guessing. Although the participants in this subgroup do not use a second-order strategy, the scores of this subgroup are above chance level. This shows that from the fact that a participants have above chance performance, one cannot conclude that the participants master the task and/or the correct strategy. First, it could only be a subgroup mastering the task. Second, it could be that only a partially correct strategy was applies, which is considerably more easy. Hence, sum scores of age groups are not always very indicative for their cognitive abilities.

The overall performance for the first-order and second-order tasks is poor compared to performance on theory of mind tasks. Only 50% of the 9 and 10 years olds show a true first-order strategy, which agrees with the percentage of children that passes the criterion in [6]. However, none of the children shows a proper second order strategy. The poor performance might be due to the instruction by scaffolding instead of learning by feedback, which was used by Flobbe et al. [6]. Note that for the theory of mind tasks, instructions are mostly very limited. The reason that we have chosen for the scaffolding explanation is that we want to have optimal performance for all ages. Since learning by feedback differs importantly between 5 and 12 years of age (eg., [3]), we avoided learning by feedback. Moreover, for a strategy analysis one should test stable performance. Feedback will result in changing performance if people are not performing in an optimal way from the start (as was found by [8]). For future research it would be interesting to train children extensively on these strategic game items in an adaptive training system over a time frame of weeks, to reveal the optimal performance children gain after extensive deliberate practice [4]. An adaptive test and training system was developed as the Mathsgarden.com (rekentuin.nl; [11]). For a different complex reasoning game, static MasterMind, we see very high performance for primary school children after extensive deliberate practice on a large item bank.

There is a second possible reason why we found few children responding optimally on first and second-order reasoning items, as compared to theory of mind tasks. It is important to note that responding with a non-optimal strategy is not necessarily resulting in non-optimal choices for all items. This is not only true for the items that we designed but also for some of the items that were included in the Flobbe et al. task [6]. The result is that sum scores might end up above chance level unless children are not following a true first- or second-order strategy. Hence, strategy analysis is important for fully understanding performance on complex reasoning tasks and its development.

References

1 Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. Infant and Child Development, 11, 73–92.

2 Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. Journal of experimental child psychology, 87(4), 299-319.

3 Duijvenvoorde, A. C. K. van, Zanolie, K., Rombouts, S. a R. B., Raijmakers, M. E. J., & Crone, E. a. (2008). Evaluating the negative or valuing the positive? Neural mechanisms supporting feedback-based learning across development. The Journal of neuroscience : the official journal of the Society for Neuroscience, 28(38), 9495-9503.

4 Ericsson, K. A. (2006). The Cambridge handbook of expertise and expert performance. Chapter: The Influence of experience and Deliberate Practice on the Development of Superior Expert Performance. Cambridge University Press.

5 Flavell, J. H., Green, F. L., & Flavell, E. R. (1986). Development of knowledge about the appearance– reality distinction. Monographs of the Society for Research in Child Development, 51 (1).

6 Flobbe, L., Verbrugge, R., Hendriks, P., & Krämer, I. (2008). Children's Application of Theory of Mind in Reasoning and Language. Journal of Logic, Language and Information, 17(4), 417-442.

7 Gopnik, A., & Astington, J. W. (1988). Children_s understanding of representational change and its relation to the understanding of false belief and the appearance–reality distinction. Child Development, 59, 26–37.

8 Hedden, T., & Zhang, J. (2002). What do you think I think you think? Strategic reasoning in matrix games. Cognition, 85, 1–36.

9 Hogrefe, G., & Wimmer, H. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. Child Development, 57, 567.

10 Jansen, B.R.J. and Maas, H.L.J. van der (1997). Statistical Test of the Rule Assessment Methodology by Latent Class Analysis. Developmental Review, 17, 321–357.

11 Klinkenberg, S., Straatemeier, M., & Maas, H. L. J. van der. (2011). Computer adaptive practice of Maths ability using a new item response model for on the fly ability and difficulty estimation. Computers & Education, 1-12.

12 McCutcheon, A. L. (1987). Latent class analysis. Newbury Park: Sage.

13 Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological bulletin, 135*, 749-773.

14 Muris, P., Steerneman, P., Meesters, C., Merckelbach, H., Horselenberg, R., van den Hogen, T., & van Dongen, L. (1999). The TOM test: A new instrument for assessing theory of mind in normal children and children with pervasive developmental disorders. Journal of Autism and Developmental Disorders, 29, 67–80.

15 Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds_ difficulty understanding false belief: Representational limitation, lack of knowledge, or pragmatic misunderstanding? British Journal of Developmental Psychology, 5, 125–137.

16 Rodgers, E. M. (2004). Interactions that scaffold reading performance. Journal of Literacy Research, 36(4), 501-532

17 Schwarz, G. (1978). Estimating the dimension of a model. The Annals of Statistics, 6(2), 461-464.

18 Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. Cognitive Psychology, 28, 225-273.

19 Straatemeier, M., Van der Maas, H.L.J., & Jansen, B.R.J. (2008). Children's Knowledge of the Earth: A New Methodological and Statistical Approach. Journal of Experimental Child Psychology, 100, 276–296.

20 Tager-Flusberg, H., & Sullivan, K. (1994). A second look at second-order belief attribution in autism. Journal of Autism and Developmental Disorders, 24, 577–586.

21 Verhoeven, L., & Vermeer, A. (2006). Verantwoording Taaltoets Alle Kinderen (TAK). Arnhem: Centraal Instituut voor Toetsontwikkeling.

22 Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. Child Development, 72, 655-684.

23 Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition, 13, 103–128.

Appendix A

In Appendix A, all items are enumerated (Tables A1, A3, and A5) and the expected accuracy patterns according to different strategies (Tables A2, A4, and A6). In the latent class analysis not all items are used (see note Table 2), because these items did not correlate well with the items of the same type for unclear reasons. Items of the same type have the same expected scores for all strategies.

Table A1. Zero-order items

Items	B1		B2		Optimal Response
	L	M	L	M	
A	3	3	2	1	L
B	3	1	2	4	R
1	3	3	1	4	R
2	4	4	1	3	L
3	1	2	3	3	R
4	1	4	4	2	L
5	4	3	1	2	L
6	1	3	3	4	R
7	1	3	4	2	L
8	4	3	2	4	R
9	3	3	1	4	R

Note. Items A and B are example items, 1 – 9 are test items. Items are coded by an enumeration of leafs (L) and marbles (M) from the left branch (B1) to the right branch (B2). See Figure 1 for the configuration of leafs and marbles. The optimal choice is left (L) or right (R).

Table A2. Expected accuracy patterns for different potential strategies

	Strategies		
Items	0-A	0-B	0-C
1, 8, 9	1	0	1
4, 7	1	0	1
2, 5	1	1	0
3, 6	1	1	0

Note. The potential strategies are 0-A, the optimal strategy, 0-B, the choice for largest sum of leafs and marbles, 0-C, the choice for largest relative gain. 1 is correct, 0 is incorrect

Table A4. Expected accuracy patterns for different potential strategies

	Strategies					
Items	1-A	1-B	0-A	0-B	0-C	0-D
1, 4, 6, 9, 15, 11	1	0	1	1	0	0
2, 5, 7, 8, 10, 12, 13	1	1	0	0	0	1
3, 14	1	1	1	1	1	0

Note. The potential strategies are 1-A, the optimal strategy, 1-B, the choice for largest relative gain, 0-A, a zero-order strategy with largest gain, 0-B, zero-order strategy with largest sum of leafs and marbles, 0-C zero-order strategy with largest relative gain, 0-D, go to the right. 1 is correct, 0 is incorrect.

Table A5. Second-order items

Item	B1		B2		B3		B4		Optimal
	L	M	L	M	L	M	L	M	Response
A	2	3	4	4	3	1	1	2	L
B	2	2	1	4	4	1	1	3	R
C	4	3	1	1	2	2	2	1	L
1	3	3	4	1	2	4	2	2	L
2	3	2	2	4	4	1	2	3	R
3	3	2	2	4	4	1	2	3	R
4	1	3	2	4	4	2	3	3	L
5	3	1	4	3	1	4	1	2	R
6	3	1	1	4	4	3	2	2	R
7	2	3	1	3	4	1	2	2	L
8	3	1	4	3	2	4	3	2	R
9	1	2	4	2	2	4	2	3	L

Note. See note Table A1

Table A6. Expected accuracy patterns for different potential strategies

	Strategies		
Items	2-A	1-A	1-B
1, 9,4,7	1	1	0
2, 3	1	0	1
5, 6, 8	1	0	0

Note. The potential strategies are 2-A, the optimal strategy, 1-A, a first-order strategy with a second choice for the child, 1-B, a first-order strategy without a second choice for the child. 1 is correct, 0 is incorrect.