Reasoning about self and others

Meijering, Ben

Publication date:
2014

Chapter 7

Summary and discussion

Discussion

This dissertation details an investigation of higher-order theory of mind in adults. It is high time to study higher-order theory of mind, as it has not yet received as much attention as first-order theory of mind (ToM). Nevertheless, higher-order ToM is not some exotic cognitive function, and people need it to engage in complex social interactions. Communication, for example, may already require higher-order ToM if the wording is ambiguous, which is not that uncommon in language. To find the most probable meaning of an ambiguous utterance, the listener has to reason about the speaker's beliefs, and account for the fact that the speaker in turn may have reasoned about the listener's knowledge.

The study of ToM in adults also needs more attention, as ToM has mostly been studied in infants and children. That is not so surprising as first-order ToM develops around the age of 4, and second-order ToM between 6 and 8. Nevertheless, adults still frequently fail to account for the mental states of others, and it is not yet evident why that happens. For example, do adults not *have* a complete theory of mind, or do they not have sufficient cognitive resources to *apply* it? This dissertation provides new insights into why adults may not always apply ToM despite the fact that they have already mastered it.

In contrast to many other studies, theory of mind was investigated by means of two-player games, instead of false-belief tasks. One obvious advantage of these games is that they do not require language processing as much as the stories in false-belief tasks. Another advantage is that games can be presented many times, in various configurations, to the same participants. One concern has been that some games do not strictly require mental state reasoning. However, Chapters 3 and 4 have demonstrated that people did interpret the two-player games in terms of mental states. The study in Chapter 4, for example, shows that people do not consider a rational computer player to be the same as a completely deterministic device, even though the outcome of both was based on the same principle. In sum, the paradigm of two-player games has proven to be successful in examining various characteristics of theory of mind.

Cognitive constraints

Some studies suggest that the development of ToM in children involves a conceptual change (Gopnik & Slaughter, 1991; Gopnik & Wellman, 1992; 2000; Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). According to this account, children first experience desires and perceptions as simple causal links between them and the world, at age 2. Later, children learn about beliefs, and even think of these as representations, at age 3. However, they cannot yet incorporate belief representations into their 'core' theory of mind, which still reflects are rather direct causal link between them and the world. At the age of 4, they learn to perceive that representations (of mental states) are the basis of psychological function.

An important prediction of this account is that if development of ToM would involve a conceptual change, ToM would not be susceptible to improvement when it has already reached maturity. The findings in this dissertation (see Chapter 2), however, imply that application of ToM is a computational process that can benefit from supporting structure. Application of higher-order ToM improved when adults were trained to account for increasingly more complex mental states, when they were prompted to take another's perspective, and when they were provided with visual cues as to how decisions are dependent on mental states. These findings show the ability to apply higher-order ToM is susceptible to improvement, and thus ToM may not be a fixed skill after all. However, adults and children do need sufficient cognitive resources to put that ability into practice.

The study in Chapter 3 shows that the application of ToM may involve a specialized cognitive function to infer the mental states of self and others. Previously, developmental studies suggested that unsuccessful inference of mental states reflects a broader problem with representations in general (Leekam, Perner, Healey, & Sewell, 2006; Perner & Leekam, 2008; Sabbagh, Xu, Carlson, Moses, & Kang, 2006; Todd & Gigerenzer, 2000). The findings in Chapter 3, however, show that reasoning about someone else's decision-making is more difficult than making the same decision oneself, even though the required reasoning steps are the same. Apparently, switching perspective makes the decision problem more difficult. One possible explanation is that the representation of a particular decision-making problem becomes more elaborate as the complexity of the involved mental states increases.

The findings in Chapter 4 corroborate the conclusions of the study in Chapter 3. Participants were presented with two-player games in which they had to reason about another player. The findings show that performance depended on whether the other player was reasoning about the participant's decision or, instead, about a mechanism. We hypothesized that the other player's mental states would be easier to infer if he would be reasoning about a mechanism, because the mechanism was completely deterministic. In contrast, if the other player would be reasoning about the participant, the participant would have to reason about a multitude of ideas the other player could be having about her. Importantly, both conditions were completely isomorphic with respect to the required reasoning steps. Still, the results showed that the response times were shorter in the mechanism games. We argue that the non-mechanism games were more difficult to solve because people had to test many possible mental state interpretations. In the mechanism games, in contrast, the other person's actions were dependent on a deterministic mechanism and people could therefore test fewer possible mental state interpretations and respond faster. Besides differential response times, the types of errors qualitatively differed between mechanism and non-mechanism games.

In sum, the studies in Chapters 3 and 4 showed that the complexity of mental states, all other task aspects controlled for, caused differential cognitive processes. The more intricate the involved mental states were, the worse the performance was, which suggests that application of ToM consumed cognitive resources. The studies in later chapters show how people try to preserve cognitive resources and still perform well.

Cognitive processes

As application of ToM and especially higher-order ToM are considered to be effortful processes (see Chapters 2, 3, and 4), it is not surprising that people use simple strategies to reduce demands on cognitive resources. Todd and Gigerenzer (Meijering, Van Rijn, Taatgen, & Verbrugge, 2012; Szymanik, Meijering, & Verbrugge, 2013; see also Chapter 5 in this dissertation; Todd & Gigerenzer, 2000) already argued that people use simple strategies or heuristics to solve many (non-social) tasks. Chapters 5 and 6 show that this may be true for inference of mental states as well: People start out reasoning about simple mental states, using basic strategies, and only account for more sophisticated mental states if their simple strategies do not yield desirable outcomes anymore.

In the study in Chapter 5, we tracked people's eye movements during a two-player game in which they had to infer the other player's mental states. The eye movements were analyzed for the presence of patterns, or eye movement sequences, that would indicate in what order people tend to construct a representation of recursive mental states. The most efficient strategy across all games would have been to construct these recursive mental states in a backward fashion, as each mental state depended on the next one. However, that strategy requires a deep understanding of the task domain, and most people tend to use more simple strategies that work across multiple domains (e.g., Gopnik et al., 2004; Todd & Gigerenzer, 2000). The eye movements indeed indicated that people inferred mental states in a more simple and forward progression, only tracking backward if previous (higher-order) mental states had to be revised.

The prevalence of a simple and forward approach can be explained by the principle of economy: Immediate decisions that ignore future possibilities *can* yield the optimal outcome in many cases (2008; Meijering et al., 2012; Szymanik et al., 2013). Furthermore, forward reasoning receives much practice across many domains, for example in causal inference (Apperly & Butterfill, 2009; Baron-Cohen, Leslie, & Frith, 1985; Gopnik et al., 2004; Gopnik & Wellman, 1992; Leslie, Friedman, & German, 2004; Onishi & Baillargeon, 2005; Premack & Woodruff, 1978; Saxe, Schulz, & Jiang, 2006; Wimmer & Perner, 1983). Thus, it is not surprising that forward reasoning, in its most simple form, is also used to construct representations of recursive mental states.

Chapter 6 provides a computational cognitive account of the use of simple strategies during inference of mental states. The model is based on previous empirical findings (reported in Chapters 2 and 5), and is validated by data from Flobbe et al.'s (but see Baker, Saxe, & Tenenbaum, 2009; Flobbe et al., 2008) developmental study. The model uses a simple strategy at first and only starts incorporating more complex mental states in the face of unexpected outcomes. The simple strategy is comparable to forward reasoning, which is later followed by backtracking, as the model starts considering future possible actions and underlying mental states. Investigating Flobbe et al.'s data, we saw response patterns that were indeed indicative of simple strategies. Few children were able to account for the fact that another person could be reasoning about them.

In sum, the findings in Chapters 5 and 6 corroborate the claim we made in earlier chapters: Application of (higher-order) ToM is a computational process that can either benefit from supporting structure, or be simplified by using simple strategies, thereby reducing cognitive demands. Given these findings, what new insights can future ToM research bring?

Looking into the future

The studies in this dissertation show that application of (higher-order) ToM is a complicated task. The fact that higher-order ToM consists of multiple procedural and declarative building blocks almost poses something like an inverse problem in the sense that there are many possible sources that could yield the same behavioral patterns. Therefore, a multidisciplinary approach to the investigation of ToM is desirable.

So far, ToM has mostly been studied in clinical settings, developmental studies, animal studies, and imaging studies, which all have produced many interesting insights and theories (Apperly & Butterfill, 2009; 2013; Baron-Cohen et al., 1985; Gopnik & Wellman, 1992; Leslie et al., 2004; Onishi & Baillargeon, 2005; Premack & Woodruff, 1978; Saxe et al., 2006; Wimmer & Perner, 1983). Most theories, however, exist only on paper, and do not directly translate to quantifiable predictions (but see Baker et al., 2009). This is why computational cognitive modeling needs to be employed more often, as a way of testing existing and new theories: Once implemented, a theory can yield quantifiable predictions that can be directly tested in one or more experiments.

This dissertation is a modest step towards a cognitive modeling approach, as we have yet to validate our model in many more domains. The model could be used to accommodate other ToM paradigms, and it could be used to generate hypotheses for clinical, developmental, imaging and other psychophysiological studies. The model could, for example, simulate application of ToM in higher-order false-belief tasks, in which one agent (e.g., Sally) has a false belief about another agent's (e.g., Anne's) beliefs. At first, the model would do the task from its own perspective, and later the model would attribute its knowledge to the other agent, Sally. Lastly, the model would reason about Sally as if she would attribute her own knowledge to yet again another agent, Anne. Scenarios such as these are currently being tested empirically by Arslan et al. (2013). The most obvious prediction is that children of 4 years old who have not yet mastered second-order ToM, but who do have first-order ToM, would not fall prey to the reality bias. Instead they would apply first-order ToM when, in actuality, they are asked to make a second-order inference.

It would also be interesting to look at transfer of ToM between various ToM domains by having one model play, for example, Marble Drop games and do higher-order false-belief tasks. Would experience in one task be beneficial in the other? How much overlap is there between the tasks with respect to demands on cognitive functions? Questions like these have recently been investigated by Taatgen (2013) in other (non-social) domains. His modeling approach has culminated in Actransfer, an extension of ACT-R (Anderson, 2007; Anderson et al., 2004), which is a theory about the nature and transfer of cognitive skills. Actransfer is particularly relevant to investigate the domain-specificity of ToM, which is still hotly debated: Does ToM require a specialized cognitive function, or does it involve general cognitive skills that are used across multiple domains? The empirical results in Chapter 3 suggest that ToM requires a specialized cognitive function, but it provides pointers to many possible explanations. Actransfer could help in finding the most primitive elements that comprise application of ToM.

Testing our model in various domains may not only yield insights that help us improve the model. Some insights will help improve the cognitive architecture (e.g., ACT-R) as a whole. If, for example, application of ToM indeed requires a specialized cognitive function to infer mental states of self and others, the architecture needs to accommodate such a module. Therefore, the study of ToM is particularly interesting for the entire cognitive sciences, as it will help constructing an integrated theory of cognition.

Samenvatting

Het onderzoek in dit proefschrift gaat over hoe mensen redeneren over andermans denken; wat hun gedachten, intenties en doelen zijn. Dit onderzoek is belangrijk voor het vormen van een geïntegreerde theorie van cognitie, omdat de vraag is of redeneren over andermans denken een speciale cognitieve functie is. Zowel het onderzoek als het redeneren over andermans denken is ingewikkeld om één heel duidelijke reden: redeneren en denken zijn onzichtbaar processen. Toch zijn mensen verassend goed in het redeneren over andermans denken; we kunnen met enige zekerheid voorspellen op welke politieke partij vrienden en familieleden zullen stemmen, omdat we weten hoe zij denken, wat hun opvattingen zijn, et cetera. Maar er zijn limieten aan het redeneren over andermans denken. Met name bij jonge kinderen wordt dat snel duidelijk, bijvoorbeeld als ze niet begrijpen dat we ze nog steeds kunnen zien als ze hun handen voor hun ogen houden. Jonge kinderen vinden het moeilijk om zich in een ander te verplaatsen. Volwassenen hebben daar minder moeite mee, maar toch wordt tijdens een spel schaak (of poker) al snel duidelijk dat het ondoenlijk is om alle gedachten van de tegenspeler te anticiperen. Het is helemaal moeilijk om te anticiperen welke gedachten de tegenspeler heeft over onze gedachten. Dit recursieve denkproces, dat in theorie oneindig is, houdt in de praktijk al snel op. In dit proefschrift laat ik zien waarom het redeneren over andermans denken, wat ik vanaf nu meta-denken zal noemen, limieten heeft.

Onderzoek naar meta-denken

Meta-denken is in het verleden voornamelijk bij kinderen onderzocht, met als achterliggend idee dat de ontwikkeling van een cognitieve functie iets kan vertellen over de uiteindelijke aard van die functie in volwassenen. Het meest populaire experiment om meta-denken bij kinderen te onderzoeken is de Sally-Anne taak. Het gaat als volgt: Sally en Anne spelen met knikkers. Als Sally eventjes weggaat, bergt ze eerst haar knikkers op in haar mandje. Terwijl Sally weg is, pakt Anne de knikkers en verstopt ze die in de speelgoeddoos. Na een tijdje komt Sally terug en de vraag aan kinderen is: Waar zal Sally naar haar knikkers zoeken? Kinderen die het meta-denken onder de knie hebben, weten dat Sally nog steeds denkt dat de knikkers in haar mandje zijn en dat ze daar zal zoeken, ook al zijn de knikkers in werkelijkheid in de speelgoeddoos. Kinderen die het meta-denken nog niet beheersen, zullen zeggen dat Sally naar de knikkers in de speelgoeddoos zal zoeken. Ze kunnen feiten en gedachten nog niet van elkaar onderscheiden.

Tot op heden is nog steeds niet duidelijk of dat komt omdat ze geen begrip hebben van mentale toestanden zoals kennis, gedachten en intenties of omdat ze nog niet de cognitieve vaardigheden hebben om over die mentale toestanden na te denken. Om dit probleem te ondervangen, heb ik onderzoek gedaan bij volwassenen, want die hebben een groter besef van onzichtbare denkprocessen. De algemene cognitieve vaardigheden zijn ook beter ontwikkeld bij volwassen en dus stelt onderzoek bij hen ons beter in staat om vast te stellen in hoeverre meta-denken een speciale cognitieve functie betreft.

Speciaal voor dit onderzoek heb ik een knikkerspel ontwikkeld waarin twee spelers om beurten een beslissing maken over het verloop van het spel. Het spel doet een beroep op het meta-denken, omdat de uitkomsten van elke beslissing afhankelijk zijn van de volgende beslissing. De ene speler moet dus nadenken over het beslisproces en de onderliggende gedachten van de andere speler.

De aard van het meta-denken

Om na te gaan of het meta-denken wordt gelimiteerd door cognitieve vaardigheden, heb ik onderzocht of training en andersoortige ondersteuning een gunstig effect hebben op het meta-denken. De uitkomsten van dit onderzoek zijn positief: Mensen presteerden beter als ze het meta-denken in een bepaalde taak stapsgewijs oefenden: Ze leerden eerst de taak vanuit het eigen perspectief te doen en later dat perspectief aan een ander toe te schrijven. Ook presteerden mensen beter als hen expliciet werd gevraagd om zich te verplaatsen in het perspectief van de ander. Deze resultaten laten zien dat suboptimale uitkomsten in een sociale interactie niet zozeer zijn toe te schrijven aan onbegrip als wel aan gebrek aan oefening. In die zin is het meta-denken dus een cognitieve vaardigheid die je kunt oefenen en niet perse een vaardigheid die men wel of niet beheerst.

De belangrijkste vraag in dit proefschrift is of meta-denken een speciale cognitieve vaardigheid is, in tegenstelling tot een vaardigheid die is opgebouwd uit meer algemene vaardigheden. De resultaten van een aantal ontwikkelingsstudies in 2006 doen vermoeden dat meta-denken niet speciaal is en bestaat uit algemene vaardigheden. Deze studies hebben laten zien dat kinderen die slecht scoorden op de Sally-Anne taak ook moeite hadden met redeneren in het algemeen. Echter, in dit proefschrift laat ik zien dat dergelijke conclusies niet zijn te trekken op basis van ontwikkelingsstudies. Het experiment in Hoofdstuk 3 laat duidelijk zien dat het meta-denken wel degelijk een speciale cognitieve vaardigheid is: volwassenen werden blootgesteld aan twee condities met als enige verschil de instructie om de taak vanuit het eigen perspectief of dat van een ander te doen. De resultaten laten zien dat zowel de kwaliteit als de snelheid van het meta-denken verschilde tussen de twee condities. Als de instructie was om de taak vanuit het eigen perspectief te doen, presteerden de volwassen sneller en beter dan wanneer de instructie was om dezelfde taak vanuit het perspectief van een ander te doen. Deze resultaten impliceren dat de aard van het menselijk redeneren afhankelijk is van het feit of een taak al dan niet over mentale toestanden gaat, los van de complexiteit van de taak.

In dit proefschrift laat ik ook zien dat in sociale interacties mensen de voorkeur geven aan simpele strategieën en pas beginnen met meta-denken als het echt niet anders kan. In een eerdere ontwikkelingsstudie werd kinderen gevraagd een computerspel te spelen, met de computer als tegenstander. De uitkomsten van die studie deden vermoeden dat de kinderen het meta-denken al enigszins beheersten. Maar door middel van computersimulaties laat ik zien dat het waarschijnlijker is dat de kinderen simpelere strategieën gebruikten als die ook tot juiste uitkomst leidden. Pas als duidelijk werd dat die strategieën niet altijd werkten, probeerden de kinderen het perspectief van de tegenspeler mee te nemen door hun eigen strategie aan de tegenspeler toe te schrijven. Zodoende ontwikkelden zij langzamerhand steeds complexere strategieën die steeds meer op meta-denken gingen lijken.

De bevindingen van dit proefschrift zijn belangrijk voor het formuleren van een geïntegreerde theorie van cognitie, omdat het meta-denken daarin een rol dient te hebben. Het is een speciale cognitieve functie die het denken in kwalitatieve zin beïnvloed. Daarnaast zijn de bevindingen in dit proefschrift van belang voor de praktijk, onder andere omdat het meta-denken is te trainen. Door het bieden van de juiste structuur kunnen mensen leren om optimalere uitkomsten te behalen in sociale interacties zoals onderhandelingen, samenwerkingsverbanden en competities.

Dankwoord

Toen ik pas begonnen was als PhD student en ik Rineke nog niet zo goed kende, vertelden andere PhD studenten me: 'Oh, she is so great!' en 'You are so lucky!' Inmiddels kan ik bevestigen dat dat waar is. Ik heb geluk gehad: behalve dat Rineke een inspirerende onderzoeker is, stond ze ook nog eens altijd voor me klaar.

Ik heb ook geluk gehad met mijn andere begeleiders, Hedderik en Niels. Van hen heb ik geleerd dat 'onderzoek doen' net zo gaaf kan zijn als detectivewerk.

In de tijd bij AI maakte ik deel uit van twee onderzoeksgroepen, de Vici Reading Group en de Cognitive Modeling Group. Ik wil huidige en voormalige deelnemers van beide groepen bedanken. De discussies in de Vici Reading Group waren bijzonder vermakelijk; wat konden we kritisch zijn op andermans papers :). Het was ook erg leerzaam om elkaars werk en presentaties te zien, in beide groepen.

Louter door organisatorische invloeden heb ik diverse kamergenoten gehad: Jelmer, Leendert, Enkhbold, Sujata, Jakub en Burcu. Ik wil Jelmer en Leendert bedanken voor hun inzichtelijke kijk op het wetenschappelijk spel; Enkhbold voor z'n gedeelde fascinatie voor (computer)spellen; Sujata en Jakub voor hun mooie verhalen over logica en logici; Burcu voor haar uitstekende richtingsgevoel in Berlijn.

Tussen de onderzoeksbedrijven door was er gelukkig vaak tijd voor ontspanning en afleiding, bijvoorbeeld tijdens koffiepauzes, waar we voor menig niet-bestaand probleem een oplossing hebben gevonden. Aswin, Bea, Burcu, Charlotte, Gyuhee, Harmen, Ioanna, Jean-Paul, Marijke, Michiel, Olarik en Trudy waren vaak van de partij, bedankt!

Hereby, I would like to express my gratitude to professors Petra Hendriks, Josef Perner, and Maartje Raijmakers for reading my dissertation. I'm proud that I can say that these great researchers have seen my work.

Ik wil ook graag mijn paranimfen bedanken, Jean-Paul en Durk. Ik ben blij dat twee goeie vrienden me bijstaan in het ceremoniële geweld van de verdediging. Beide beschikken namelijk ruimschoots over 'theory of mind'.

Tenslotte, pa, ma, Rob, Gonny, Charlotte: ik heb het gehaald!

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY: Oxford University Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036–1060.

Apperly, I. A. (2011). *Mindreaders: The cognitive basis of "theory of mind."* Hove, UK: Psychology Press.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*(4), 953–970.

Apperly, I. A., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: Evidence from adults' performance on a non-inferential theory of mind task. *Cognition, 106*(3), 1093–1108.

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffit, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *The Quarterly Journal of Experimental Psychology, 63*(6), 1201–1217.

Arslan, B., Hohenberger, A., & Verbrugge, R. (2012). The development of second-order social cognition and its relation with complex language understanding and memory. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 1290–1295). Austin, TX: Cognitive Science Society.

Arslan, B., Taatgen, N. A., & Verbrugge, R. (2013). Modeling developmental transitions in reasoning about false beliefs of others. In R. L. West, & T. C. Stewart (Eds.), *Proceedings of the 12th International Conference on Cognitive Modelling* (pp. 77–82). Ottawa, CA: Carleton University.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York, USA: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412.

Bacharach, M., & Stahl, D. O. (2000). Variable-frame level-n theory. *Games and Economic Behavior, 32*(2), 220–246.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*(3), 110–118.

Baker, C., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition, 113*(3), 329–349.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind?" *Cognition, 21*(1), 37–46.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.

Birch, S. A. J., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences, 8*(6), 255–260.

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*(5), 382–386.

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*(1), 25–31.

Borst, J. P., Taatgen, N. A., & Van Rijn, H. (2010a). The problem state: A cognitive bottleneck in multitasking. *Journal of Experimental Psychology: Learning, memory, and cognition, 36*(2), 363–382.

Borst, J. P., Taatgen, N. A., Stocco, A., & Van Rijn, H. (2010b). The neural correlates of problem states: Testing fMRI predictions of a computational model of multitasking. *PLoS ONE, 5*(9).

Brassard, G., & Bratley, P. (1996). *Fundamentals of algorithmics*. Englewood Cliffs, N.J.: Prentice Hall.

Bull, R., Phillips, L. H., & Conway, C. A. (2008). The role of control functions in mentalizing: Dual-task studies of theory of mind and executive function. *Cognition, 107*(2), 663–672.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*(5), 187–192.

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*(2), 73–92.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development, 9*(4), 377–395.

Cohen, A. S., & German, T. P. (2010). A reaction time advantage for calculating beliefs over public representations signals domain specificity for "theory of mind." *Cognition, 115*(3), 417–425.

Colman, A. M. (2003). Depth of strategic reasoning in games. *Trends in Cognitive Sciences, 7*(1), 2–4.

Coull, G. J., Leekam, S. R., & Bennett, M. (2006). Simplifying second-order belief attribution: What facilitates children's performance on measures of conceptual understanding? *Social Development, 15*(2), 260–275.

De Villiers, J. (2007). The interface of language and theory of mind. *Lingua, 117*(11), 1858–1878.

De Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development, 17*(1), 1037–1060.

De Weerd, H., Verbrugge, R., & Verheij, B. (2013). How much does it help to know what she knows you know? An agent-based simulation study. *Artificial Intelligence, 199-200*, 67–92.

Dumontheil, I., Apperly, I. A., & Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescense. *Developmental Science, 13*(2), 331–338.

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.

Fu, W., & Gray, W. D. (2004). Resolving the paradox of the active user: Stable suboptimal performance in interactive tasks. *Cognitive Science, 28*(1), 901–935.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of "theory of mind." *Trends in Cognitive Sciences, 7*(2), 77–83.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

German, T. P., & Hehman, J. A. (2006). Representational and executive selection resources in "theory of mind": Evidence from compromised belief-desire reasoning in old age. *Cognition, 101*(1), 129–152.

Ghosh, S., & Meijering, B. (2011). On combining cognitive and formal modeling: A case study involving strategic reasoning. In J. Van Eijck & R. Verbrugge (Eds.), *Proceedings of the Workshop on Reasoning About Other Minds: Logical and Cognitive Perspectives* (pp. 79–92). Groningen, NL: University of Groningen.

Ghosh, S., Meijering, B., & Verbrugge, R. (2010). Empirical reasoning in games: Logic meets cognition. In T. Ågotnes, N. Alechina, & B. Logan (Eds.), *Proceedings Third Logics for Resource Bounded Agents Workshop* (pp. 15–34). Lyon, FR.

Goodie, A. S., Doshi, P., & Young, D. L. (2012). Levels of theory-of-mind reasoning in competitive games. *Journal of Behavioral Decision Making, 25*(1), 95–108.

Gopnik, A., & Slaughter, V. (1991). Young children's understanding of changes in their mental states. *Child Development, 62*(1), 98–110.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language, 7*(1-2), 145–171.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111*(1), 3–32.

Gray, W. D., Sims, C. R., Fu, W., & Schoelles, M. J. (2006). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review, 113*(3), 461–482.

Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science, 6*(3), 346–359.

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

Hiatt, L. M., & Trafton, J. G. (2010). A cognitive model of theory of mind. In D. Salvucci & G. Gunzelmann (Eds.), *Proceedings of the 10th International Conference on Cognitive Modeling* (pp. 91–96). Philadelphia, PA: Drexel University.

Hollebrandse, B., Hobbs, K., De Villiers, J., & Roeper, T. (2008). Second order embedding and second order false belief. In A. Gavarro & M. J. Freitas (Eds.), *Language Acquisition and Development: Proceedings of GALA 2007*. Cambridge Scholar Press.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65–70.

Jansen, B. R. J., & Van der Maas, H. L. J. (2002). The development of children's rule use on the balance scale task. *Journal of Experimental Child Psychology, 81*(4), 383–416.

Johnson, E. J., Camerer, C., Sen, S., & Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory, 104*(1), 16–47.

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

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. Hoboken, N.J.: John Wiley & Sons, Inc.

Kelley, H. H., & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology, 31*(1), 457–501.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition, 89*(1), 25–41.

Kong, X., Schunn, C. D., & Wallstrom, G. L. (2010). High regularities in eye movement patterns reveal the dynamics of the visual working memory allocation mechanism. *Cognitive Science, 34*(2), 322–337.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330*(6012), 1830–1834.

Leekam, S. R., Perner, J., Healey, L., & Sewell, C. (2006). False signs and the non-specificity of theory of mind: Evidence that preschoolers have general difficulties in understanding representations. *British Journal of Developmental Psychology, 26*(4), 485–497.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind." *Trends in Cognitive Sciences, 8*(12), 528–533.

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology, 46*(3), 551–556.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4*(1), 6–14.

McColgan, K. L., & McCormack, T. (2008). Searching and planning: Young children's reasoning about past and future event sequences. *Child Development, 79*(5), 1477–1497.

McKelvey, R. D., & Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica, 60*(4), 803–836.

Meijering, B., Van Maanen, L., Van Rijn, H., & Verbrugge, R. (2010). The facilitative effect of context on second-order social reasoning. In R. Catrambone & S. Ohlsson (Eds.), *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (pp. 1423–1428). Austin, TX: Cognitive Science Society.

Meijering, B., Van Rijn, H., Taatgen, N. A., & Verbrugge, R. (2012). What eye movements can tell about theory of mind in a strategic game. *PLoS ONE, 7*(9).

Meijering, B., Van Rijn, H., Taatgen, N. A., & Verbrugge, R. (2013). Reasoning about diamonds, gravity, and mental states: The cognitive costs of theory of mind. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 3026–3031). Austin, TX: Cognitive Science Society.

Meijering, B., Van Rijn, H., Taatgen, N., & Verbrugge, R. (2011). I do know what you think I think: Second-order theory of mind in strategic games is not that difficult. In L. Carslon, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2486–2491). Austin, TX: Cognitive Science Society.

Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological Bulletin, 135*(5), 749–773.

Nyamsuren, E., & Taatgen, N. A. (2013). Set as an instance of a real-world visual-cognitive task. *Cognitive Science, 37*(1), 146–175.

O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development, 67*(2), 659–677.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255–258.

Osborne, M. J., & Rubinstein, A. (1994). *A course in game theory*. Cambridge, MA: MIT press.

Perner, J., & Leekam, S. R. (2008). The curious incident of the photo that was accused of being false: Issues of domain specificity in development, autism, and brain imaging. *The Quarterly Journal of Experimental Psychology, 61*(1), 76–89.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that..." Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39*(3), 437–471.

Pinheiro, J. C., & Bates, D. M. (2000). Mixed-effects models in S and S-PLUS. Springer Verlag.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515–526.

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition, 117*(2), 230–236.

Raijmakers, M. E. J., Mandell, D. J., Van Es, S. E., & Counihan, M. (2013). Children's strategy use when playing strategic games. *Synthese*.

Ramsey, R., Hansen, P. C., Apperly, I. A., & Samson, D. (2013). Seeing it my way or your way: Frontoparietal brain areas sustain viewpoint-independent perspective selection processes. *Journal of Cognitive Neuroscience, 25*(5), 670–684.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology, 80*(3), 201–224.

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Kang, L. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science, 17*(1), 74–81.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance, 36*(5), 1255–1266.

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: A case of a selective deficit in inhibiting self-perspective. *Brain, 128*(5), 1102–1111.

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology, 16*(2), 235–239.

Saxe, R., Schulz, L. E., & Jiang, Y. V. (2006). Reading minds versus following rules: Dissociating theory of mind and executive control in the brain. *Social Neuroscience, 1*(3-4), 284–298.

Suddendorf, T., Nielsen, M., & Gehlen, von, R. (2011). Children's capacity to remember a novel problem and to secure its future solution. *Developmental Science, 14*(1), 26–33.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology, 30*(3), 395–402.

Szymanik, J., Meijering, B., & Verbrugge, R. (2013). Using intrinsic complexity of turn-taking games to predict participants' reaction times. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 1426–1432). Austin, TX: Cognitive Science Society.

Taatgen, N. A. (2013). The nature and transfer of cognitive skills. *Psychological Review, 120*(3), 439–471.

Thoermer, C., Sodian, B., Vuori, M., Perst, H., & Kristen, S. (2012). Continuity from an implicit to an explicit understanding of false belief from infancy to preschool age. *British Journal of Developmental Psychology, 30*(1), 172–187.

Todd, P. M., & Gigerenzer, G. (2000). Précis of simple heuristics that make us smart. *Behavioral and Brain Sciences, 23*(1), 727–780.

Van Maanen, L., & Verbrugge, R. (2010). A computational model of second-order social reasoning. In D. Salvucci & G. Gunzelmann (Eds.), *Proceedings of the 10th International Conference on Cognitive Modeling* (pp. 259–264). Philadelphia, PA: Drexel University.

Van Rij, J., Van Rijn, H., & Hendriks, P. (2010). Cognitive architectures and language acquisition: A case study in pronoun comprehension. *Journal of Child Language, 37*(3), 731–766.

Van Rij, J., Van Rijn, H., & Hendriks, P. (2013). How WM load influences linguistic processing in adults: A computational model of pronoun interpretation in discourse. *Topics in Cognitive Science, 5*(3), 564–580.

Van Rijn, H., Van Someren, M., & Van der Maas, H. L. J. (2003). Modeling developmental transitions on the balance scale task. *Cognitive Science, 27*(2), 227–257.

Verbrugge, R., & Mol, L. (2008). Learning to apply theory of mind. *Journal of Logic, Language and Information, 17*(4), 489–511.

Weber, R., Camerer, C., Rottenstreich, Y., & Knez, M. (2001). The illusion of leadership: Misattribution of cause in coordination games. *Organization Science, 12*(5), 582–598.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655–684.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103–128.

Winner, E., & Leekam, S. R. (1991). Distinguishing irony from deception: Understanding the speaker's second-order intention. *British Journal of Developmental Psychology, 9*(2), 257–270.

Zhang, J., & Hedden, T. (2003). Two paradigms for depth of strategic reasoning in games: Response to Colman. *Trends in Cognitive Sciences, 7*(1), 4–5.

Zhang, J., Hedden, T., & Chia, A. (2012). Perspective-taking and depth of theory-of-mind reasoning in sequential-move games. *Cognitive Science, 36*(3), 560–573.

Publication list

Journal publications

Meijering, B., Van Rijn, H., Taatgen, N. A., & Verbrugge, R. (2012). What eye movements can tell about theory of mind in a strategic game. *PLoS ONE, 7*(9). Doi:10.1371/journal.pone.0045961

Ghosh, S., **Meijering, B.**, & Verbrugge, R. (2014). Strategic reasoning: Building cognitive models from logical formulas. *Journal of Logic, Language, and Information*. Doi:10.1007/s10849-014-9196-x

Szymanik, J., Robaldo, L., & **Meijering, B.** (2014). On the identification of quantifiers' witness sets: A study of multi-quantifier sentences. *Journal of Logic, Language, and Information*. Doi:10.1007/s10849-014-9197-9

Peer-reviewed conference papers

Bergwerff, G., **Meijering, B.**, Szymanik, J. Verbrugge, & R. Wierda, S. M. (2014). Computational and algorithmic models of strategies in turn-based games. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society

Meijering, B., Taatgen, N. A., Van Rijn, H., & Verbrugge, R. (2013). Reasoning about mental states in sequential games: As simple as possible, as complex as necessary. In R. L. West & T. C. Stewart (Eds.), *Proceedings of the 12th International Conference on Cognitive Modeling* (pp. 173–178). Ottawa: Carleton University.

Meijering, B., Van Rijn, H., Taatgen, N. A., & Verbrugge, R. (2013). Reasoning about diamonds, gravity, and mental states: The cognitive costs of theory of mind. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 3026–3031). Austin, TX: Cognitive Science Society.

Szymanik, J., **Meijering, B.**, & Verbrugge, R. (2013). Using intrinsic complexity of turn-taking games to predict participants' reaction times. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 1426–1432). Austin, TX: Cognitive Science Society.

Meijering, B., Van Rijn, H., Taatgen, N., & Verbrugge, R. (2011). I do know what you think I think: Second-order theory of mind in strategic games is not that difficult. In L. Carslon, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2486–2491). Austin, TX: Cognitive Science Society.

Ghosh, S., & **Meijering, B.** (2011). On combining cognitive and formal modeling: A case study involving strategic reasoning. In J. van Eijck and R. Verbrugge (eds.), *Proceedings of the Workshop on Reasoning About Other Minds: Logical and Cognitive Perspectives* (pp. 79-92). Groningen, 2011.

Meijering, B., Van Maanen, L., Van Rijn, H., & Verbrugge, R. (2010). The facilitative effect of context on second-order social reasoning. In R. Catrambone & S. Ohlsson (Eds.), *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (pp. 1423–1428). Austin, TX: Cognitive Science Society.

Ghosh, S., **Meijering, B.**, & Verbrugge, R. (2010). Empirical reasoning in games: Logic meets cognition. In T. Ågotnes, N. Alechina, & B. Logan (Eds.), *Proceedings Third Logics for Resource Bounded Agents Workshop* (pp. 15–34). Lyon, FR.

Abstract

Meijering, B., Van Rijn, H., Taatgen, N. A., & Verbrugge, R. (2011). Second-order theory of mind in strategic games. *Interdisciplinary Workshop on Cognitive Neuroscience, Educational Research, and Cognitive Modeling*. Delmenhorst, Germany.

In preparation

Meijering, B., Van Rijn, H., Taatgen, N. A., & Verbrugge, R. (Submitted). *Reasoning about self versus others: Changing perspective is hard.*

Meijering, B., Van Rijn, H., Taatgen, N. A., & Verbrugge, R. (Under revision). *Integrating recursive theory of mind in decision making in sequential games.*

Meijering, B., Taatgen, N. A., & Verbrugge, R. (Under revision). *Modeling inference of mental states: As simple as possible, as complex as necessary.*