Reasoning about self and others

Meijering, Ben

Publication date:
2014

Chapter 2

Integrating recursive application of theory of mind in decision making in sequential games

Abstract

In collaborative, competitive, and negotiation situations we need to reason about each other's goals, intentions, beliefs, and desires, which requires a so-called Theory of Mind (ToM). This study investigates decision making in which ToM has to be applied recursively: "I think that you think that I think…". Participants were presented with sequential games in which the payoff for one player depended on the decisions of another player. Previous studies typically found suboptimal decisions in these games. One possible explanation is an overall inability to apply recursive ToM. Instead, we argue that suboptimal decisions are caused by unsuccessful integration of recursive ToM in the decision making process. This hypothesis is tested by means of three experimental manipulations that each should facilitate this integration process. First, each player's decision options are introduced and explained in a stepwise fashion during the training phase. Second, participants are prompted to predict the other player's decision. Third, participants are presented with a concrete and realistic task representation that visually cues the recursive structure of the decision making problem. The results show that performance was better in those conditions that specifically targeted the integration of recursive mental states in the decision making process.

Parts of this chapter were previously published in the *proceedings of the 32nd and 33rd Annual Conference of the Cognitive Science Society* (2010, 2011).

Introduction

Making decisions can be a complicated process, especially when the actions of others have to be factored in. To understand, or even predict, the actions of others we need to infer their goals, beliefs, desires, et cetera. It is difficult to infer these mental states because we cannot directly observe them. Still, children already learn to reason about the mental states of others from the age of 4. They develop a so-called theory of mind (ToM), an understanding that others may have beliefs that differ from their own and do not need to reflect reality (e.g., Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). However, making decisions based on mental state inferences remains challenging, even for adults (e.g., Flobbe, Verbrugge, Hendriks, & Krämer, 2008; Hedden & Zhang, 2002; Zhang, Hedden, & Chia, 2012).

Developmental studies have shown that children typically make optimal decisions if they are required to infer relatively simple mental states (Flobbe et al., 2008; Raijmakers, Mandell, Van Es, & Counihan, 2013). In Raijmaker et al.'s (2013) sequential games, for example, children performed well if they had to infer mental states such as "The other player *intends* to stop the game when it is his turn". In contrast, few children were able to make optimal decisions if they were required to infer more complex mental states such as "The other player *believes* that I *intend* to continue when it is my turn". Most children could not incorporate such second-order, or recursive, beliefs into their decision making process (for similar findings see Flobbe et al., 2008).

Some studies have shown that adults, too, find it difficult to make decisions based on second-order mental states (Flobbe et al., 2008; Hedden & Zhang, 2002; Johnson, Camerer, Sen, & Rymon, 2002; McKelvey & Palfrey, 1992; Verbrugge & Mol, 2008; Zhang et al., 2012). In Hedden and Zhang's (2002) sequential games, for example, adult participants initially did not seem to realize that the other player was reasoning about them. As a consequence they made suboptimal decisions that were based on inaccurate mental state representations.

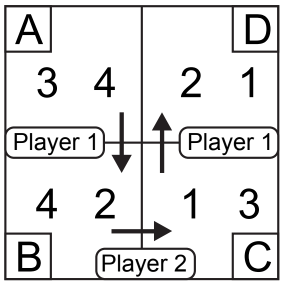

Figure 2.1: Example of a matrix game; adapted from Hedden and Zhang (2002); labels (A – D, Player 1 / 2) and arrows are added for illustrative purposes and were not depicted during the experiment.

Figure 2.1 depicts an example of Hedden and Zhang's sequential games, which are so-called matrix games. Each cell of a matrix game contains a pair of rewards, or so-called payoffs, that range from 1 to 4. The left payoff of a pair is Player 1's payoff; the right payoff is Player 2's. Both players alternately decide whether to stop the game in the current cell, or to continue it to the next. Each player's goal is that the game stops in the cell that contains his or her highest possible payoff, irrespective of the payoff for the other player. Based on the assumption that each player is rational, both players should know that their outcome depends on the decisions of the other player: Either player can decide to stop or continue the game when it is their turn. Crucially, as each player's decision is based on beliefs about the next decision, recursive beliefs have to be incorporated into the decision making process.

In the game in Figure 2.1, for example, Player 1 should decide to continue the game from A to B, as she could have inferred that a rational Player 2 will decide to stop the game in that cell. Player 2 would not continue to C, as he would have inferred that a rational Player 1 would continue to D, which contains a smaller payoff for him than B. As B contains a higher payoff for Player 1 than A, Player 1 should decide to continue from A to B. In sum, Player 1 needs to reason about Player 2's belief about Player 1's intention, thus applying second-order ToM by attributing first-order ToM to Player 2.

Most of the earlier mentioned studies concluded that adults and children do not have sufficient ability or cognitive capacity to apply second-order ToM (Flobbe et al., 2008; Hedden & Zhang, 2002; Raijmakers et al., 2013; Verbrugge & Mol, 2008; Zhang et al., 2012). Lacking the ability to apply ToM recursively, these studies claim, one makes suboptimal decisions that are based on inaccurate mental state representations. However, as children typically perform well in false-belief tasks that require them to keep track of stories with multiple agents that each have their own set of recursive beliefs (e.g., Apperly, Back, Samson, & France, 2008; Apperly et al., 2010; Hollebrandse, Hobbs, De Villiers, & Roeper, 2008; Perner & Wimmer, 1985; Sullivan, Zaitchik, & Tager-Flusberg, 1994; Wimmer & Perner, 1983), it seems that in some tasks sufficient ability or capacity is available to apply second-order ToM. We therefore hypothesize that suboptimal decisions are due to unsuccessful integration of recursive mental states in the decision making process. In many sequential games, participants have to attribute recursive mental states to the other player and, on top of that, they have to combine these attributions to choose their own best possible course of action. Suboptimal decisions arise if the integration of recursive mental states breaks down, even if the second-order representations are accurate.

To test whether suboptimal decisions in sequential games are due to unsuccessful integration of second-order mental states, we devised three experimental manipulations that each should scaffold the integration of mental states: (1) To make participants aware of the dependency between all decision points, each subsequent decision is introduced and explained in a stepwise fashion during the training phase; (2) To train integration of the other player's mental states, participants are prompted to predict the other player's decision; (3) To make the recursive structure of the decision making problem more salient, participants are presented with a new visual task representation. All three manipulations should clarify the "who, what, where" while making decisions in sequential games.

Method

Participants

Ninety-three first-year psychology students (63 female) with a mean age of 21 years (ranging between 18 and 31 years) participated in exchange for course credit. All participants had normal or corrected-to-normal visual acuity. Informed consent as approved by the Ethical Committee Psychology of the University of Groningen was obtained before testing.

Design

The experimental design comprised three factors: training, prompting predictions, and task representation. All factors were administered between participants. The experiment consisted of two phases: a training phase, followed by a test phase.

Training

The training phase was included to familiarize participants with the rules of sequential games. Participants were randomly assigned to one of two training procedures. In one training procedure participants were presented with Hedden and Zhang's (2002) 24 original training

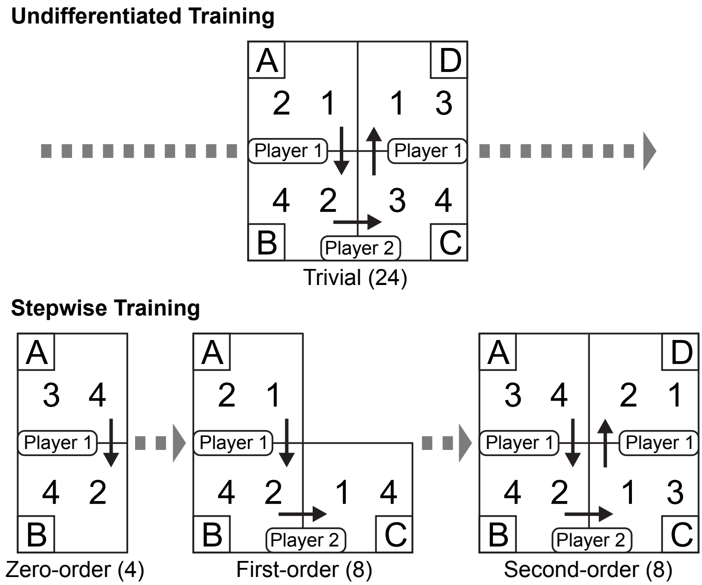

Figure 2.2: Schematic overview of the Undifferentiated and Stepwise training procedures. Undifferentiated training consists of 24 so-called trivial games (see text for explanation). Stepwise training consists of 4 zero-order games, 8 first-order games, and 8 second-order games. The actual training items all had different payoff distributions.

games (see Figure 2.2; top panel). These training games are considered easier to play than truly second-order games such as in Figure 2.1, because Player 2 does not have to reason about Player 1's last possible decision: Player 2's payoff in B is either lower or higher than both his payoffs in C and D. Consequently, Player 1 does not have to attribute ToM to Player 2. These games are therefore referred to as trivial games by Hedden and Zhang. This training procedure will henceforth be referred to as Undifferentiated training, as all games have three decision points.

In the other training procedure participants were presented with three blocks of games that are simple at first and become increasingly more complex with each subsequent block (see Figure 2.2; bottom panel). This procedure will henceforth be referred to as Stepwise training. The first block consisted of 4 games with just one decision point. These games are so-called zero-order games, as they do not require application of ToM. The second block consisted of 8 games with two decision points. These games require application of first-order ToM, as the participant is required to reason about the other player. The third block consisted of 8 games with three decision points that require application of second-order ToM, as the participant has to reason about the other player, and take into account that the other player is reasoning about them.

We hypothesize that the Stepwise training procedure provides scaffolding to support representation of increasingly more complex mental states. Stepwise introduction, explanation, and practice of each additional decision point helps participants integrate mental states of increasing complexity into their decision making process.

Prompting predictions

The second factor, prompting participants for predictions, was manipulated in the test phase. Hedden and Zhang (2002) prompted their participants to predict Player 2's decision (in B), before making a decision themselves. By prompting participants for predictions, participants were explicitly asked to take the other player's perspective, and we hypothesize that these prompts helped participants to integrate the other player's perspective in their decision making process.

We tested this hypothesis by means of two test blocks. In the first, we asked half of the participants, assigned to the Prompt group, to predict Player 2's move before making their own decision. Participants assigned to the No-Prompt group, in contrast, were not explicitly asked to predict Player 2's move. The second test block was added to test whether prompting had long-lasting effects on performance. No participant was asked to make predictions, and performance differences between the two groups would indicate lasting effects of prompting.

Task representation

The third factor is task representation. Before the training phase started, participants were assigned to one of two task representations, which did not change anymore during the remainder of the experiment. The Matrix Game was one of these task representations, and we devised a second. The new task representation was devised to clarify the recursive structure of the decision making problem. Depicted in Figure 2.3, the new task representation shows a more intuitive display of who decides where and what the consequences of each decision are. We argue that this new representation, henceforth referred to as Marble Drop, provides scaffolding that supports the integration of decisions and underlying mental states.

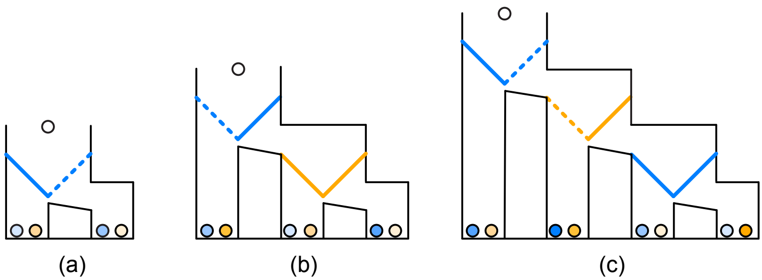

Figure 2.3: Examples of zero-order (a), first-order (b), and second-order (c) Marble Drop games. The blue player (i.e., Player 1) has to obtain the darkest possible blue marble, the orange player (i.e., Player 2) the darkest possible orange marble. The dashed lines are added for illustrative purposes and represent the trapdoors that proficient ToM players should remove to obtain their darkest possible marble. See text for additional explanation.

Importantly, this representation is isomorphic to matrix games and thus requires the same reasoning.

Figure 2.3 depicts examples of zero-order, first-order, and second-order Marble Drop games. Participants are told that a white marble is about to drop, and that its path can be manipulated by removing trapdoors. Their goal is to let the white marble drop into the bin containing the darkest possible marble of their target color, blue in these example games, by controlling only the blue trapdoors. Player 2's goal is to obtain the darkest possible orange marble, but Player 2 can only control the orange trapdoors. The marbles are ranked from light to dark, with darker marbles preferred over lighter marbles, yielding payoff structures isomorphic to those in matrix games.

Stimuli

Payoffs

The payoffs in matrix games are numerical, ranging from 1 to 4, whereas the payoffs in Marble Drop games are color-graded marbles that have a one-to-one mapping to the numerical values in the matrix games. The colors of the marbles are four shades of orange and blue, taken from the HSV (i.e., hue, saturation and value) space. A sequential color palette is computed by varying saturation, for a given hue and value. This results in four shades (with saturation from .2 to 1) for each of the colors orange (hue = .1, value = 1) and blue (hue = .6, value = 1).

Payoff structures

The payoff structures are selected so that the order of ToM reasoning mastered by the participants can be derived from their first decision[1]. The total set of payoff structures, balanced for the number of decisions to continue or stop a game, is limited to 16 items. These items are listed in Appendix A. Detailed discussion of the rationale behind the exclusion

1 We look at the entire set of first decisions, as an individual decision cannot discriminate between multiple strategies. For example, in one particular game both guessing and applying second-order ToM might yield a correct decision. However, by looking at the entire set of decisions, we can discriminate guessing from applying second-order ToM, as guessing would only yield a correct decision in 50% of the games.

criteria is given in Appendix B.

Procedure

To familiarize participants with the rules of sequential games, they were first presented with a training block that either consisted of Undifferentiated training or Stepwise training. The instructions, which appeared on screen, explained how to play sequential games and what the goal of each player was. The instructions also mentioned that participants were playing against a computer-simulated player, as Hedden and Zhang (2002) have shown that inclusion of a cover story did not affect ToM performance. Each training game was played until either the participant or the computer-simulated player decided to stop, or until the last possible decision was made. After each training game participants were presented with accuracy feedback indicating whether the highest attainable payoff was obtained. In case of an incorrect decision, an arrow pointed at the cell / bin that contained the highest attainable payoff. As the feedback never referred to the other player's mental states, participants had to infer these themselves.

The two test blocks consisted of second-order games. As mentioned above, the procedure for participants in the Prompt and No-Prompt groups differed in the first test block. Participants in the Prompt group were first asked to enter a prediction of Player 2's decision before they were asked to make a stop-or-continue decision at their own decision point. Participants in the No-Prompt group, in contrast, were not asked to make predictions. They were only asked to make decisions. Accuracy feedback was presented both after entering a prediction and after entering a decision, but the arrow was not shown anymore in the test blocks. This block consisted of 32 trials; each of the 16 payoff structures was presented twice, but not consecutively as the items were presented randomly.

The second test block was the same for all participants. They were asked to make decisions only.

Results and discussion

To account for multiple sources of random variation (i.e., participants and payoff structures were both sampled from a larger population), the data were analyzed by means of linear mixed-effects (LME) models (Baayen, 2008; Baayen, Davidson, & Bates, 2008). We included random intercepts to allow the intercepts of the regression models to vary across participants and items. Random slopes were included to allow the effects (i.e., slopes) of training, prompting, and task representation to vary across items (Barr, Levy, Scheepers, & Tily, 2013). The correctness of the decisions was analyzed by means of *logistic* LMEs, as correctness of decisions is a binary variable. These models are provided by the lme4 package (version 0.999375-42; Pinheiro & Bates, 2000) in R (www.r-project.org, version 3.0.1). Two separate models were fit to the data because the factors training and prompting were manipulated in different blocks. All figures depict means and standard errors, which are represented by error bars.

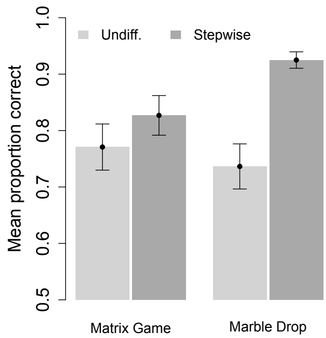

Figure 2.4. Mean proportions correct decisions; depicted separately for Undifferentiated training (light grey) and Stepwise training (dark grey) for both Matrix Games and Marble Drop games.

Effect of training and representation

Type of training was manipulated before either test block was administered. Its effect on the correctness of decisions was analyzed in the first test block, together with the effect of task representation. Figure 2.4 depicts the mean proportions of correct decisions.

As can be seen in Figure 2.4, the effect size of Stepwise training (in contrast to Undifferentiated training) significantly varied with task representation, $\chi^2(1) = 6.4$, $p = 0.011$. Figure 2.4 shows that stepwise training had a significant positive effect on the correctness of decisions, $\chi^2(1) = 12.1$, $p < 0.001$, but this effect was driven by participants assigned to Marble Drop games, $\beta = 2.063$ (SE = 0.491), $z = 4.203$, $p < 0.001$. In fact, the effect of stepwise training was not significant for participants assigned to Matrix Games, $\beta = .427$ (SE = .432), ns.

Even though there was no significant main effect of task representation, participants assigned to stepwise training did have a significantly higher probability of making a correct decision if they were assigned to Marble Drop games instead of Matrix Games, $\beta = 1.307$ (SE = 0.483), $z = 2.707$, $p = 0.007$. Participants assigned to undifferentiated training did not have an advantage if they were assigned to Marble Drop games, $\beta = -.326$ (SE = .439), ns. In other words, the effectiveness of the Marble Drop task representation, positive in any case, varied with the type of training received by the participants.

In sum, both the Stepwise training procedure and the Marble Drop task representation positively affected the probability of making a correct decision, the latter factor primarily by means of an interaction.

Effect of prompting

As prompting was manipulated in the first test block, we analyzed its short-term effect during the first test block and its longer-term effect in the second test block. Figure 2.5 depicts the accuracy, or proportion correct, of the decisions and predictions.

Log-likelihood ratio comparisons indicated that the factor task representation did not make the models fit the decisions better. Therefore, we report the statistics of a full factorial model that includes terms for *prompting*, *test block*, and an interaction between the two.

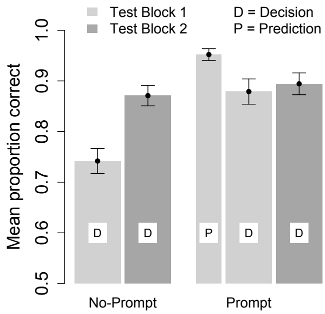

Figure 2.5. Mean proportion of correct predictions and decisions, depicted separately for the No-Prompt and Prompt groups, the first and second test block.

As can be seen in Figure 2.5, the extent to which the proportion of correct decisions increased from the first test block to the second depended on whether participants were prompted to predict the other player's decision, $\chi^2(1) = 12.2$, $p < 0.001$. There was a significant main effect of *test block*, $\chi^2(1) = 78.8$, $p < 0.001$, but it was mainly driven by the No-Prompt group, $\beta = 1.072$ (SE = 0.12), $z = 9.96$, $p < 0.001$. The probability of making a correct decision increased just slightly in the Prompt group, $\beta = .351$ (SE = .156), $z = 2.256$, $p = .024$.

There was also a main effect of prompting participants for predictions, $\chi^2(1) = 13$, $p < .001$. However it was mostly present in the first test block, as the difference between the prompting conditions in the second test block was small and just significant, $\beta = .639$ (SE = .314), $z = 2.032$, $p = 0.04$. Thus, during Test Block 2, participants that were not prompted for predictions in Test Block 1 performed almost as well as participants that *were* prompted. Nevertheless, prompting did have a stronger positive effect in the short term during Test Block 1, $\beta = 1.36$ (SE = .304), $z = 4.467$, $p < 0.001$.

Figure 2.5 also shows that participants performed close to ceiling with respect to their predictions, which also require second-order theory of mind. However, their proportion of correct decisions was significantly lower, $t(46) = -3.1827$, $p < 0.01$. Thus, a correct prediction did not always yield a correct decision, which implies that it is not trivial to incorporate a second-order inference in the decision making process. This is remarkable, as the application of theory of mind is required when making a prediction, and not anymore when a decision has to be made afterwards. Nevertheless, this finding supports our hypothesis that integration of mental states, when making decisions, is the crux of suboptimal performance.

General conclusions

In this study we investigated decision making in the context of sequential games. Crucially, participants were asked to make decisions that required them to infer second-order, or recursive, mental states. Previous studies have found suboptimal performance in such games and concluded that incorrect decisions were caused by insufficient ability to apply second-order ToM (Flobbe et al., 2008; Hedden & Zhang, 2002; McKelvey & Palfrey, 1992; Raijmakers et al., 2013; Zhang et al., 2012). Other studies, however, have found that participants have few difficulties applying second-order ToM if they are not required to make decisions (Flobbe et al., 2008). We therefore argue that the crux of suboptimal performance in sequential games is the integration of second-order ToM in the decision making process. Our results confirm this idea and show that decision making in sequential games can be improved by scaffolding the integration of second-order ToM.

Our hypothesis is best supported by the finding that correct predictions do not always result in correct decisions. Interestingly, a prediction of what the other player would do requires the application of second-order ToM, as participants have to reason about the other player, who in turn reasons about their last possible decision. Participants performed almost at ceiling when asked to predict the other player's decision. In other words, the participants had sufficient ability to apply second-order ToM. Nevertheless, they performed significantly worse when asked to make a decision based on their predictions. This finding implies that participants found it difficult to integrate second-order ToM in the decision making process, not to apply second-order ToM.

The other findings, that decision making improves by providing participants with Stepwise training, prompts for predictions, and the Marble Drop task representation, support our hypothesis as well. As mentioned previously, these conditions should scaffold the integration of recursive mental states. In Stepwise training, participants practice how to make decisions based on mental states of increasing complexity. When prompted to predict the other player's decision, participants are asked to take the other player's perspective, which is an essential step that should precede and be the basis of their own decision making. The Marble Drop task representation provides clear visual cues as to how the games are structured. In fact, a previous eye tracking study (Meijering, Van Rijn, Taatgen, & Verbrugge, 2012) has shown that participants use these cues while reasoning. For example, the participants in Meijering et al.'s (2012) study kept fixating the trapdoors during the entire experiment, even though the trapdoors remained a constant factor that did not change across the training and test phases. Meijering et al. argued that the trapdoors were used as placeholders for mental states, thereby providing scaffolding for the decision making process. In other words, the Marble Drop task representation, together with the other factors, facilitated the integration of second-order ToM in the decision making process.

The effects were most prominent in the first test block. Participants who were not assigned to stepwise training, who were not prompted to take the other player's perspective, and who played abstract Matrix Games during the entire experiment, did learn in the long run how to incorporate complex mental states into their decisions. In the second-test block the differences in performance between the conditions failed to reach significance. Thus, it seems that participants benefitted most from supporting structure early on in the task when they had not yet developed a strategy for which their cognitive resources were sufficient, as we will explain below.

Before we continue with a cognitive explanation, we would first like to address the concern that people were playing against a computer-simulated player instead of a real player. Based on Hedden and Zhang's (2002; 2012) findings we did not construct a cover story to make participants believe they were playing against a real player. In Hedden and Zhang's study it did not matter whether participants thought to be playing against a real or a computer-simulated player. Furthermore, the level of intelligence participants attributed to the real player did not affect their default mental model of the other player. In other words, the participants were aware of the nature of the other player, but that awareness did not affect what beliefs, goals, and intentions they attributed to that player. Moreover, other studies have shown, too, that people do attribute (human) mental states to computer-simulated players (Flobbe et al., 2008; Meijering et al., 2012; Meijering, Van Rijn, Taatgen, & Verbrugge, 2013).

Given that sequential games evoke the application of (recursive) ToM, we argue that our findings are generalizable to other ToM settings, including everyday social interactions. Still, generalizability depends on the extent to which supportive measures are effective in reducing demands on cognitive resources. A slightly different notion of generalizability is the extent to which experience on our task will improve application of ToM in other domains. Here, we are more conservative because Flobbe et al.'s (2008) study has shown that performance on one particular ToM task does not necessarily correlate with performance on another ToM task. As Flobbe et al. have shown that application of ToM may not be a unitary skill, it is conceivable that training in one ToM task does not necessarily generalize to other ToM tasks (also see Thoermer, Sodian, Vuori, Perst, & Kristen, 2012).

A cognitive explanation for the facilitative effects of training, prompting, and task representation may be in terms of reducing demands on executive functions. Examples of executive functions are planning, resistance to interference, set-shifting, and working memory, which all help to combine alternate perspectives and find the best possible decision (Apperly & Butterfill, 2009; Bull, Phillips, & Conway, 2008; Dumontheil, Apperly, & Blakemore, 2010). Our experimental manipulations may have reduced demands on executive processes in three ways. First, in the Stepwise training condition, participants received naturally delineated chunks of instruction and training of each successive ToM-order. A clear outline of the task at hand helped planning and reduced demands on working memory. Second, by prompting for predictions, we may have structured participants' reasoning by providing them with an efficient method of 'solving' games. A structured method may not only help planning and thereby reduce demands on working memory; it may also help set-shifting, that is, switching between goals, beliefs, and intentions. Third, the Marble Drop task representation visually cued possible actions and consequences, and provided placeholders for mental states. These cues and placeholders allowed participants to preserve working memory capacity and help them planning their own actions.

Support for these explanations comes from studies that have shown that application of ToM is an effortful process (e.g., Apperly et al., 2010; Lin, Keysar, & Epley, 2010). In Lin et al.'s (2010) study, for example, working memory capacity was positively correlated with efficiency in applying ToM. Participants with low working memory capacity were less efficient in applying ToM than participants with high working memory capacity. Moreover, a second experiment in Lin et al.'s study demonstrated that participants' ability to apply ToM was significantly reduced by a secondary task. These findings show that even though people may have the capacity to apply ToM, they may fail in correctly using ToM, which was already noted by Keysar et al. (2003). Lin et al.'s findings imply that, the other way around, efficiency may improve if demands on working memory, and other executive functions, are reduced. Our results seem to corroborate this notion.

A similar explanation has been proposed for suboptimal behavior in the multitasking setting. Borst et al. (2010a; 2010b) have shown that multitasking in itself is not difficult, but that instead working memory constraints cause the often-claimed difficulties associated with multitasking. If both tasks cause high working memory load, performance breaks down. However, performance does not deteriorate if at most one of the tasks causes high working memory load. By reducing working memory load in our task, participants had more freely available cognitive resources to incorporate mental states into their decision making process (see also Van Rij, Van Rijn, & Hendriks, 2010; 2013).

To conclude, decision making can be a complex task, depending on the variables involved. In this study, decisions were dependent on mental states. Participants had to apply second-order ToM and use the outcome as a basis for their decision making. This proved especially difficult: Correct predictions, which required inference of second-order mental states, did not always result in correct decisions. We argue that participants failed to make optimal decisions because they had difficulties integrating the complex mental states. Our experimental manipulations specifically and successfully targeted this integration process, showing that decision making on the basis of mental states appears to be flexible in the sense that it is susceptible to improvement. Generalizing our findings to everyday life, one could argue that the complex decision making that we engage in in many social settings is a skill that we can develop. We can improve it by taking measures that help us incorporate the mental states of others.

Appendix A

The table below lists the payoff structures that were used in the experiment. The prediction and decision columns list whether the correct responses were either to stop (i.e., 0) or to move (i.e., 1) at the corresponding decision points.

Payoffs Player 1				Payoffs Player 2					
A	B	C	D	A	B	C	D	Prediction	Decision
Zero-order payoff structures									
1	3			3	2				1
3	1			2	1				0
4	2			2	4				0
3	4			2	4				1
First-order payoff structures									
2	1	3		1	2	3		1	1
2	1	3		3	2	1		0	0
2	3	1		1	2	3		1	0
2	3	1		3	2	1		0	1
3	2	4		2	3	4		1	1
3	2	4		4	3	2		0	0
3	4	2		2	3	4		1	0
3	4	2		4	3	2		0	1
Second-order payoff structures									
3	1	2	4	2	3	4	1	0	0
3	1	2	4	4	2	3	1	0	0
3	2	1	4	1	3	4	2	0	0
3	2	1	4	4	2	3	1	0	0
3	4	1	2	1	3	2	4	1	0
3	4	1	2	2	3	1	4	1	0
3	4	1	2	3	2	1	4	1	0
3	4	1	2	4	2	1	3	1	0
3	4	1	2	1	3	4	2	0	1
3	4	1	2	2	3	4	1	0	1
3	4	1	2	3	2	4	1	0	1
3	4	1	2	4	2	3	1	0	1
3	1	2	4	2	3	1	4	1	1
3	1	2	4	4	2	1	3	1	1
3	2	1	4	1	3	2	4	1	1
3	2	1	4	4	2	1	3	1	1
Trivial payoff structures									
2	4	3	1	1	2	4	3	1	1
3	4	2	1	1	2	3	4	1	0
2	1	4	3	4	3	1	2	0	0
3	4	1	2	4	3	1	2	0	1
2	3	4	1	1	2	3	4	1	1
3	4	1	2	4	3	2	1	0	1

Payoffs Player 1				Payoffs Player 2					
A	B	C	D	A	B	C	D	Prediction	Decision
2	1	3	4	4	3	1	2	0	0
3	4	2	1	1	2	4	3	1	0
3	4	2	1	4	3	2	1	0	1
2	1	4	3	4	3	2	1	0	0
3	4	1	2	1	2	4	3	1	0
2	4	3	1	1	2	3	4	1	1
3	4	2	1	4	3	1	2	0	1
3	4	1	2	1	2	3	4	1	0
2	3	4	1	1	2	4	3	1	1
2	1	3	4	4	3	2	1	0	0
3	2	4	1	4	3	1	2	0	0
2	4	1	3	4	3	2	1	0	1
2	1	4	3	1	2	3	4	1	1
3	4	1	2	4	1	2	3	1	0
2	1	3	4	1	2	4	3	1	1
2	3	1	4	4	3	2	1	0	1
3	4	2	1	4	1	2	3	1	0
3	1	4	2	4	3	2	1	0	0

Appendix B

Payoff structures are excluded if Player 1's payoff in A is either a 1 or a 4, because Player 1 would not need to reason about the Player 2's decision. It is obvious that Player 1 should continue the game if his payoff in A is a 1 and stop if his payoff in A is a 4. The game in Figure B.1b is an example of a game in which Player 1 should decide to stop in A. In line with Hedden and Zhang (2002), we focused on so-called 2- and 3-starting games, associated with payoff structures in which Player 1's first payoff was a 2 or a 3, respectively.

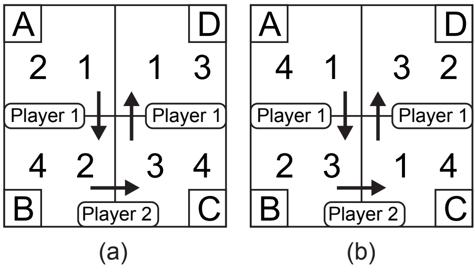

Figure B.1: Two example matrix games. The game in *a* is a so-called trivial game. See text for explanation. The game in *b* does not require any ToM reasoning at all, because Player 1's maximum payoff is already available in A.

We also excluded payoff structures in which Player 2's payoff in B was either a 1 or a 4, because Player 2 would not need to reason about Player 1's last possible decision between C and D. Accordingly, first-order reasoning on the part of Player 1 would suffice.

Of the remaining payoff structures, we excluded the so-called trivial ones in which Player 2's payoff in B was either lower or higher than both his payoffs in C and D. Figure B.1a depicts an example of such a game: Player 2 does not need to reason about Player 1's last possible decision, as his payoffs in C and D are both more preferable than his payoff in B.

The next two exclusion criteria are based on the type of Player 2 that a participant could be reasoning about. In line with Hedden and Zhang, we distinguished between a hypothesized zero-order Player 2 and a hypothesized first-order Player 2. Hedden and Zhang consider these Player 2 types to be either myopic or predictive, respectively. A participant, always assigned to the role of Player 1, might be reasoning about a zero-order Player 2, which would *not* reason about the participant's last possible decision. Hedden and Zhang consider such a Player 2 to be myopic, as it only considers Player 2's own payoffs in B and C. In contrast, a participant might be reasoning about a hypothesized first-order Player 2, which *does* reason about the participant's last possible decision.

As Player 1's decision depends on Player 2's decision, payoff structures that yield the same answer for zero-order and first-order Player 2s cannot inform us on the level of ToM reasoning on the part of Player 1. These payoff structures are considered non-diagnostic, and are therefore excluded from the final set of stimuli. In contrast, Hedden and Zhang included these payoff structures as long as the decision (and thus prediction) of Player 2's move at the second decision point was opposite for imagined zero- and first-order Player 2s. As half of the participants were not prompted for predictions in the first test block and none of them in the second test block, we had to exclude these items.

We selected a final set of stimuli, which we were able to (double-)balance for both the number of *stop* and *continue* decisions of Player 1 and the number of *stop* and *continue* decisions of Player 2. As this was only possible for 3-starting games, we excluded the 2-starting games. This left us with 16 unique payoff structures, all 3-starting games.

Using the same criteria as mentioned above, we selected 4 zero-order and 8 first-order payoff structures for the Stepwise Training condition. These training payoff structures, as well as the 16 second-order payoff structures, are listed in Appendix A. The trivial games used in the Undifferentiated Training condition, which were described earlier, are also listed in Appendix A.