The Role of Social Cognition in Reading Comprehension: A Developmental Perspective

Marcella Pavias, student Developmental Psychopathology in Education and Child Studies, Leiden University

Name	M. Pavias
Student number	1103865
Research Master	Developmental Psychopathology in Education and Child Studies
Specialization	Educational Studies
Supervisor	Dr. L. van Leijenhorst
Second reader	Prof.dr. P. W. van den Broek
Date	August 2014

Abstract

This eye-tracking study investigates whether age-related changes in the ability to take perspective influence narrative text comprehension. Thirty-two typically-developing children (*M* = 11.73; *SD* = 0.74) and 34 young adults (*M* = 21.02; *SD* = 1.98) read stories in which the need to use perspective-taking abilities was systematically varied. The offline measure (after reading) suggested that adults were better and faster at making inferences in general, and both 10-12-year-olds and 18-25-year-olds were faster in making an inference in the complex perspective-taking condition (which required them to take the perspective of one of the story characters and imagine how this character would react to the intentions, thoughts, or feelings of another story character) compared to the control condition (which required them to make an inference about physical causality with regard to an object). The reading process itself revealed that 18-25-year-olds read stories faster across all conditions we examined. In addition, both 10-12-year-olds as well as adults revealed the longest reading times in the most difficult condition in which complex perspective-taking was needed to draw inferences. Stories in which the interaction between two story characters has to be taken into account are processed differently compared to stories in which no social-cognitive information is needed, as well as compared to when one only has to take the perspective of one story character. Narratives in which perspective-taking is crucial for comprehension are more difficult to process, even for adults, but are better represented in the situation model readers construct.

Key words: reading comprehension, perspective-taking, development, social cognition

The Role of Social Cognition in Reading Comprehension: A Developmental Perspective

Reading comprehension is an important predictor of educational achievement. Many students fail to reach a sufficient level of reading comprehension skills. Almost 13 % of 15-year-old students in the Netherlands have low literacy skills and this probably results in difficulties in school and difficulties with functioning in society (Kordes, Bolsinova, Limpens, & Stolwijk, 2013). Although all fourth grade students (approx. 10 years old) in The Netherlands are able to locate and retrieve information from different parts of the text, only 48 % were able to interpret and integrate ideas and information, and 10 % were not even able to make straightforward inferences about the attributes, feelings, and motivations of main characters (Mullis, Martin, Foy, & Drucker, 2012). To be able to prevent reading comprehension difficulties it is important to understand the factors and mechanisms that underlie these difficulties. In this study, we attempt to clarify from a developmental viewpoint whether age-related changes in the ability to take perspective influence comprehension of a narrative text.

For readers, the fundamental objective for comprehension is to construct a representation of the information conveyed by the text; a situation model. This is an integrated mental representation of the text that the reader creates (Kintsch, 1998; Zwaan & Radvansky, 1998) and is the foundation for further comprehension of the text, for recalling the text, for answering questions about it, and for other off-line measures of reading comprehension (Van den Broek & White, 2012). For this representation to be accurate it needs to include idea units that are connected in meaningful relations that are integrated into a coherent whole together with background knowledge (Van den Broek & White). To achieve this, a reader often needs to infer these relations.

Referential and causal relations in a text are crucial for comprehension. Causal relations inform the reader about how an event described in one sentence influences or causes an event in other parts of the text (Van den Broek & White). These relations are a substantive part of the situation model.

During the construction of this model readers produce various types of causal connections in order to link together the events in the narrative they are reading (Kendeou, Van den Broek, White & Lynch, 2009); the reader makes inferences. Therefore, to comprehend a text it is necessary that the reader makes these inferences. The ability to make inferences improves with age. Children from second-, fifth-, and eighth-grade as well as undergraduate students are capable to make inferences, but the ability to do so improves when getting older (Casteel & Simpson, 1991). Even when 6-15-year-olds have the same knowledge available to make the required inference, older children outperform younger children and make more inferences (Barnes, Dennis, & Haefele-Kalvaitis, 1996).

In narratives one often has to take the goals, needs, and feelings of the protagonist into account. Perspective-taking ability permits individuals to anticipate the behaviour and reactions of others (Davis, 1983). In the context of narrative texts this might possibly imply that for individuals who have better perspective-taking abilities the construction of the situation model is easier and done with less processing effort due to the fact they can better anticipate what is coming with regard to the protagonists in the story. To our knowledge there is no evidence for this in the literature. However, when focusing on the neural representation of the situation model, regions associated with considering the mental state of another person are indeed implicated for narrative comprehension (Mano, Harada, Sugiura, Saito & Sadato, 2009). In this fMRI study in which the response to sentence pairs was compared, the protagonist could either be aware of the emotion-inducing event in the second sentence or could not be aware of this event because the event took place in another location than the

location in which the story character was located. Compared to the control condition (in which the story character in the second sentence was different from the character in the first sentence) these conditions elicited activation in brain regions that are implicated in emotional comprehension. Additionally, the condition in which the character was not present elicited more response in the right temporo –parietal junction and the precuneus/posterior cingulated cortex compared to the condition in which the character was present during the event that could trigger an emotional response. This suggests that these regions are implicated in spatial perspective-taking (Mano et al., 2009). The brain regions of Theory of Mind (ToM), for which perspective-taking is an important ability, continue to develop between adolescence and adulthood (Blakemore, 2008). In a study in which adults (ages 11-16) and adolescents (ages 24-40) were asked to choose an appropriate ending after looking at three cartoon frames, adults outperformed adolescents when it was necessary for successful completion of the task to infer how the story character would react to the other character's emotional state (affective ToM condition). Adolescents also made more errors in this condition compared to when it was not necessary to infer the mental states of the characters in the cartoon. Additionally, the affective Tom condition elicited more activation in the left ventromedial prefrontal cortex than when it was not necessary to infer the mental state. For adults this difference was not found. This differential neural response in the ventromedial prefrontal cortex between the adult and adolescent group is a sign of developmental changes in affective ToM processing (Sebastian et al., 2012).

Additional support for the idea that socio-emotional processing contributes to narrative comprehension comes from a computational study using the landscape model. The Landscape Model portrays the dynamic processes of online comprehension and at the same time investigates how these processes lead to the construction of a stable memory representation of a text (Van den Broek, Risden, Fletcher & Thurlow, 1996). In a study on reading and recall of

narrative texts the generation of emotional inferences was examined using the Landscape Model (Van den Broek, 1995). A model in which emotional inferences were included was a better match to the human data than one in which only causal and referential inferences were included (Marotto, Barreyro, Cevasco, & Van den Broek 2011). Additionally, it appeared that emotional inferences play a role in comprehension of natural texts. Reading narrative texts can also be linked to the development of empathy. Performance of 17-38-year-olds on the Reading the Mind in the Eyes task (a measure of social intelligence (Baron-Cohen, Wheelwright, Hill, Raste, & Plimb, 2001)), in which individuals infer the mental state of a person based on a picture of their eyes, could be predicted from narrative text exposure (Mar, Oatley, & Peterson, 2009). Participants who scored higher on the narrative scale of an adjusted version of the Author Recognition Test (Mar, DeYoung, Higgins, & Peterson, 2006), and thus were assumed to have had more narrative text exposure during their life, were better able to complete the Reading the Mind in the Eyes task.

Even though the ability to take the perspective of another continues to develop throughout childhood and adolescence (Gurucharri & Selman, 1982; see also Choudhury, Blakemore, & Charman, 2006; Van der Graaff et al., 2014), and is a crucial part of the development of social cognition (Lapsley, 2006), little is known about the role of perspective taking abilities in the development of reading comprehension. However, the research discussed above suggests that the ability to take perspective might also influence reading comprehension and during development individuals develop higher-order perspective taking skills. Adolescents (aged 14-18), for example, are less able to carry out instructions in which they have to take the perspective of another person into account than adults (aged 19-27). In this computerized version of the Director Task (Keysar, Barr, Balin, & Brauner, 2000), the director (a fictional character positioned opposite to the participant's view) instructed participants to move a particular object. However, the director could see some, but not all of

the objects the participants saw due to the fact that the view on some objects was blocked to the director. The perspective of the director was needed in order to execute the instructions accurately. The adults outperformed the adolescents when asked to move objects in this communicative task (Dumontheil, Apperly, & Blakemore, 2010). This suggests that the ability to perform complex perspective-taking tasks continues to develop in late adolescence.

To investigate the role of perspective taking in reading comprehension, and whether this role is similar in different stages of development, reading comprehension is studied in children (10-12-year-olds) and young adults (18-25-year-olds). The general research question is: *Do age-related changes in perspective taking skills influence the ability to make causal inferences that require social cognition?*

For this purpose individuals of the different age groups read three different types of short narratives in which the need to use perspective-taking abilities for comprehension was systematically varied.

Reading comprehension can be assessed with offline and online tasks. The former are measures that take place when the reader has completed the reading process, thus after reading the text, such as a recall or answering questions. The latter is a task that takes place during reading, such as the thinking aloud protocol or eye tracking. An advantage of online reading comprehension measures is that the processes during reading can be studied instead of only the result of the reading process. Eye tracking provides an opportunity to combine behavioral offline measures with online measures of reading comprehension. For this purpose we collected both online (story reading times) and offline (accuracy & choice reaction time) data from all participants using the combination of eye-tracking and the computerized reading task.

With regard to development it is expected that adults are better at making inferences than children as the ability to make inferences improves with age (Barner et al., 1996; Casteel &

Simpson, 1991; Hulme & Snowling, 2009; Hypothesis 1a). We expect that all groups perform well above chance level with regard to accuracy because the stories are developed at an age-appropriate level for the children (Hypothesis 1b).

Furthermore we anticipate that adults will be faster than children to choose the correct answer (Hypothesis 2).

With regard to the different conditions, it is expected that children and young adults are faster in making an inference in which the reader does not have to take the perspective of one of the protagonists to be able to make the inference, and the inference is based on physical causality with regard to an object, compared to stories in which the perspective of the protagonist and the reaction of the other character on the protagonist has to be taken into account (Hypothesis 3).

With regard to reading patterns it is expected that there will be differences in reading time in the different conditions and in age group when reading the stories. In general, we expect children to have a longer reading time across all conditions we examined (Hypothesis 4). In addition we expect that in this online measure participants have a longer reading time in the complex condition as the perspective of two characters needs to be taken into account simultaneously and this requires more social cognitive abilities (Hypothesis 5).

Method

Participants

Two age groups were recruited. The first group consisted of 36 children (19 girls) between the ages of 10;1 and 12;9 (M = 11.73; SD = 0.74) recruited from a Dutch primary school. After permission from the headmaster of the school parents of 58 children were contacted through a letter. Sixty-two percent of the parents gave written consent. The second group encompassed 38 young adults (35 females) between the ages of 18;1 and 25;5 (M = 21.02; SD

= 1.98) recruited from Leiden University and of which the majority participated in exchange for course credits (71%) and others volunteered to participate.

After informed consent of the parents or the participant, participants were screened for word decoding difficulties (using the three-minute-test), and for other exclusion criteria such as diagnosed developmental disorders, e.g. Attention Deficit and Hyperactivity Disorder, neurological, or reading disabilities, such as dyslexia, or the use of drugs that act upon the central nervous system (e.g. stimulants and amphetamines) (using the background information questionnaire). Data from participants who reported these difficulties or drug use, as well as data from participants with below-average word-decoding skills were excluded from all further analyses. This resulted in a sample consisting of 32 children (17 girls) between the ages of 10;1 and 12;9 (M = 11.67; SD = 0.73) and 34 adults (32 females) between the ages of 18;1 and 25;5 (M = 20.98; SD = 2.02). In addition, a short version of Raven's Progressive Matrices was used to estimate participants' IQ. All individuals had IQ scores and vocabulary scores in the normal range, and IQ scores did not differ between age groups, p = .12.

Materials

Story-completion task. All participants read 45 four-sentence stories that were written based on cartoons by Sebastian et al. (2012), and in which we systematically varied the need for social cognitive abilities for comprehension. Each story consisted of a first sentence in which two characters were introduced, a second sentence which described a situation, a third sentence in which an action took place, and a final fourth sentence, the target sentence in this experiment, in which two possible endings of the story were presented. Participants had to choose the best fitting final sentence on each trial. One variant was plausible when the participant made an inference that logically followed from the first three sentences. The other version was less plausible. Plausibility was verified by at least three different researchers (see

Table 1 for an example stimulus). Stories were presented as a whole to facilitate a close-to-natural reading experience, compared to sentence-by-sentence presentation.

We created three conditions in which the need for perspective-taking abilities for comprehension was varied. Each individual read 15 stories in each condition. Two random orders of the 45 stories were created and participants were arbitrarily assigned one of these versions. In all conditions there was a causal dependency between the previous sentences and the target sentence. This means that to understand the two versions of the target sentence, and to choose which is the most appropriate, one needs to use knowledge about psychological and physical causality in order to become aware of the causes of the events in these target sentences, and create a coherent mental representation of the story. The first *control* condition required the participants to make an inference about physical causality with regard to an object. In the second *simple perspective-taking* condition participants need to make an inference which requires them to take the perspective (intention, thoughts, or feelings) of one of the story characters. In the third *complex perspective-taking* condition participants need to make an inference that requires them to imagine how one of the story characters would react to the intentions, thoughts, or feelings of the other story character.

The length of the target sentences varied between 3 and 25 syllables, ($M_{plausible}$ = 10.44 , $SD_{plausible}$ = 4.68; $M_{lessplausible}$ = 10.04, $SD_{lessplausible}$ = 4.64). The number of syllables for the plausible and less plausible target sentences did not differ, p = 0.69. There were also no differences in number of syllables of the target sentence between the conditions, p = 0.31.

Participants read each three-sentence story on a computer monitor, in a self-paced manner by pressing the left mouse key to advance to a fixation dot (500 ms) after which the target sentences appeared. If participants did not advance within 30 seconds, the fixation dot automatically appeared. Participants indicated which of the two target sentence options they wanted to select by clicking on a box next to the sentence using the computer mouse.

Following their choice participants were asked to rate by clicking one of 5 boxes, which corresponded to a 5-point scale, how certain they were about their choice. If they did not select one of the target sentence options within 20 seconds the rating screen appeared. Participants had 10 seconds to indicate how certain they were about their choice. After this, and prior to the next story, a drift-check dot appeared. For an example trial see Figure 1. Choice reaction times as well as eye-tracking data were collected during the task.

Eye tracking. Eye-tracking data was collected with the EyeLink 1000 eye tracker (SR Research Ltd, Mississauga, Ontario, Canada) using the Desktop Mount with a sampling rate of 500 Hz. The stories were presented on a Dell P1913S computer screen, with a refresh rate of 60 Hz and a screen-resolution of 1024x768 pixels, at approximately 60 cm distance from the participant's eyes. Head movement was minimized by using a chin and head rest. All text was presented in black Verdana 22-point on a white background with anti-aliasing to make sure the letters looked smooth on the screen. Movements of the right eye were tracked, but viewing was binocular. The eye tracker was calibrated at the start of the experiment with a thirteen-point calibration with a pacing interval of 1000 ms. A drift check was performed before each trial, the eye tracker was recalibrated if necessary.

Perspective taking. The Interpersonal Reactivity Index (IRI; Davis, 1980) was used to assess differences in empathy and more specifically to measure differences in perspective taking. This index consists of four scales; the Perspective Taking scale (PT), the Fantasy scale (FS), the Empathic Concern scale (EC), and the Personal Distress scale (PD). The PT scale is used to assess the extent to which the participants spontaneously adopt the psychological point of view of others (e.g. "I try to look at everybody's side of a disagreement before I make a decision"). The FS scale measures whether individuals are likely to transpose themselves into imaginary settings (e.g. "I really get involved with the feelings of the characters in a novel"). EC evaluates feelings of concern for others (e.g. "When I see someone being taken

advantage of, I feel kind of protective towards them"). The PD scale assesses the extent to which individuals feel personally distressed when observing others who are experiencing negative situations (e.g. "When I see someone who badly needs help in an emergency, I go to pieces"; Davis, 1983). In the current study the focus lies on the PT scale. The reliability and validity of the Dutch version of the IRI proved the overall internal consistency, construct validity, and factor structure to be sufficient and indicate that the scales are valuable when measuring the perception of people about their empathic tendencies (De Corte et al., 2007). However, an adaptation of the scales was used for all participants, because the current version is not suitable for children due to the relatively difficult words used in this questionnaire. In the current sample the reliability is sufficient, Cronbach's α ranged from .68 to .81 for the adults and from .47 to .65 for the children (Table 2).

Background information questionnaire. Background information about the participants was collected using a questionnaire designed for this study. It contains questions about gender, age, nationality and education of the participants, mother tongue, languages used in daily life, reading problems, and problems with the Dutch language.

Intellectual ability. The participants were screened with Raven's Standard Progressive Matrices (Raven, Raven & Court, 2000) to assess Spearman's *g*, or general intelligence. Raven's SPM consists of 60 items, categorized in five sets (A through E) of 12 items each. Each item consists of a figure from which one part is missing. Either six (sets A and B) or eight (sets C through E) pieces are displayed below the figure from which the participant has to select the one piece that completes the figure. The different sets and items within a set increase in difficulty. Based on the number of correct items, estimated IQ scores were obtained using international norms (Raven et al., 2000). Participants had 30 minutes to complete the test. Participants with an outlying low score on this test did not qualify to

participate in this study. Hereby group differences in comprehension due to other factors than the ability to take perspective are minimized.

Vocabulary. Vocabulary was assessed with the Peabody Picture Vocabulary Test-III-NL (PPVT-III-NL; Schlichting, 2005). The test is reliable and valid according to the Commissie Testaangelegenheden Nederland (COTAN; Evers, Braak, Frima, & Van Vliet-Mulder, 2009-2011). This test was adapted into a paper-and-pencil format to allow group administration in the child group. Participants were required to select and mark the picture that matched a printed word out of four presented pictures. Participants were presented with 132 (children) or 72 (adults) age-appropriate items taken from the original test: the children were given sets 7 through 17, and adults sets 12 through 17. Participants had 30 minutes to complete the test. The number of correct items for each participant was converted into an age-appropriate norm score, based on a normal population distribution with a mean of 100 and standard deviation of 15 (Schlichting, 2005). As the PPVT-NL was adapted, two aspects had to be taken into account when converting the raw scores into standard scores. First, whereas the original test imposes no time limit, the adapted version sets a time limit of 30 minutes to complete the task. Because all participants finished the task within 30 minutes this modification did not affect the results. Second, whereas the original test uses an individualized criterion for ending the test, in the group administration participants within an age group received the same items and, thus, a uniform end to the test. It is conceivable that this adaptation led to underestimation (if a participant would have received additional items in individual testing) or overestimation (if a participant received items beyond what he/she would have received in individual testing) of a participant's receptive vocabulary score. Underestimation would not pose a problem because the purpose of the vocabulary test was to ensure that participating children were in the normal range with respect to receptive vocabulary and because all participants were able to complete the test within 30 minutes.

Overestimation also did not occur as scoring was stopped when a participant reached items that he/she would not have received in individual testing.

The Three-Minute-Test. The Drie-Minuten-Toets (DMT); Krom, Jongen, Verhelst, Kamphuis & Kleintjes, 2010) was used to assess the decoding capacities of the participants to control for technical reading capacities. Standardized scores on this word-decoding test (DMT; Verhoeven, 1995) were obtained. This test is developed by CITO (*Centraal Instituut voor Toetsontwikkeling*), the Dutch national institute for measurement in education, which also provides national norms for these tests. This test is valid and reliable for primary school children in grades 1 through 6 according to the COTAN (Evers et al., 2009-2011). For adults ceiling effects are expected, but as this test was only used to exclude participants who score below what is expected for their age, this is not problematic.

Reading attitude. In addition the Elementary Reading Attitude Survey (ERAS; McKenna & Kear, 1990) was used to assess the participants' attitude towards reading. This survey consists of 20 4-point Likert items (e.g. "How do you feel about learning from a book"). The survey has good construct validity and good internal consistency; Cronbach's α of the total score ranged from .87 to .89 in children from grade one through six (McKenna & Kear). The ERAS is translated into Dutch and the Garfield pictures that indicate the emotional states ranging from very happy to very upset were replaced by smileys that range from very happy to very sad as these figures are more common to Dutch children. Three raw scores were interpreted: the total score of the first ten items which represent the attitude towards recreational reading, the total of the remaining items which represent the attitude towards academic reading, and a composite score of the total survey. The raw scores were not converted into percentile ranks as the corresponding national ranks are normed for an American sample.

Procedure

Eye tracking and questionnaire data were collected from individual participants in a quiet room at their school (10-12-year-olds), or in the lab (18-25-year-olds). First the reading task was administered. Participants were given the instruction for the reading task. Participants were told that they would be reading incomplete short stories on a computer screen. They were instructed to read for comprehension at a normal pace (no faster or slower than they would normally read stories) and to choose the best final sentence for each story. Participants received three practice stories after they were positioned in front of the screen with their head in the chinrest. Care was taken that all participants understood the task after practice and that all participants were seated comfortably with their head in the chinrest. The instruction and practice items took approximately 15 minutes to complete. The 45 stories were presented all at once consecutively. After each story it was monitored if eye tracking needed recalibration. The story-completion task took approximately 30 minutes to complete.

After reading the stories participants were given a short break with a small snack and something to drink and then completed the DMT, Raven's SPM, PPVT-III-NL, the adapted version of the IRI, the ERAS, and filled out the questionnaire about their background. The procedure for the children was similar, with two exceptions, first, the experimenter asked them the questions from the background information questionnaire before the start of the experiment, second the PPVT-III-NL and Raven's SPM were administered in a group session at their school before the individual experiment took place. At the end of the experiment participants were thanked for their participation and given a small present (10-12-year-olds) or course credit (some 18-25-year-olds).

Results

Demographic Variables

Table 3 shows the demographic variables of the sample for the two age groups separately. Analyses indicated that 10-12-year-olds and 18-25-year olds did not differ with regard to personal distress as measured by the IRI (p = .54), and academic reading attitude as measured by the ERAS (p = .63). They did differ with regard to a standardized vocabulary score, t (64) = 3.44, p = .001, d = .84 (a medium effect). The adults had a higher vocabulary score than children but both groups had sufficient word knowledge, as indicated by PPVT norm-scores ranging from 86 through 125. The groups also differed with regard to word-decoding skills, t(45.48) = 7.53, p < .001, d = 1.87, perspective taking, t(64) = 3.52, p =.001, d = 0.86, fantasy, t(64) = 2.43, p = .02, d = 0.60, and empathic concern, t(64) = 2.0, p = .049, d = 0.49, which is to be expected from a developmental point of view. The finding that the 18-25-year-olds are more inclined to take other people's perspective than children suggests that this ability is further developed among the young adult group.

Story-completion task

All data were analyzed by repeated measures analysis of variance with condition as the within-subject variable and age group as a between-subject factor. Estimated Marginal Means (EMMs) are reported and when the assumption of sphericity is violated the degrees of freedom are corrected using Greenhouse-Geisser estimates of sphericity.

Story-completion Task Accuracy To examine our *first* hypothesis, that adults' increased ability to make inferences during reading would result in higher accuracy on the Story-completion task, we submitted the percentage of correct answers to a repeated measures ANOVA with Age group (10-12-year-olds, 18-25-year-olds) as between subject-factor, and task condition (control, simple perspective-taking, complex perspective-taking) as within subject-factor. As expected, adults outperformed children in the Story-completion task, as

indicated by a significant effect of age group, F (1,64) = 4.22, p = .04, η_p^2 = .06. However, this was a small effect, even though adults chose the most plausible ending more often (M = 0.97; SE = 0.0006) compared to children (M = 0.95; SE = 0.0006) accuracy was high in both age groups (see Figure 2). There were no accuracy differences between the control, simple perspective-taking, and complex perspective-taking conditions (p = .34), and there was no significant age group x condition interaction (p = .69). Thus, accuracy results show that participants did not have difficulties completing the task and both groups perform well above chance level.

Story-completion Task Choice Reaction Times In the following analyses on reaction times all data from incorrectly answered trials are excluded. To test our *second* hypothesis that adults would be faster in choosing the most plausible story ending because of their increased ability to make inferences during reading, and to examine our *third* hypothesis whether participant's choice reaction times in both age groups were influenced by the type of inference that had to be made (i.e. whether perspective taking was needed or not), we submitted participants choice reaction times to a repeated measures ANOVA with age group as between and condition as within subjects factors. A main effect of age group revealed that adult participants were on average 1199 ms faster than the children across all conditions, F(1,64) = 41.53, p < .001, η_p^2 = .39 we examined (see Table 4). In addition, there was a main effect of the type of inference that had to be made on choice reaction times, F (2, 128) = 3.94, p = .022, η_p^2 = .06. Bonferroni corrected post hoc tests indicate that the time to select the most plausible ending was on average 180 ms shorter in the complex perspective-taking condition compared to the control condition, p = .02, but not compared to the simple perspective-taking condition, p = 1 (see Table 5 for means and standard errors). There was no interaction between age group and condition, p = .76 (see Figure 3). Contrary to what we expected,

individuals were fastest in the complex perspective-taking condition, not in the control condition, and the data did not reflect age related differences in perspective-taking ability.

Story-completion Task Story Reading Times To examine possible age and condition effects during reading (*fourth* and *fifth* hypothesis), we calculated total story reading times by adding all fixation durations for each trial, and submitted these story reading times to a repeated measures ANOVA with condition (control, simple perspective-taking, complex perspective-taking) as within subjects factor and age group (10-12-year-olds, 18-25-year-olds) as between subjects factor. As expected, adults read the stories faster compared to children (average difference = 976 ms), $F(1,64) = 12.59$, $p = .001$, $\eta_p^2 = .16$ (see Table 4). There was no age group x condition interaction, ($p = .66$). However, a main effect of perspective-taking condition, $F\ (1.74,111.56) = 72.89$, $p < .001$, $\eta_p^2 = .53$, was present (see Figure 4). Post-hoc Bonferroni corrected comparisons revealed that participants spent on average 837 ms more time to read the stories in the complex perspective-taking condition compared to the simple perspective-taking condition ($p < .001$), and 516 ms more in the complex perspective-taking condition compared to the control condition($p < .001$; see Table 5). In addition, participants took on average 321 ms longer to read stories in the control condition compared to the simple perspective-taking condition ($p < .001$).

Discussion

In this study the role of age-related changes concerning perspective taking skills in the ability to make causal inferences that require social cognition was explored. Adults were better and faster at making inferences in general. Both 10-12-year-olds and 18-25-year-olds are faster in making an inference that required them to infer how one of the story characters would react to the intentions, thoughts, or feelings of the other story character. The 18-25-year-olds read the stories faster across all conditions that we examined. In addition, both 10-

12-year-olds and adults revealed the longest reading times in the most difficult condition (complex perspective-taking condition).

As anticipated, the results imply that both 10-12-year-olds as well as 18-25-year-olds possess the ability to make inferences and that this ability improves with age, as indicated by the fact that both groups performed well above chance level and that the 18-25-year-olds were more accurate and faster at making inferences (Hypothesis 1 and 2 were confirmed). This is in line with previous research that indicated that children who are approximately 7-years-old are able to draw inferences from texts, but that this ability develops with age (Casteel & Simpson, 1991) and that 15-year-olds are better at drawing inferences than 6-year-olds (Barnes et al., 1996).

Contrary to what we expected, both 10-12-year-olds and 18-25-year-olds were faster in selecting the most plausible ending in the complex perspective-taking condition (when the perspective of one of the story characters and the reaction of that character on the other character was needed) compared to the control condition (when it was not necessary to take the perspective of the story characters; Hypothesis 3 was rejected). Although inferences that require individuals to take the interaction between two story characters into account simultaneously are processed differently compared to stories in which it is not necessary to take the perspective of the story character, the offline measure suggests that in the most complex condition it is actually the least difficult to make the inference needed to choose the most plausible ending.

The results of the online measure challenge the conclusion that could be drawn based on the Story-completion task choice reaction times. Across all conditions adults read faster than children (Hypothesis 4 was confirmed), however, all participants were slowest to read the complex perspective-taking stories (Hypothesis 5 was confirmed). Stories that require people to imagine how one of the story characters would react to the intentions, thoughts, or

feelings of the other story character are more difficult to process. Knowledge of situations in which social cognition plays a role is needed and more complex inferences need to be made in order to be able to choose the most plausible ending after reading the story. This finding could be taken to suggest that these longer reading times ultimately facilitate comprehension (as reflected in decreased time to choose the most plausible ending offline). University students are able to drawn inferences online, during reading of the text, and these inferences aid the reader in offline tasks (O'Brien, Shank, Myers, & Rayner, 1988). The longer reading times in the complex perspective-taking condition resulted in faster offline choices, this was also true for the 10-12-year-olds. O'Brien and Myers (1985) argued that more difficult texts will result in a stronger memory trace. In their study participants performed better during recall of texts in which a difficulty emerged (an unpredictable word). The undergraduate students recalled text components before a difficulty better than components from texts that did not contain an unpredictable word. O'Brien and Myers proposed that readers reprocess parts of the text when encountering a difficulty. In stories in which complex-perspective taking is important for comprehension of the text readers need to take the perspective of a story character into account, as well as the reaction of that character on the other character that is present in the story. This requires more social-cognitive abilities and takes longer to process. O'Brien and Myers suggested that the slower reading process when encountering a difficulty results in more elaborative processing in order to understand the text. In our study, this longer processing time results in that individuals can better anticipate on what is coming with regard to the interaction between story characters; this could explain why both 10-12-year-olds as well as 18-25-year-olds are faster in choosing the most plausible ending in the complex perspective-taking condition for which we anticipated them to be the slowest. They have processed the complex perspective-taking narratives more deeply and this facilitates their comprehension of these stories.

In addition, this demonstrates the importance of examining the reading process and not just the result of the reading process. Offline measures alone would probably not have provided us with the insight that stories in which complex-perspective taking is needed to understand the situation in the text are more difficult than stories in which no perspective-taking, or simple-perspective taking, is needed.

We cannot provide a straightforward answer to whether age-related perspective taking skills influence the ability to make causal inferences that require social cognition. The 18-25-year-olds yielded a higher score with regard to the extent to which they spontaneously adopt the psychological point of view of others than the 10-12-year olds. The ability to take perspective was more developed in the adult group. Additionally they outperformed the 10-12-year-olds on the offline and online reading measures. However, the pattern of results for these measures was the same for children and adults; there were no interactions between age group and condition. This could be due to the fact that the stories were designed so they would be appropriate for children, therefore the materials might have been relatively easy for young adults. Additionally, it could also be that there are no differences at a behavioural level, whilst there still might be age-related differences at a neurological level.

Sebastian and colleagues (2012), in a study on affective Tom (when you have to infer how people are feeling) and cognitive Tom (when you have to infer beliefs and motivations) also did not find an interaction between age group (11-16-year-olds and 24-40-year-olds) and cartoon type (Affective ToM, Cognitive Tom, Physical Causality) for reaction times. The neuroimaging data on the other hand did reveal age-related differences with regard to affective ToM. Blakemore (2008) also suggests that the ‘social brain’ continues to develop between adolescence and adulthood.

Other, more sophisticated, online measures might provide us with more insight into what is different about the stories in which complex-perspective taking abilities are needed in

order to comprehend the text. What happens during the longer reading times can be explored with more depth. It can be expected, for example, that the longer reading times are due to rereading previous parts of the text. To what parts of the text readers regress might be different for adults than for children as they do not possess perspective-taking abilities to the same extent, adults might for example regress back to words in these types of stories in a more consistent way than children during the reading process, and this leaves possibilities to explore this process. In future research other measures to capture perspective taking ability of the participants could be used. The Director Task (Keysar et al., 2000), for example, could measure the ability to perform complex perspective-taking tasks. However, at this moment the IRI is one of few perspective taking measures that can be easily administered to adults and, with some adaptations, to children as well. When these two measures would be used in conjunction they would be able to inform us on the ability to spontaneously adopt another's point of view as well as the ability to perform tasks in which perspective-taking is needed.

To conclude, stories in which the interaction between two protagonists has to be taken into account are processed differently compared to stories in which no social-cognitive information is needed and different compared to when one only has to take the perspective of one story character. They are more difficult to process. This corresponds to the social complexity of daily life where situations in which interactions between persons take place are more difficult to understand than situations in which the perspective of one person needs to be taken into account, or when perspective taking is not necessary. Grasping various perspectives in a narrative is essential for comprehension of narrative texts. Perspective-taking ability improves throughout development and the results presented here demonstrate that narratives in which perspective-taking is crucial for comprehension are more difficult to process, even for adults, but are better represented in the situation model readers construct.

References

Barnes, M. A., Dennis, M., Haefele-Kalvaitis, J. (1996). The effecys of knowledge availability and knowledge accessibility on coherence and elaborative inferencing in children from six to fifteen years of age. *Journal of Experimental and Child Psychology, 61*, 216-241.

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

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*(2), 241-251.

Blakemore, S. J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience, 9*(4), 267-277.

Casteel, M. A., & Simpson, G. B. (1991). Textual coherence and the development of inferential generation skills. *Journal of Research in Reading, 14*(2), 116-129.

Choudhury, S., Blakemore, S-J., & Charman, T. (2006). Social cognitive development during adolescence. *Social Cognitive and Affective Neuroscience, 1*(3), 165-174.

Davis, M. H.(1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85-104.

Davis, M. H. (1983). Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach. *Joural of Personality and Social Psychology, 44*(1), 113-126.

De Corte, K., Buysse, A., Verhofstadt, L. L., Roeyers, H., Ponnet, K., & Davis, M. H. (2007). Measuring empathic tendencies: reliability and validity of the Dutch version of the interpersonal reactivity index. *Psychologica Belgica, 47*(4), 235-260.

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

Evers, A., Braak, M.S.L., Frima, R.M., & Van Vliet-Mulder, J.C. (2009-2011). *COTAN Documentatie*. Amsterdam: Boom test uitgevers.

Gurucharri, C., & Selman, R. L. (1982). The development of interpersonal understanding during childhood, preadolescence, and adolescence: A longitudinal follow-up study. *Child Develoment, 53*(4), 924-927.

Hulme, C., & Snowling, M. J. (2009). *Developmental Disorders of Language Learning and Cognition*. Oxford, UK: Wiley-Blackwell.

Kendeou, P., Van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101*(4), 765-778.

Keysar, B., Barr, d. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: the role of mutual knowledge in comprehension. *Psychological Science, 11*, 32-38.

Kintsch, W. (1998). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163-182.

Kordes, J., Bolsinova, M., Limpens, G., & Stolwijk, R. (2013). *Resultaten PISA-2012: Praktische kennis en vaardigheden van 15-jarigen.* Arnhem: Cito

Krom, R., Jongen, I., Verhelsy, N., Kamphuis, F., & Kleintjes, F. (2010). *DMT en AVI. Groep 3 tot en met 8*. Arnhem: Cito.

Lapsley, D. K. (2006). Moral stage theory. In M. Killen, & J. Smetana (Eds.), *Handbook of Moral Development* (pp. 37-66). Mahwah, NJ: Lawrence Erlbaum

Mano, Y., Harada, T., Sugiura, M., Saito, D.N., & Sadato, N. (2009). Perspective-taking as part of narrative comprehension: A functional MRI study. *Neuropsychologia*, *47*, 813-824.

Mar, R. A., DeYoung, C. G., Higgins, D. M., & Peterson, J. B. (2006). Self-liking and self-competence separate self-evaluation from self-deception: Associations with personality, ability, and achievement. *Journal of Personality*, *74*, 1047-1078.

Mar, R. A., Oatley, K., & Peterson, J. B. (2009). Exploring the link between reading fiction and empathy: Ruling out individual differences and examining outcomes. *Communication, 34*, 407-428.

Marotto, C. M., Barreyro, J. P., Cevasco, J., & Van den Broek, P. (2011). Generation of emotional inferences during text comprehension: Behavioral data and implementation through the Landscape Model. *Escritos de Psicologia, 4*(1), 9-17.

McKenna, M. C., & Kear, D. J. (1990). Measuring attitude toward reading: A new tool for teachers. *The Reading Teacher*, *43*(9), 626-639.

Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 international results in reading.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

O'Brien, E. J., & Myers, J. L. (1985). When comprehension difficulty improves memory for text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 12-21.

O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 410-420.

Raven, J., Raven, J. C., & Court, J. H. (2000, updated 2004). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Section 3: The Standard Progressive Matrices. San Antonio, TX: Harcourt Assessment.

Sebastian, C. L., Fontaine, N. M., Bird, G., Blakemore, S. J., De Brito, S. A., McCrory, E. J. P., & Viding, E. (2012). Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. *Social Cognitve and Affective Neuroscience, 7*, 53-63.

Schlichting, L. (2005). *Peabody Picture Vocabulary Test-III-NL. Handleiding.* Amsterdam, The Netherlands: Harcourt Test.

Van den Broek, P. (1995). A 'landscape' model of reading comprehension: Inferential processes and the construction of a stable memory representation. *Canadian Psychology, 36*(1), 53-54.

Van den Broek, P., Risden, K., Fletcher, C. R., & Thurlow, R. (1996). A 'landscape' view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In B. K. Britton & A.C. Graesser (Eds.), *Models of understanding text* (pp. 165-187). Hillsdale, NJ: Erlbaum.

Van den Broek, P., & White, M. J. (2012). Cognitive processes in reading and the measurement of comprehension. In Espin, C. A., McMaster, K. L., Rose, S., & Wayman, M. M. (Eds.). A measure of success: *The influence of curriculum-based measurement on education* (pp. 329-339). Minneapolis, MN: University of Minnesota Press.

Van der Graaff, J., Branje, S., De Wied, M., Hawk, S., Van Lier, P., Meeus, W. (2014). Perspective taking and empathic concern in adolescence: Gender differences in developmental changes. *Developmental Psychology, 50*(3), 881-888.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123* (2), 162-185.

Table 1

Example of the Stimuli

Condition	Example				
	Introduction	Situation	Action	Plausible	Less plausible
Control	Mary and Claire are at the bus stop.	It is snowing.	The roads are very slippery.	The bus is way too late.	The bus is exactly on time.
Simple	Peter and Anne are ice-skating.	Peter loses his balance.	Peter falls really hard.	Peter starts to cry.	Peter keeps skating.
Complex	Tom and his friend John are walking on the street.	The ice cream truck drives by.	John points at the truck.	Tom buys two ice creams and gives one to John.	Tom buys two ice creams and eats them both.

Table 2

Reliability of the Interpersonal Reactivity Index in the Current Sample

Scale	Cronbach's α	
	adults	kids
Perspective Taking	.73	.56
Fantasy	.81	.61
Empathic Concern	.68	.65
Personal Distress	.78	.47

Table 3

Means and Standard Deviations for Scores on the IRI, PPVT, DMT, and ERAS

Variable	*M* (*SD*) 10-12-year-olds	18-25-year-olds
IRI perspective taking	3.15 (0.57)	3.62 (0.50)
IRI fantasy	3.39 (0.67)	3.77 (0.60)
IRI empathic concern	3.38 (0.57)	3.63 (0.45)
IRI personal distress	3.03 (0.58)	2.94 (0.61)
PPVT norm score	101.16 (9.17)	108.32 (7.75)
DMT	77.88 (15.65)	178.56 (7.72)
ERAS academic reading	27.50 (4.33)	28.03 (4.44)
ERAS recreational reading	28.53 (5.90)	33.32 (5.81)

Note. IRI = Interpersonal Reactivity Index; PPVT = Peabody Picture Vocabulary Test; DMT = Drie Minuten Toets, ERAS = Elementary Reading Attitude Survey.

Table 4

Estimated Marginal Means of Accuracy, Reaction Time, and Reading Time as a Function of Condition and Age Group

Condition	Accuracy		Reaction Time in ms		Total Duration in ms	
	adults	kids	adults	kids	adults	kids
Complex	0.96	0.95	3501.92	4699.38	5094.17	6137.21
	(0.11)	(0.11)	(136.57)	(140.78)	(223.49)	(230.36)
Simple	0.97	0.94	3535.04	4784.60	4316.66	5240.63
	(0.09)	(0.10)	(146.08)	(150.58)	(181.61)	(187.20)
Control	0.97	0.96	3705.90	4856.95	4619.45	5581.68
	(0.08)	(0.09)	(139.96)	(141.17)	(191.67)	(197.57)
Total	0.97	0.95	3580.96	4780.31	4676.76	5653.17
	(0.006)	(0.006)	(129.56)	(133.57)	(191.64)	(197.54)

Note. Standard errors are in parentheses.

Table 5

Estimated Marginal Means of Accuracy, Reaction Time, Number of Fixations, and Reading Time as a Function of Condition

Condition	Accuracy	Reaction Time in ms	Total Duration in ms
Complex	0.96 (0.008)	4100.65 (98.07)	5615.69 (160.48)
Simple	0.95 (0.007)	4159.82 (1053.74)	4778.64 (130.41)
Control	0.96 (0.005)	4281.43 (104.90)	5100.56 (137.63)

Note. Standard errors are in parentheses.

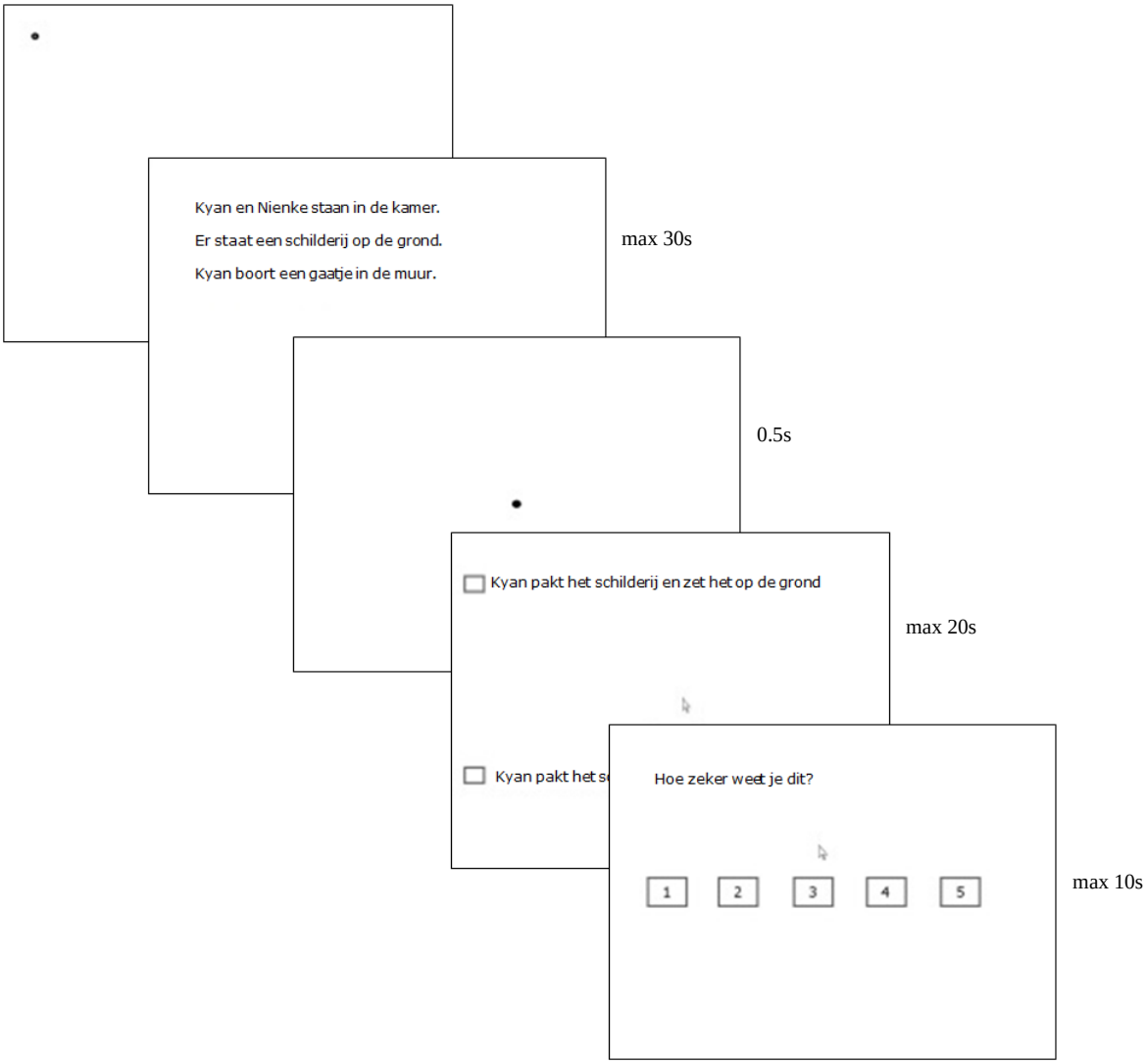

Figure 1. Trial Example. After a drift-check the story appeared. Participants read the story and subsequently pressed the left mouse key to advance to a fixation dot (500ms). If they did not press within 30 seconds, the fixation dot appeared automatically. Hereafter the target sentences appeared and participants selected one of them (20 s time limit) and following this, participants were asked to rate how certain they were about their choice (10s time limit).

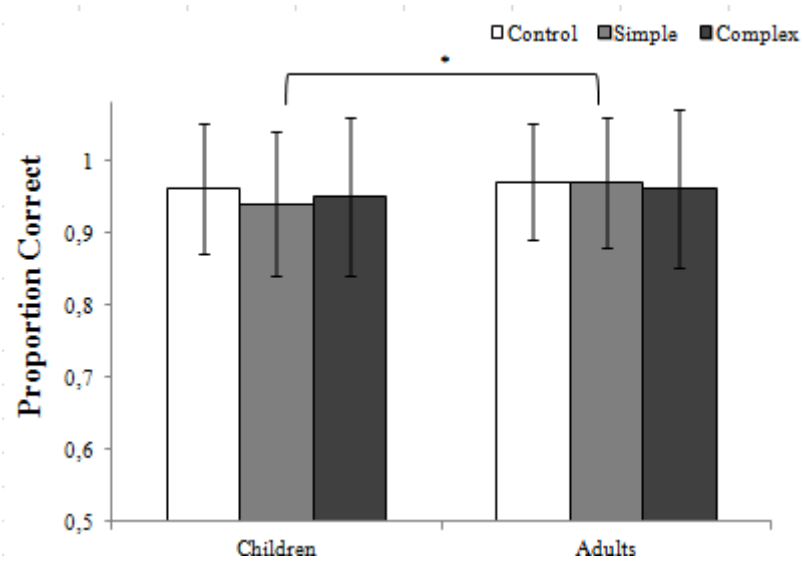

Figure 2. Mean proportion correct for each condition. The 18-25-year-olds outperformed the 10-12-year-olds. There were no differences between the three conditions on accuracy. Standard errors are represented in the figure by the error bars attached to each column.

* $p < .05$.

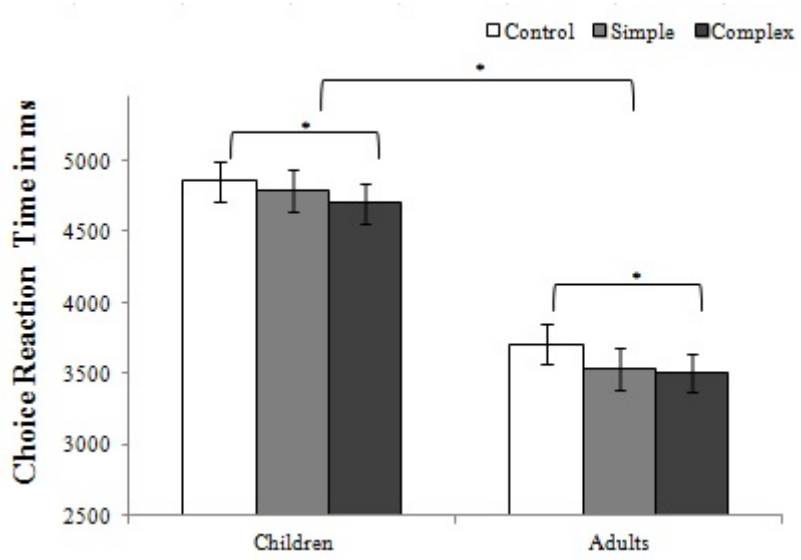

Figure 3. Mean task choice reaction time (ms). The 18-25-year-olds were faster than the 10-12-year-olds. Participants were faster in the complex perspective-taking condition compared to the control condition. There was no interaction between age group and condition. Standard errors are represented in the figure by the error bars attached to each column.

* $p < .05$.

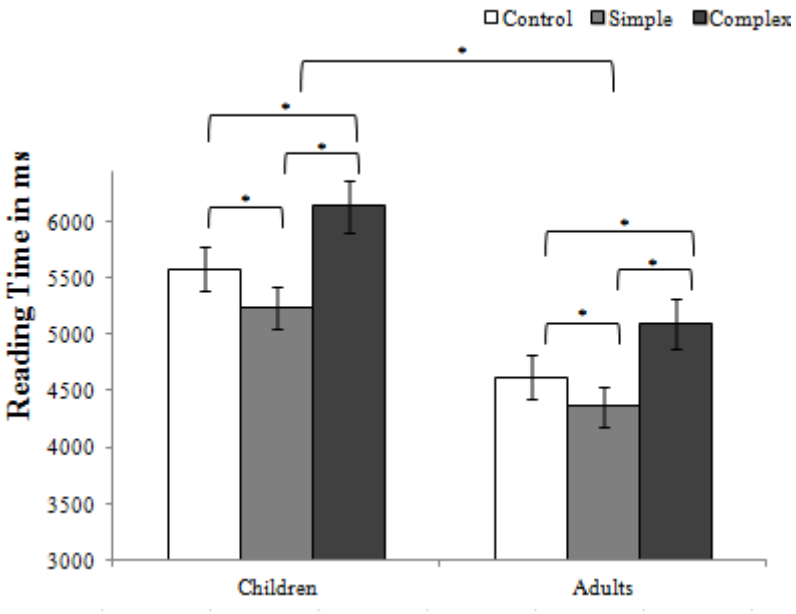

Figure 4. Mean Story-completion task reading time (ms). The 18-25-year-olds were faster than the 10-12-year-olds. Participants were faster in the complex perspective-taking condition compared to the control condition and compared to the simple perspective-taking condition. They were also faster in the simple perspective-taking condition compared to the control condition. There was no interaction between age group and condition. Standard errors are represented in the figure by the error bars attached to each column.

* $p < .05$.