The Spanish Journal of Psychology (2022), **25**, e26, 1–13.

doi:10.1017/SJP.2022.22

CAMBRIDGE UNIVERSITY PRESS

How Chronological Age, Theory of Mind, and Yield are Interrelated to Memory and Suggestion in Young Children

Nieves Pérez-Mata, Amparo Moreno, Margarita Diges and Miriam Peláez

Universidad Autónoma de Madrid (Spain)

Abstract. We investigated the interrelations between chronological age, theory of mind (ToM), Yield (as a measure of individual suggestibility), memory and acceptance of experimental suggestion in a sample of children between 3 and 7 years old ($N = 106$). One week after participants interacted with 'a Teacher', they were asked to recall activities carried out with the Teacher (direct experience) and the contents of a story read to them by the Teacher (indirect experience). Data were examined with an analysis of developmental trajectories, which allows establishing the predictor value of socio-cognitive developmental factors regardless of participants' chronological age. It also estimates predictor values in interaction with the age and determines whether age is the best predictor for performance. As in previous research, results showed that chronological age was the main predictor of memory performance, both for direct experience (i.e., activities performed) and indirect experience (i.e., contents of the story). However, ToM and Yield, together with participants' ages, modulated their acceptance of the external suggestions received (presented only once, one week after the event). A turning point was observed at age 4.6. Below this age, the greater the mentalist skills (higher ToM), the lower was the vulnerability to external suggestion. Still, children below this age characterized individually as being suggestible (Yield medium or high) were more vulnerable to suggestion the younger they were. Thus, developmental socio-cognitive factors might modulate young children's vulnerability to external suggestions, even if received only once.

Received 5 May 2022; Revised 21 September 2022; Accepted 27 September 2022

Keywords: analysis of developmental trajectories, chronological age, memory, suggestibility, theory of mind

At present, there is a significant number of accusations of alleged child sexual abuse based solely on the child's statement (Azzopardi et al., 2019; Smith et al., 2018). Thus, interest in understanding how suggestions affect memory in young children has increased because of its relevance in the forensic field (Ceci & Bruck, 2006; Principe et al., 2014). Performance in memory tasks improves between ages 3 and 6, and, at the same time, vulnerability to suggestion decreases. Although evidence for an 'age-memory-suggestion' relationship has been accumulating over decades (Brainerd & Reyna, 2012; Bruck & Ceci, 1999; Bruck & Melnyk, 2004; Malloy & Quas, 2009; Principe et al., 2014), an explanation based solely on *chronological age* is still lacking because other socio-cognitive factors, which are imperfectly correlated with age, also play a role in developmental differences.

For this reason, researchers in this field have been interested in understanding how developmental socio-cognitive factors, in addition to age, underlie the 'age-memory-suggestion' relationship (for an updated

Correspondence concerning this article should be addressed to Nieves Pérez-Mata. Universidad Autónoma de Madrid. Facultad de Psicología. Departamento de Psicología Básica. Calle Iván Pavlov, No. 6, Ciudad Universitaria de Cantoblanco. 28049 Madrid (Spain).

E-mail: nieves.perez@uam.es

The authors thank Soledad Herrera, María López Rodil and Nuria Malibrán for their help in the collecting and coding the data; Karina Solcoff assisted in the preparation of the material. We also thank Dr. Pilar Aivar for her suggestions and comments prior to submission. Finally, special thanks to parents, children and school staff of the two schools - C. E. I. P. Vicente Aleixandre in Torrejón de Ardoz (Madrid) and E. I. Bärber Inhelder in Madrid - who participated in the present study.

Conflicts of interest: None.

Funding statement: This work was supported by the Transfer Knowledge Project 088501–FUAM (Foundation of the Universidad Autónoma de Madrid) co-directed by the first and third authors, and the National Research SEJ2004–07655 financed by Dirección General de Investigación Científica y Técnica (DGICYT).

Data Sharing: The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

How to cite this article:
Pérez-Mata, N., Moreno, A., Diges, M., & Peláez, M. (2022). How chronological age, theory of mind, and yield are interrelated to memory and suggestion in young children. *The Spanish Journal of Psychology*, *25*. e26. Doi:10.1017/SJP.2022.22

review see Klemfuss & Olaguez, 2020). In this context, our study examines the predictive value of two socio-cognitive factors -*theory of mind* (ToM) and *individual suggestibility* (measured by Yield), defined as individual characteristic- in memory performance and acceptance of experimental suggestions in young children. Specifically, we want to determine to what extent the age of the participants could modulate the predictive value of those two factors.

ToM, Memory, and Suggestion Acceptance in Children

According to Perner (1991), ToM involves comprehending how mental representations are understood as such. People do not have copies of reality, but they represent it. Therefore, mind and reality are, to a certain extent, independent. An important change in the development of representational skills occurs between the ages of 3 and 6, and this change may be important for memory and suggestion.

For more than two decades, Perner has been analyzing the relationship between episodic memory (in an autobiographical sense) and ToM (Perner, 2000; Perner & Ruffman, 1995). Perner has proposed that ToM is a prerequisite to mentally represent episodic experiences lived in the past. Thus, children with higher mentalist skills (i.e., with a more developed ToM) could generate richer episodic representations of memory.

In the same vein, following Tulving's proposal (1985), Perner et al. (2007) have revised the defining features of episodic memory and, hence, the difference between *recalling* the past and *knowing* it. According to their theory, only a direct experience of events allows reviving them later on. In contrast, indirectly experienced (known) events cannot result in that kind of recall. Following this distinction, in the present study children were asked to *recall* a series of activities (a direct experience) performed with an adult ('a Teacher') in interaction during the experimental session a week ago. This was compared with *knowing* the narrative contents of a story that the Teacher told the child (which is closer to an indirect experience) during that.

It has been found that children between 3 and 11 years old are able to give more than half of the details provided in a previously told story (indirectly experienced) when they are asked about its content immediately or one week later (Caprin et al., 2016; Elischberger, 2005; Gobbo et al., 2002). Interestingly, when they must remember an experience directly lived, they recall almost the whole event in a free recall task (Ornstein et al., 1992; Principe et al., 2006). Therefore, it seems that the details from events directly experienced are more accessible for children than the contents from a story which was only read to them.

Moreover, Perner (2000) linked ToM development and richness of episodic memory with children's resistance to suggestion. That is, children with higher ToM abilities would be better able to correctly discriminate between a memory coming from a really lived experience and a memory coming from a false external suggestion (see also, Bright-Paul et al., 2008; Perner et al., 2007; Scullin & Bonner, 2006; and Welch-Ross, 2000).

Unfortunately, a very inconsistent relationship has been found between mentalist skills and acceptance of suggestion in previous studies. This inconsistency is most likely due to methodological differences between studies (Klemfuss & Olaguez, 2020), involving a varied range of ToM tasks and different number of tasks (one, two, five, or six). Also, the effect of task order is not clear regarding the representational skills needed to better resist suggestion.

One of the aims of the present study was to more closely explore the relationship between ToM, memory and acceptance of suggestion. For that, we used the Wellman and Liu's (2004) graduated Scale of ToM. This Scale includes six tasks with increasing difficulty, from *diverse desires to real-apparent emotion*, and it has revealed a robust sequence of understanding through hundreds of preschoolers from different countries and cultures (USA, Canada, Australia, Germany, Italy, Japan, and so on, e.g., Wellman, 2012; Westra & Carruthers, 2017). The use of this Scale is expected to provide a more sensitive measure of ToM, compared to when only one or two independent tasks are used and dichotomous scores are obtained (pass or fail the task). A more sensitive and robust measure of ToM would be more informative in determining the possible association of ToM with memory and suggestion acceptance.

Yield, Memory, and Suggestion Acceptance in Children

The term *suggestibility* refers to an individual characteristic that makes a person more likely to respond in a particular way to suggestion (Gudjonsson, 2003). Gudjonsson assumed that cognitive and social factors underline people's suggestibility. He designed the Gudjonsson Suggestibility Scale 1 (GSS1) and Scale 2 (GSS2) to evaluate the answers to suggestive questions (*Yield*) and response changes to negative feedback from previous responses (*Shift*). These scales have been adapted for use with children. In our study, we used the Spanish adaptation (Diges et al., 2010) of the *Book Suggestibility Scale for Children* (BSSC; Melinder et al., 2005). It is important to emphasize that Yield and Shift imply two very different socio-cognitive factors. Yield indicates the tendency of a person to spontaneously incorporate into memory a false suggestion given by an interviewer, which mainly depends on cognitive factors; Shift, however, indicates the willingness of the

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press

person to abandon their previous belief in response to the challenge posed by an interviewer, and is linked to social factors related with the confidence that the person has in the contents of his or her own memory.

Research with children has shown a consistent positive relationship between Yield (as a measure of individual suggestibility) and acceptance of details suggested in an experimental session (Karpinski & Scullin, 2009; Quas et al., 2005; Scullin & Bonner, 2006). Thus, the higher Yield score of the participant, the greater the acceptance of experimental suggestions. However, the pattern of results for Shift has been very inconsistent (Diges et al., 2010; Karpinski & Scullin, 2009; Quas et al., 2005; Scullin & Bonner, 2006), which has led researchers to put aside these values. In the present study, we will only use the Yield scores obtained by our participants as a measure of individual suggestibility.

Another matter of interest is how the individual suggestibility could affect the recall of an event. A link between a weak memory representation and an increased vulnerability to suggestion has been repeatedly proposed (e.g., Ceci et al., 1988; Malloy & Quas, 2009; Principe et al., 2014; Tousignant et al., 1986). Thus, when people with weak memory traces retrieve information from past experiences, they may infer events that did not take place or add details that were later suggested to them. This would mean that people with a higher individual suggestibility would retain memories that could, potentially, be more vulnerable to external suggestions. Caprin's et al. results (2016) are consistent with this proposal, but in Klemfuss and Olaguez's review (2020) a clear association between suggestibility and memory was not found, which could be due to the high methodological variability among the studies reviewed.

The Present Study

A sample of young children between 3 and 7 years old individually interacted with an unknown adult ('The Teacher'). One week later, a different adult (an interviewer) asked participants about the activities that they carried out with 'The Teacher' and also about the contents of a story that 'The Teacher' had told them. Furthermore, once during the interview the interviewer suggested false information about some of the activities performed. This false information was provided in order to examine: The interrelations between ToM, Age, memory and suggestion, on the one hand, and the interrelations between Yield, Age, memory and suggestion, on the other hand.

Repeated introduction of misinformation has been shown to be successful in obtaining a powerful and irreversible effect of suggestion in preschoolers (Ceci & Bruck, 2006; Garven et al., 1998; Otgaar et al., 2019; Peláez et al., 2019; Sauerland et al., 2019). In fact, when repeated suggestions are combined with other suggestive techniques (e.g., social reinforcement, 'others have told me…', etc.), even older children are very vulnerable to suggestion (Garven et al., 1998; Wood & Garven, 2000). However, in this study we decided to provide the suggested information only once after a one-week delay. This is unlike the more usual procedures, in which false information is repeatedly suggested during an interview or over the course of several interviews. We introduced this change because we were interested in examining whether young children with a higher mentalist skill or a lower individual suggestibility value are able to resist such a suggestion when it is given in a less persistent manner (i.e., only once during a single delayed interview).

If the relationship proposed by Perner (2000; Welch-Ross, 2000) between mentalist skill and the emergence of episodic (autobiographical) memory is correct, then ToM would predict memory performance for the activities performed with 'the Teacher' (details directly experienced), even beyond chronological age. That is, we expect that the higher the mentalist skill of the participant, the better the result on memory performance. On the other hand, we expect that the chronological age would be a good predictor of memory for the narrative content of the story read to the participants by the Teacher (details not directly experienced). In this case, the older the participant, the better the memory of the details of the story. Furthermore, since mentalist skill has also been linked to resistance to suggestion (Perner, 2000), we also expect that ToM would predict vulnerability to the experimental suggestion. In this case, a lower vulnerability to suggestion was expected for participants with higher ToM scores.

Additionally, according to previous results (Karpinski & Scullin, 2009; Quas et al., 2005; Scullin & Bonner, 2006), a higher individual suggestibility, measured by Yield, would be associated with higher acceptance of the experimental suggestion, beyond participant's chronological age. Finally, we were interested in exploring the potential relationship between Yield and memory performance as well. In this case, no hypothesis was advanced due to the inconsistent pattern of results previously found (Klemfuss & Olaguez, 2020).

To analyze the data, we applied an analysis of developmental trajectories (Thomas et al., 2009), which is detailed below.

Method

Participants

One hundred and six children (50 boys and 56 girls) between the ages of 3 and 7 (range from 39 to 87 months;

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press

M = 61.95, SD = 16.34) took part in the study. Participants were recruited from two schools situated in Madrid that offered early childhood education. Informed consent was obtained from children's parents and/or guardians before they participated in the study.

Materials and Variables

The materials and variables used in the study were as follows:

The adapted Spanish version of the *Book Suggestibility Scale for Children* (BSSC) was used (Diges et al., 2010). It included a reading of a story, with a duration of approximately 15 minutes, complemented with 16 colored drawings showing characters and actions performed by them. After reading the story[1], the child was assessed on ToM tasks, which took between 3 and 6 minutes (M = 3.5 minutes), depending on the number of ToM tasks they completed; and then 18 yes/no questions about the story were asked (14 were misleading questions intermingled with 4 true questions). The answers to the 14 misleading questions gave the Yield score on the Scale. Immediately after answering the questions, the interviewer said to the child: 'Oh! I am afraid some of your answers are wrong, so I am going to ask you the questions again to see if you get them right this time' (negative feedback), and the questions were asked again. The changes in the answers gave the Shift score on the Scale. Due to the inconsistency of the Shift measure, and following the recommendations of previous research (Karpinski & Scullin, 2009; Quas et al., 2005; Scullin & Bonner, 2006), only the Yield score was used as a measure of individual suggestibility (ranging from 0 to 14).

The adapted Spanish version of the *ToM* Scale devised by Wellman and Liu (2004) was used (Diges et al., 2010), which includes six increasingly difficult tasks (*diverse desires*, *diverse beliefs*, *knowledge access*, *explicit false belief*, *belief emotion*, and *real-apparent emotion*). If the child performed the first task (*diverse desires*) correctly, he or she was given the second one (*diverse beliefs*) and so forth, until reaching the last one (*real-apparent emotion*). If the child failed two consecutive tasks, the assessment stopped, and his or her ToM score was the number of tasks correctly completed before the two consecutive fails (see Wellman & Liu, 2004). Therefore, the Tom score ranged from 0 to 6.

Two quantitative measures of *free recall* of Session 1 were obtained: (1) *Recall of activities* was calculated using a guide designed for this purpose that included 37 entries describing the child-teacher interaction activities performed during Session 1 (e.g., 'we put the tape in the recorder'); and (2) *Recall of the Story* was calculated using a guideline provided by one of the Suggestibility Scale's authors (M. H. Scullin). This guideline has 91 entries on the characters, their characteristics and the actions narrated in the story (e.g., 'The aliens helped the boy').

Therefore, there were two scores for the participant's delayed recall: *Activities* in the interaction with 'the Teacher' (direct experience), and *content of the story* (indirect experience). Two independent raters separately coded the free recall reports using both guidelines ('recall of activities' and 'recall of the story'). Each marked entry from the corresponding guideline scored one point. Raters' scores were compared, and inter-rater agreement was reached for each of the reports coded. The participant's final score corresponded with the points in which both raters were in complete agreement. The inter-judge agreement reached was above 80% for both recall tasks.

A *Yes/No recognition* task was also used, which consisted of 20 questions about Session 1. To keep the child focused on the task, the questions were divided into two sets of 10 questions (five true and five misleading in each set). The order of the sets was counterbalanced, and half of the children answered in the opposite order of the other half. These questions were about the activities the child and 'the Teacher' had performed (or not performed) during the first session (e.g., a true question: 'Did you help the Teacher put the tape in the recorder?'; a misleading question: 'Did the Teacher allow you to look through the video camera?'). None of these questions referred to the contents of the story, to clearly separate memory for the details of the story from memory about the activities performed by the child in Session 1. *Hits on recognition* were 'Yes' responses to the 10 true questions, and *suggestion acceptance* were 'Yes' responses to the 10 misleading questions. In both cases, scores ranged from 0 to 10.

The adapted Spanish version of the *Vocabulary Scale* from McCarthy Scales of Children's Abilities was used (MSCA; McCarthy, 2004) in order to examine whether richness of vocabulary could modulate participants' performance. This Scale included two parts: Part I–Pictorial Vocabulary Subscale consisted of nine items (maximum score = 9 points), while Part II–Oral Vocabulary Subscale had 10 items (maximum score = 20 points). Thus, the maximum score for the entire Vocabulary Scale was 29 points.

[1]In the original Scale the free recall task is presented immediately after the reading, but neither the quantity nor the quality measures of the free recall are considered in calculating the individual suggestibility score of the Scale. In addition, we were interested in analysing participants' recall of all the activities carried out in Session 1, including the contents of the story. In this way, both memories (activities and story) were subjected to a delay of one week.

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press

Procedure

The study took place in two sessions led by two different interviewers in a small room in the participant's own school. In Session 1, the interviewer first asked the child his or her name and age, and then asked the child to pass her a folder containing the story (Suggestibility Scale) and proceeded to read it while simultaneously showing the corresponding drawings. After that, the interviewer presented the ToM tasks to evaluate the child's mentalist skills. Next, the questionnaire of the Suggestibility Scale was administrated, and negative feedback was given about his or her performance ('Oh! I am afraid some of your answers are wrong…'). Finally, the whole set of questions was asked again. Session 1 lasted approximately 20–25 minutes.

In Session 2, a week later, another interviewer introduced herself to the child and asked for his or her name and age. Then, she told the child that she wanted to know what he or she had done with the first 'Teacher' the previous week: 'Someone told me that the other day you were here with another Teacher. Do you remember her name? (If the child did not remember it, the interviewer reminded him or her of the Teacher's name). Did you have fun with her? … Therefore, you had a lot of fun with her! Well, now I would like you to tell me everything you did with the Teacher (her name) because I need your help to know everything you did together because I would like to do the same activities with another child tomorrow' (this request led to the free recall of Session 1). If the child did not remember anything, two reminders were given ('I have been told she read you a story', and 'I have been told she showed you some characters'). Once the child had finished his or her recall, the interviewer asked if anything else had happened in the session ('Did anything else happen?'). Then, the first set of 10 yes/no questions were presented, followed by the Pictorial and Oral Vocabulary Subscales. Finally, the child responded to the second block of 10 yes/no recognition questions, and the interviewer thanked the child for participating, adding that he or she had done it very well. Session 2 lasted approximately 15 minutes.

Procedures followed in the study were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975.

Results

Statistical Methods

To test the predictions made, we used an analysis of development trajectories (Thomas et al., 2009). This analysis generates a function that links performance in each of the experimental tasks with a specific factor regarding participants' socio-cognitive development. Thus, the main advantage of this method is that it establishes the predictive value of each socio-cognitive factor independently of the chronological age of the participant and the predictive value of each factor in interaction with the age of the participants. Alternatively, it allows demonstrating whether age is the variable that best explains participants' performance in the experimental tasks used.

In the present study, we were interested in analyzing how our participants' ToM development (i.e., ToM-Low, ToM-Medium, or ToM-High) could affect their performance in the memory tasks used (free recall and recognition) and in the acceptance of false suggestions about activities not carried out during Session 1. Additionally, we were also interested in analyzing how individual suggestibility, measured through Yield (i.e., Yield-Low, Yield-Medium, or Yield-High), could be affecting participants' memory performance and their acceptance of false suggestions. Moreover, the use of developmental trajectory analyses would allow us to determine to what extent the effects of ToM or Yield on performance were independent of the participants' chronological age, or whether performance would be strongly modulated, or even exclusively determined, by age.

To run this type of analysis, all recorded interviews were transcribed and inter-rater agreement was reached on all recall measures. Next, we examined the predictive value of ToM on the dependent measures, alongside or beyond the chronological age of the participants. For this purpose, regression analyses were carried out, in which ToM groups (Low, Medium, High) and the continuous variable *rescaled age*[2] were entered as predictors. The developmental trajectory analysis made possible to further examine the potential modulatory role played by age through an ANCOVA, including age as a covariate. That is, the ANCOVA reported whether the socio-cognitive variable (the group variable) actually influenced the dependent variable, once the influence of the continuous quantitative covariate Age was removed. The same procedure was followed to analyze the predictive value of Yield.

[2]To rescale age, the youngest participant with the worst performance in each critical developmental socio-cognitive variable is first selected. In our case, these included, on the one hand, the youngest child with the lowest ToM score and, on the other hand, the youngest child with the highest Yield score. For these two participants, age in months was transformed into zero, and the ages of the remaining participants of the sample were rescaled by subtracting from their age (in months) the age of the youngest participant with the worst performance in each of those variables respectively. The rescaled age by participants' age on each socio-cognitive variable was then entered into the regression analyses as a continuous variable.

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press

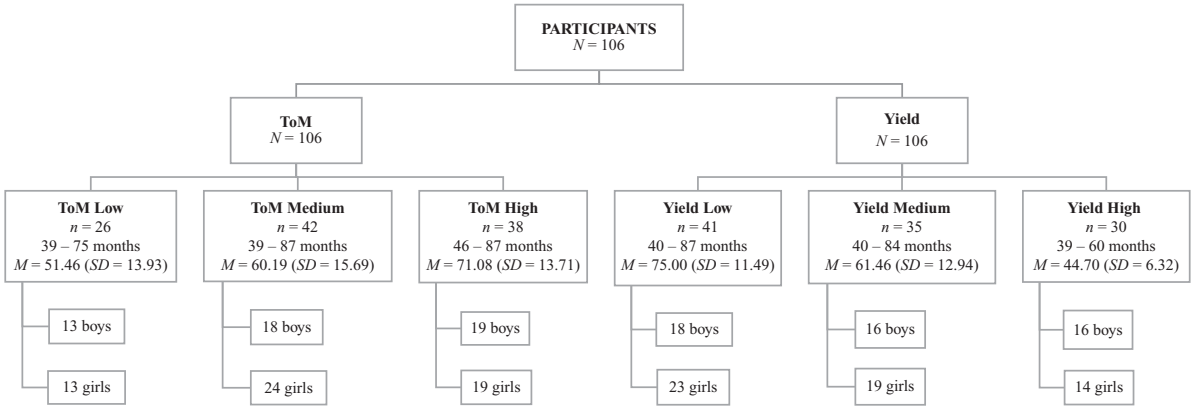

Figure 1. Number of Participants, Sex, and Age (Range, Mean, and Standard Deviation in Months) in each Group of ToM and Yield

In our study, three groups of ToM (Low, Medium, High) were established following the criteria of Wellman and Liu (2004). The 'ToM-Low' group included participants who were only able to correctly resolve the '*diverse desires*' task or did not correctly resolve any task (n = 26). The 'ToM-Medium' group included participants who successfully completed the '*diverse beliefs*' task and/or '*knowledge access*' task (n = 42). The 'ToM-High' group included participants who were able to correctly resolve the '*false belief*' task or overcome one of the two most difficult tasks ('*belief-emotion*' and '*real-apparent emotion*'; n = 38). Regarding *Individual* Suggestibility, given that the Scale does not have normative data, we tried to also generate three balance groups based on the distribution of the participants' Yield scores (maximum score = 14; higher scores indicated higher vulnerability to suggestion as an individual characteristic). The group 'Yield-High' (the most vulnerable participants) included children who reached scores between 10 and 14 (n = 30), the group 'Yield-Medium' included participants with scores between 3 and 9 (n = 35), and the group 'Yield-Low' (the less vulnerable participants) included children with scores between 0 and 2 (n = 41) (see Figure 1).

Next, we describe the results obtained for ToM and Yield in *free recall (activities and story)*, *recognition*, and *suggestion acceptance*. The effect size (η^2) of ToM, Yield, Age, and interactions are reported for each dependent measure following Guéguen et al. (2011). Thus, an effect size was *small* when $\eta^2 < .06$, *medium* when $.06 \leq \eta^2 \leq .13$, and *large* when $\eta^2 \geq .14$. The level of significance for all analyses was established at $p \leq .05$, and for the multiple post-hoc comparisons Bonferroni corrections were used.

Additionally, pictorial and oral vocabulary scores were also introduced into each regression analysis and ANCOVA as a covariate to examine whether these scores played a modulatory role in the expected relationships between the variables. These linguistics measures showed no significant effects in any of the analyses conducted, although participants' vocabulary scores correlated positively with chronological age, ToM, correct recall of activities and the story, and distortions of the story ($rs \geq .24$, $ps < .05$), and negatively with Yield and suggestion acceptance ($rs \geq .45$, $ps < .01$). Therefore, we will no longer discuss these linguistics measures.

ToM as a Predictor of Free recall, Recognition and Suggestion Acceptance

As expected, ToM scores correlated positively with participants' chronological age, $r = .52$, $p < .01$, correct recall of activities, $r = .39$, $p < .01$, and correct recall of the story, $r = .30$, $p < .01$, but negatively with acceptance of suggestion, $r = -.45$, $p < .01$ (Table 1). The results of the regression analyses using ToM and Age[3] as predictor variables for *free recall* (*activities* and *story*[4]) and *recognition* (*hit rate* and *suggestion acceptance*) are presented below (Table 2).

Free recall of activities. Contrary to what was expected, ToM did not predict *correct recall*, and most variance was explained by Age ($R^2 = .38$), $F(2, 98) = 30.03$, $p < .001$ (Table 2). Age effects persisted in the ANCOVA, $F(1, 95) = 33.49$, $p < .001$, $\eta^2 = .26$, with a large effect size. Multiple comparisons indicated that 3-year-old (y-o) children remembered significantly fewer activities than 5-y-o, 6-y-o, and 7-y-o children,

[3]For brevity, Age refers to rescaled chronological age.

[4]Scores from 101 participants were included in the analyses for all the free recall measures. Five participants had to be excluded: One participant for technical reasons, and four participants because they did not provide a free recall of Session 1. Furthermore, for the recall of activities, since the number of ToM tasks completed by each participant could vary, the recall score of activities was also considered in relative terms, thus the proportion of recall was calculated from the total number of ToM tasks presented to each participant. For recognition, scores from the entire sample were included in the statistical analyses (N = 106).

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press

Table 1. *Pearson Correlations between Chronological Age in Months, ToM Scores on the Wellman and Liu's ToM Scale, Yield Scores on the Suggestibility Scale BSSC and the Dependent Measures*

	Age in months	ToM scores	Yield scores	Correct Recall of Activities	Distortions of Activities	Correct Recall of the Story	Distortions of the Story	Hit rate in Recognition	Suggestion Acceptance
Age in months	1.00	.52**	–.78**	.61**	.09	.49**	.22*	–.19	–.58**
ToM scores		1.00	–.55**	.39**	–.10	.30**	.01	.04	–.45**
Yield scores			1.00	–.59**	.02	–.40**	–.22*	.27**	.59**

Note. **Correlation is significant at the .01 level (bilateral). *Correlation is significant at the .05 level (bilateral).

*p*s < .001; and 4-y-o children also differed from 6-y-o and 7-y-o children, *p*s < .01. The remaining comparisons were not statistically significant, *p*s > .30 (Table 3). Furthermore, as the number of *distortions* was very low (it ranged between 0 and 3, see Table 3), the regression model did not reach statistical significance, $F(2, 98) < 2.4$ ($R^2 = .04$; Table 2).

Free recall of the story. For *correct recall*, only Age was a predictor variable of participants' performance ($R^2 = .24$), $F(2, 98) = 15.56$, $p < .001$ (Table 2), as expected. In the ANCOVA, Age showed a large effect size, $F(1, 95) = 21.07$, $p < .001$, $\eta^2 = .18$. Multiple comparisons indicated that 3-y-o children remembered significantly fewer details of the story than 6-y-o and 7-y-o children, *ps* < .01; in addition, 4-y-o and 7-y-o children differed from each other, $p = .01$. However, 5-y-o children did not differ from the other Age groups, *ps* ≥ .2 (Table 3).

With respect to *distortions*, a significant regression model was reached, $F(2, 98) = 5.36$, $p = .01$ ($R^2 = .10$, Table 2); and in the ANCOVA, there was a significant effect of Age, $F(1, 95) = 8.73$, $p < .01$, $\eta^2 = .08$, with a medium effect size. However, multiple comparisons did not achieve statistical significance (*ps* > .75) because the range of distortions was very low for the number of details included in the story (between 0 and 7 distortions from a total of 91 details, Table 3).

Hit rate in recognition. The regression model was marginally significant, $F(2, 103) = 2.71$, $p = .07$ ($R^2 = .05$), but no differences were observed among the groups due to the low variability of the means obtained (Table 3).

Suggestion acceptance. ToM and Age predicted participants' performance ($R^2 = .37$), $F(2, 103) = 30.09$, $p < .001$ (Table 2 and Figure 2A). ANCOVA reaffirmed the predictive value of ToM, $F(2, 100) = 4.89$, $p = .01$, $\eta^2 = .09$, and a larger weight of Age, $F(1, 100) = 32.82$, $p < .001$, $\eta^2 = .25$, but the interaction was not significant, $F = 1.5$. Multiple comparisons showed that participants with Tom-Low accepted significantly more suggestions than the other two groups of ToM, *ps* < .01, but ToM-Medium and ToM-High groups did not differ from each other, $p = .33$ (Table 3). With respect to Age, the youngest participants (3-y-o) accepted significantly more suggestions than the remaining Age groups, *p*s < .001. The other Age groups did not differ from each other, *p*s > .05 (Table 3).

Yield as a Predictor of Free Recall, Recognition and Suggestion Acceptance

As expected, Yield scores correlated negatively with participants' chronological age, $r = –.78$, $p < .01$, correct recall of activities, $r = –.59$, $p < .01$, correct recall of the story, $r = –.40$, $p < .01$, and distortions of the story, $r = –.22$, $p < .05$, but positively with acceptance of suggestions, $r = .59$, $p < .01$, and also with the hit rate in recognition, $r = .27$, $p < .01$ (Table 1). The results of the regression analyses of Yield and Age as predictor variables for *free recall* (*activities* and *story*) and *recognition* (*hit rate* and *suggestion acceptance*) are described below (Table 2).

Free recall of activities. Yield, along with Age, seemed to predict *correct recall* ($R^2 = .41$), $F(2, 98) = 34.42$, $p < .001$ (Table 2). However, in the ANCOVA, Age was the only variable with a medium effect size, $F(1, 95) = 7.06$, $p = .01$, $\eta^2 = .07$. Therefore, the higher the participants' age, the higher the amount of correct recall (Table 3). Given that the number of *distortions* was very low (Table 3), the regression analysis did not show a statistically significant model, $F(2, 98) < 1$ (Table 2).

Free recall of the story. The regression model reached statistical significance, $F(2, 98) = 15.73$, $p < .001$, but Yield was not a good predictor variable, unlike Age, ($R^2 = .24$; see Table 2). In the ANCOVA, only Age achieved a marginally significant effect, $F(1, 95) = 3.45$, $p = .07$, $\eta^2 = .04$ (Table 3).

Regarding *distortions*, although the regression analysis showed a statistically significant model, $F(2, 98) = 3.33$, $p = .04$ ($R^2 = .06$; see Table 2), neither Yield nor Age were good predictor variables because the number of distortions was very low (Table 3).

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press

Table 2. *Summary of Regression Analyses of ToM, Yield, and Age Rescaled on the Dependent Measures*

Variables	*B*	*SE B*	*b*	*t*	*Sig*
Free Recall of Activities: Correct Recall					
ToM and Age					
ToM	.28	.22	.12	1.31	.20
Age	.07	.01	.56	6.27	< .001
Yield and Age					
Yield	–.71	.03	–.31	–2.69	.01
Age	.05	.01	.39	3.41	< .01
Free Recall of Activities: Distortions					
ToM and Age					
ToM	–.19	.10	–.21	–1.95	.05
Age	.01	.01	.18	1.64	.11
Yield and Age					
Yield	.09	.12	.11	.72	.48
Age	.01	.01	.16	1.10	.27
Free Recall of the Story: Correct Recall					
ToM and Age					
ToM	–.40	1.22	–.32	–.33	0.74
Age	.30	.06	.50	5.14	< .001
Yield and Age					
Yield	–.93	1.54	–.08	–.61	.55
Age	.26	0.08	.43	3.37	< .01
Free Recall of the Story: Distortions					
ToM and Age					
ToM	–.42	.18	–.25	–2.36	< .05
Age	.03	.01	.33	3.08	< .01
Yield and Age					
Yield	–.30	.23	–.19	–1.30	.20
Age	.01	.01	.08	.58	.57
Hits in Recognition					
ToM and Age					
ToM	.03	.02	.14	1.30	.20
Age	–.01	.00	–.25	–2.31	< .05
Yield and Age					
Yield	.06	.03	.29	2.05	.04
Age	.00	.00	.08	.26	.80
Suggestion Acceptance					
ToM and Age					
ToM	–.06	.03	–.19	–2.18	.03
Age	–.01	.00	–.49	–5.56	< .001
Yield and Age	.07	.04	.23	1.96	.05
Yield	–.06	.03	–.19	–2.18	.03
Age	–.01	.00	–.41	–3.41	< .01

Hit rate in recognition. The regression model was significant, $F(2, 103) = 4.00$, $p = .02$, ($R^2 = .07$; see Table 2); but no differences were observed among the groups due to the low variability of the means obtained (Table 3).

Suggestion acceptance. Yield and Age explained part of the variance ($R^2 = .36$), $F(2, 103) = 29.40$, $p < .001$ (Table 2 and Figure 2B). Furthermore, in the ANCOVA, Yield, $F(2, 100) = 9.34$, $p < .001$, $\eta^2 = .16$, Age, $F(1, 100) = 24.45$, $p < .001$, $\eta^2 = .20$, and the Yield x Age interaction, $F(2, 100) = 7.26$, $p < .01$, $\eta^2 = .13$, reached statistical significance. The effect sizes for Yield and Age were large, but the effect size for the interaction was medium. Multiple comparisons showed that participants with a high score in Yield accepted significantly more suggestions than participants with Yield-Medium or Yield-Low, *ps* < .001 (Table 3). In addition, as noted above, 3-y-o children accepted significantly more suggestions than the remaining Age groups, *ps* < .001.

Finally, to examine the interaction obtained, we conducted an additional univariate ANOVA, introducing Age as a factor and Yield as the dependent measure (range of Yield scores in the Scale = 0 – 14). A large effect size of Age was obtained, $F(4,105) = 36.44$, $p < .001$, $\eta^2 = .60$. Multiple comparisons indicated that Yield scores in the 3-y-o group ($M = 10.52$, $SD = 3.87$; $min = 1$, $max = 14$) were significantly higher than in the remaining Age groups, *ps* ≤ .01. Likewise, Yield scores for the 4-y-o participants ($M = 6.94$, $SD = 4.35$; $min = 0$, $max = 14$) were higher than those for the 6-y-o ($M = 1.54$, $SD = 1.24$; $min = 0$, $max = 4$) and 7-y-o participants ($M = 1.33$, $SD = .89$; $min = 0$, $max = 3$), *ps* < .001; while the 5-y-o ($M = 4.26$, $SD = 3.30$; $min = 0$, $max = 14$), 6-y-o, and 7-y-o participants showed no differences from each other, *ps* > .05.

As in previous research, our results show that both, Yield -a measure of *suggestibility* as an individual characteristic- and the *acceptance of suggestion* from an external source about an event experienced, decrease with age, with 3-y-o children being especially vulnerable.

ToM, Yield and Suggestion Acceptance

The two developmental socio-cognitive variables studied (ToM and Yield) appeared to contribute to participants' *acceptance of suggestion* (Figure 2A and 2B), but these two variables correlated negatively with each other, $r = –.55$, $p < .01$ (Table 1). That is, the higher the mentalist skills (i.e., higher ToM scores), the lower the individual suggestibility (i.e., lower Yield score). Thus, we conducted an ANOVA to examine the extent to which participants' mentalist skills influenced their suggestibility score. For that, mentalist skill (ToM-Low, ToM-Medium, ToM-High) was added as a factor and Yield scores (range from 0 to 14) as the dependent measure.

A large effect size of ToM was obtained, $F(2, 105) = 17.98$, $p < .001$, $\eta^2 = .26$. Multiple comparisons indicated that the participants with ToM-Low had a significantly higher Yield score ($M = 9.46$, $SD = 5.03$; $min = 0$, $max = 14$) than the participants with ToM-Medium ($M = 5.41$, $SD = 4.38$; $min = 0$, $max = 14$) and ToM-High (High: $M = 3.03$, $SD = 3.06$; $min = 0$, $max = 12$), *ps* < .001. The Medium and High groups also differed from each other, $p = .04$.

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press

Table 3. *Descriptive Statistics of ToM, Yield, and Chronological Age in Dependent Measures*

		Free Recall of Activities: Correct Recall		Free Recall of Activities: Distortions		Free Recall of the Story: Correct Recall		Free Recall of the Story: Distortions	
	n	*M (SD)*	*Min-Max*	*M (SD)*	*Min-Max*	*M (SD)*	*Min-Max*	*M (SD)*	*Min-Max*
ToM-Low	24	1.20 (1.66)	0–6.33	.46 (.72)	0–2	3.67 (6.75)	0–30	.75 (1.39)	0–5
ToM-Medium	40	2.41 (1.61)	0–6	.30 (.72)	0–3	7.50 (9.98)	0–37	.78 (1.51)	0–7
ToM-High	37	3.01 (1.96)	0–8.67	.22 (.58)	0–3	8.65 (10.35)	0–46	.43 (0.83)	0–4
Yield-Low	40	3.51 (1.78)	1–8.67	.35 (.77)	0–3	10.73 (10.50)	0–46	.98 (1.42)	0–5
Yield-Medium	35	2.18 (1.38)	0–6	.23 (.55)	0–2	7.14 (9.83)	0–37	.60 (1.35)	0–7
Yield-High	26	.76 (1.29)	0–6	.35 (.69)	0–2	1.12 (2.25)	0–9	.19 (.63)	0–3
3 years-old	27	.93 (1.38)	0–6	.30 (.67)	0–2	1.59 (2.94)	0–9	.33 (.73)	0–3
4 years-old	18	1.62 (1.32)	0–4.5	.22 (.55)	0–2	3.72 (6.41)	0–21	.28 (.75)	0–3
5 years-old	19	2.75 (1.18)	1–5.3	.26 (.45)	0–1	7.84 (7.72)	0–24	.74 (1.66)	0–7
6 years-old	25	3.32 (1.89)	1–6.5	.28 (.61)	0–2	10.72 (11.26)	0–37	.96 (1.49)	0–5
7 years-old	12	3.93 (1.92)	1.5–8.5	.58 (1.16)	0–3	15.08 (13.49)	0–46	1.08 (1.51)	0–4

		Hits in Recognition		Suggestion Acceptance	
	n	*M (SD)*	*Min-Max*	*M (SD)*	*Min-Max*
ToM-Low	26	.55 (.22)	0–1	.33 (.30)	0–1
ToM-Medium	42	.57 (.17)	.25–1	.13 (.23)	0–.88
ToM-High	38	.57 (.14)	.25–.88	.06 (.11)	0–.55
Yield-Low	41	.51 (.14)	.13–.75	.03 (.06)	0–.25
Yield-Medium	35	.56 (.15)	.25–.88	.13 (.17)	0–.75
Yield-High	30	.63 (.22)	0–1	.35 (.32)	0–1
3 years-old	31	.60 (.23)	0–1	.38 (.32)	0–1
4 years-old	18	.62 (.17)	.25–.90	.11 (.12)	0–.40
5 years-old	19	.52 (.11)	.25–.75	.08 (.13)	0–.40
6 years-old	26	.52 (.16)	.15–.75	.04 (.07)	0–.25
7 years-old	12	.55 (.11)	.40–.75	0 (0)	0–0

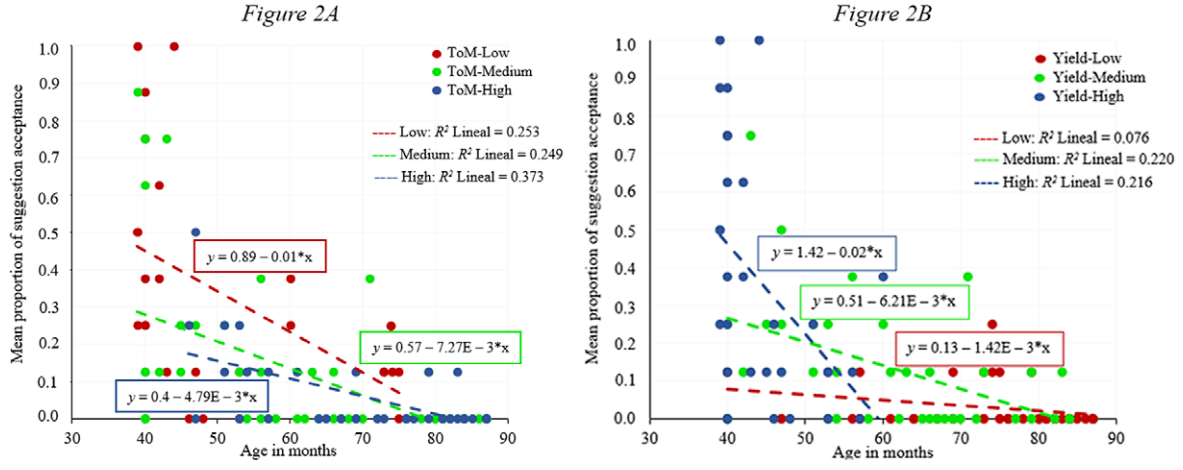

Figure 2. Scatter Plot of Mean Proportions of Suggestion Acceptance as a Function of ToM and Age (2A), and as a Function of Yield and Age (2B)

Discussion

The present study had two main goals. First, to examine, in a sample of young children, the relationships between mentalist skills (ToM) and memory for a directly experienced event, and suggestion acceptance for activities that were not actually performed -it is important to point out that activities were carried out one week

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press

earlier and the suggestion was provided only once. Second, to establish the relationships between young children's suggestibility, as an individual characteristic (Yield), and memory performance and suggestion acceptance. Additionally, although not less important, we were interested in the predictor value of these two socio-cognitive factors -ToM and Yield- together with or beyond the effect of participants' chronological age.

Following Perner's proposal (2000), we expected ToM to be a good predictor when participants had to remember the interaction they directly experienced with 'the Teacher' a week earlier, but our results did not support that prediction. Still, ToM was positively correlated with correct recall of the activities carried out during Session 1 (Table 1). In fact, in our study, the factor that best explained memory performance for the activities was participants' chronological age. Yield was also unable to predict participants' memory for those activities. In that sense, the pattern of results obtained was similar to that found by other authors: That is, the older the children are, the greater correct recall is for the activities that were carried out by them (e.g., Ceci & Howe, 1978; Ornstein et al., 1992; see also the review by Malloy & Quas, 2009).

Furthermore, as expected, chronological age was the most relevant factor for the narrative memory of the story. In this case, there was not a clear autobiographical component involved because participants only had to recount the contents of a story previously heard. Other authors have obtained a similar result as well (Kleinknecht & Beike, 2004). In the same vein, when preschoolers have been asked to recall what happened in video recordings previously watched, memory performance (free recall, in response to open-ended prompts, or to specific questions) has been much better predicted by their chronological age than by their ToM scores (Melinder et al., 2006).

A seemingly surprising result is that, in general, our participants recalled more details of the story (30%) than of the activities performed (17%), which disagrees with previous results (Gobbo et al., 2002; Ornstein et al., 1992; Principe et al., 2006). One reason for this might be that the activities conducted were too similar to those that involve a pedagogic purpose of some sort, and they might not have been salient enough for our participants. 'The Teacher' presented materials and the children had to answer questions about the activities performed with those materials. Children were quite likely to assume that the interaction with 'the Teacher' could have a pedagogic intent, since usual pedagogical signals (e.g., shared attention, eye contact between adult-child) were present during Session 1. Thus, we cannot discard that young participants were interpreting those signals as indicating that knowledge transmission of some sort – by the adult– was about to take place (Csibra & Gergely, 2009), instead of interpretating the situation as an interaction where they and 'the Teacher' were doing activities together. In fact, in a different study conducted in our laboratory (Peláez et al., 2019), the interviewer and the child carried out activities that were more interactive and attractive for participants, and memory performance for activities, tested a week later, was very similar to what has been found in previous research. In that sense, it seems that the story with drawings used in the present study turned out to be very attractive for children[5]. This might explain why recall of the story approximated memory performance typically found for recalling activities in other studies after a delay of one week (Gobbo et al., 2002; Ornstein et al., 1992; Principe et al., 2006).

Furthermore, both ToM and Yield influenced *suggestion acceptance*, as expected. In the case of ToM (Figure 2A), the slopes of the lines showed a purely quantitative pattern: As the children's chronological age increased their mentalist skills also increased, and this, in turn, was accompanied by a progressive decrease in their vulnerability to external suggestion. Thus, below 4.6 years old, both ToM and Age played a role in the acceptance of external suggestion, even when the suggestion was presented once during a single interview. Below this age, when participants had *medium* or *high* mentalist skills, their suggestion acceptance was low (mean about .10), but when they had *low* mentalist skills, the mean of suggestion acceptance trebled and increased above .30.

In the case of Yield (Figure 2B), there was an interaction with Age. Thus, when our participants had *low* individual suggestibility, regardless of their chronological age, the acceptance of suggestion was close to floor effect. However, when they had *high* individual suggestibility, the suggestion acceptance increased, and youngest participants were more vulnerable to external suggestion (mean above .30). When participants had *medium* suggestibility, it was observed that as the child's age increased, the vulnerability to suggestion gradually decreased. Note that for *medium* suggestibility, the acceptance of external suggestion of the youngest participants was below .20, while for the youngest participants with *high* individual suggestibility this value was close to .40. This pattern indicates that when children are individually characterized by being easily suggestible, they can be vulnerable to external suggestion even when it occurs only once. In fact, below 4.6 years old, Yield was relevant regarding acceptance of external suggestions, because it maintained a linear relationship

[5]The story was about the amusing adventures of a boy and two funny and extravagant aliens while they were looking for the boy's lost puppy.

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press

with age for children characterized by *high* and *medium* individual suggestibility.

In summary, there seems to be a turning point at 4.6 years old, in which, depending on the level of development of the socio-cognitive factors analyzed in the present study, children may become more or less vulnerable to suggestions from an external source of information even if they are given only once.

A possible limitation of our study was the reduced number of children in the ToM-Low group, despite the fact that the three-year-old group was the largest in the sample (Figure 1). There could be two explanations for this. First, formal schooling has been shown to be a relevant factor in encouraging ToM development. Hughes et al. (2014) obtained consistent results with a "pedagogical experiences" hypothesis. In several ToM tasks, they found that children from UK outperformed both the Japanese and Italian children, but there were no significant differences between the latter two groups. That advantage in ToM performance was due to schooling in the UK, which occurs a year earlier than in Italy and Japan. In the present study, our youngest participants attended early childhood education at a school, and most of them had also been in a nursery school the previous two years. This means that our participants could have obtained better scores in the ToM Scale than expected according to their age due to their early schooling, which could benefit their ToM development. Second, the first task of the ToM Scale (*diverse desires*) seems to be a quite easy task to resolve successfully. A child could only fail if he or she thinks that everybody shares his or her desires and refuses to accept the possibility that someone might want or like something different from what he or she wants or likes (Westra & Carruthers, 2017). However, that kind of egocentric reasoning is shown at an early age, around 14 months, but not at 18 months (Repacholi & Gopnik, 1997). Thus, the *diverse desire* task could be quite easy for our participants[6]. Besides that, their early schooling might also contribute to a higher probability in overcoming a higher number of ToM tasks.

Moreover, ToM tasks are not always sensitive enough to accurately detect differences in mentalist abilities (Quesque & Rossetti, 2020), and perhaps the criteria established to discriminate among different ToM level through the Wellman and Liu's Scale is not always sufficient. A consequence is that predictive capacity of ToM tasks for performance in different cognitive tasks has been seriously limited (Klemfuss & Olaguez, 2020). Thus, the controversy about ToM persists after decades of research, and it is not easy to resolve it at the moment.

Results of the present study clearly indicated that neither mentalist skills (ToM) nor individual differences in suggestibility (Yield) seemed to play a crucial role in memory performance for our participants. As has been consistently found, chronological age was actually the factor that best explained memory performance (Malloy & Quas, 2009). However, in addition to age, suggestion acceptance was more influenced by the developmental socio-cognitive factors examined in the present study than memory.

Specifically, ToM along with age linearly predicted suggestion: That is, the lower the mentalist skills and age of the participants, the greater the vulnerability to external suggestion. For Yield and age, a clear linear relationship was also shown when participants had *high* suggestibility scores (see the steep slope in the Figure 2B); for participants with *medium* suggestibility a lineal relationship was also observed, although with a decrease in vulnerability to suggestion that was softer with increasing age. It is worth noting that at 4.6 years old, and even when participants have received the suggestion only once during a single interview, a turning point is observed in the acceptance of that suggestion. The suggestion acceptance of children below that age was linearly related to their mentalist skills (ToM), or to individual suggestibility when they had obtained *high* or *medium* scores in Yield.

Thus, although chronological age continues to appear as a key factor, when young children have a poor memory of a past event, and they are exposed to only one source of suggestion at one point in time, some developmental socio-cognitive skills –such as ToM or Yield– could modulate their vulnerability to that suggestion. Future research should examine if other socio-cognitive factors along with situational factors that involve a good memory for the event and a persistent suggestion (e.g., Garven et al., 1998; Otgaar et al., 2019; Peláez et al., 2019; Peláez et al., 2021; Sauerland et al., 2019) could play a relevant factor in young children's suggestion acceptance. To the extent that research findings can establish the specific influence of these factors on young children, practitioners will be able to do a better job when dealing with forensic cases in which children are involved as alleged victims and their statements are the main prosecution evidence in court proceedings.

[6]We had 31 participants younger than 48 months of age. Fifty-five percent of them passed the *diverse desire* task ($n = 17$) and 45 % failed to pass it ($n = 14$). Of the total sample ($N = 106$), only 23% of the participants failed to overcome this task.

References

Azzopardi, C., Eirich, R., Rash, C. L., MacDonald, S., & Madigan, S. (2019). A meta-analysis of the prevalence of child sexual abuse disclosure in forensic settings. *Child Abuse & Neglect*, *93*, 291–304. https://doi.org/10.1016/j.chiabu.2018.11.020

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press

Brainerd, C. J., & Reyna, V. F. (2012). Reliability of children's testimony in the era of developmental reversals. *Developmental Review*, *32*(3), 224–267. https://doi.org/10.1016/j.dr.2012.06.008

Bright-Paul, A., Jarrold, C., & Wright, D. B. (2008). Theory-of-mind developmental influences suggestibility and source monitoring. *Developmental Psychology*, *44*(4), 1055–1068. https://doi.org/10.1037/0012-1649.44.4.1055

Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology*, *50*, 419–439. https://doi.org/10.1146/annurev.psych.50.1.419

Bruck, M., & Melnyk, L. (2004). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology*, *18*(8), 947–996. https://doi.org/10.1002/acp.1070

Caprin, C., Benedan, L., Ciaccia, D., Mazza, E., Messineo, S., & Piuri, E. (2016). True and false memories in middle childhood: The relationship with cognitive functioning. *Psychology, Crime & Law*, *22*, 473–494. https://doi.org/10.1080/1068316X.2016.1168420

Ceci, S. J., & Bruck, M. (2006). Children's suggestibility: Characteristic and mechanisms. *Advances in Child Development and Behavior*, *34*, 247–281. https://doi.org/10.1016/S0065-2407(06)80009-1

Ceci, S. J., & Howe, M. J. A. (1978). Age-related differences in free recall as a function of retrieval flexibility. *Journal of Experimental Child Psychology*, *26*(3), 432–442. https://doi.org/10.1016/0022-0965(78)90123-6

Ceci, S. J., Toglia, M. P., & Ross, D. F. (1988). On remembering… more or less: A trace strength interpretation of developmental differences in suggestibility. *Journal of Experimental Psychology: General*, *117*(2), 250–262. https://doi.org/10.1037/0096-3445.117.2.201

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Diges, M., Moreno, A., & Pérez-Mata, N. (2010). Efectos de sugestión en preescolares: Capacidades mentalistas y diferencias individuales en sugestionabilidad [Suggestion effects in preschoolers: Mentalist skills and individual differences in suggestibility]. *Infancia y Aprendizaje*, *33*(2), 235–254. https://doi.org//10.1174/021037010791114634

Elischberger, H. B. (2005). The effects of prior knowledge on children's memory and suggestibility. *Journal of Experimental Child Psychology*, *92*(3), 247–275. https://doi.org/10.1016/j.jecp.2005.05.002

Garven, S., Wood, J. M., Malpass, R. S., & Shaw, J. S. (1998). More than suggestion. The effect of interviewing techniques from the McMartin Preschool case. *Journal of Applied Psychology*, *83*(3), 347–359. https://doi.org/10.1037/0021-9010.83.3.347

Gobbo, C., Mega, C., & Pipe, M.-E. (2002). Does the nature of the experience influence suggestibility? A study of children's event memory. *Journal of Experimental Child Psychology*, *81*(4), 502–530. https://doi.org/10.1006/jecp.2002.2662

Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions. A handbook*. John Wiley & Sons. http://doi.org/10.1002/9780470713297

Guéguen, N., Lourel, M., Pascual, A., & Mouda, F. (2011). L'importance d'un effet: Méthodologie simple de détermination et d'évaluation de l'«effect size» [A simple method to determine and evaluate "effect size"]. *Cahiers D'études et de Recherches Francophones /Sante*, *21*(2), 103–105. https://doi.org/10.1684/san.2011.0240

Hughes, C., Devine, R. T., Ensor, R., Koyasu, M., Mizokawa, A., & Lecce, S. (2014). Lost in translation? Comparing British, Japanese, and Italian children's theory-of-mind performance. *Child Development Research*, *2014*, 1–10. https://doi.org/10.1155/2014/893492

Karpinski, A. C., & Scullin, M. H. (2009). Suggestibility under pressure: Theory of mind, executive function, and suggestibility in preschoolers. *Journal of Applied Developmental Psychology*, *30*(6), 749–763. https://doi.org/10.1016/j.appdev.2009.05.004

Kleinknecht, E., & Beike, D. R. (2004). How knowing and doing inform an autobiography: Relations among preschoolers' theory of mind, narrative, and event memory skills. *Applied Cognitive Psychology*, *18*, 745–764. https://doi.org/10.1002/acp.1030

Klemfuss, J. Z., & Olaguez, A. P. (2020). Individual differences in children's suggestibility: An updated review. *Journal of Child Sexual Abuse*, *29*, 158–182. https://doi.org/10.1080/10538712.2018.1508108

Malloy, L. C., & Quas, J. A. (2009). Children's suggestibility: Areas of consensus and controversy. In K. Kuehnle & M. Connell (Eds.), *The evaluation of child sexual abuse allegations. A comprehensive guide to assessment and testimony* (pp. 267–298). John Wiley & Sons.

McCarthy, D. (2004). *Escalas McCarthy de aptitudes y psicomotricidad para niños* [The McCarthy Scales of Children's Abilities] *(MSCA)*. Madrid TEA.

Melinder, A., Endestad, T., & Magnussen, S. (2006). Relations between episodic memory, suggestibility, theory of mind, and cognitive inhibition in the preschool child. *Scandinavian Journal of Psychology*, *47*(6), 485–495. https://doi.org/10.1111/j.1467-9450.2006.00542.x

Melinder, A., Scullin, M. H., Gunnerød, V., & Nyborg, E. (2005). Generalizability of a two-factor measure of pre-school age children's suggestibility in Norway and the USA. *Psychology, Crime & Law*, *11*(2), 123–145. https://doi.org/10.1080/10683160512331316299

Ornstein, P. A., Gordon, B. N., & Larus, D. M. (1992). Children's memory for a personal experienced event: Implications for testimony. *Applied Cognitive Psychology*, *6*(1), 49–60. https://doi.org/10.1002/acp.2350060103

Otgaar, H., Chan, J. C. K., Calado, B., & La Rooy, D. (2019). Immediate interviewing increases children's suggestibility in the short term, but not in the long term. *Legal and Criminological Psychology*, 24(1), 24–40. https://doi.org/10.1111/lcrp.12137

Peláez, M., Pérez-Mata, N., & Diges, M. (2019). Influencia del conocimiento previo y la repetición de entrevistas: memoria y sugestión en una muestra de preescolares [Influence of prior knowledge and repetition of interviews: Memory and suggestion in a sample of preschool students]. *Colombia Forense*, *6*(1), 1–23. https://doi.org/10.16925/2145-9649.2019.01.02

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press

Peláez, M., Pérez-Mata, N., & Diges, M. (2021). How does prior knowledge affect children's memory and suggestibility? C. Pracana & W. Wang (Eds.), *Psychological applications & developments VII. Advances in psychology and psychological trends series* (pp. 317–329). InScience Press. https://doi.org/10.36315/2021padVII27

Perner, J. (1991). *Understanding the representational mind*. The MIT Press.

Perner, J. (2000). Memory and theory of mind. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 317–312). Oxford University Press.

Perner, J., Kloo, D., & Gornik, E. (2007). Episodic memory development: Theory of mind is part of re-experiencing experienced events. *Infant and Child Development*, *16*(5), 471–490. https://doi.org/10.1002/icd.517

Perner, J., & Ruffman, T. (1995). Episodic memory and autonoetic consciousness: Developmental evidence and a theory of childhood amnesia. *Journal of Experimental Child Psychology*, *59*(3), 516–548. https://doi.org/10.1006/jecp.1995.1024

Principe, G. F., Greenhoot, A. F., & Ceci, S. J. (2014). Young children's eyewitness testimony memory. In T. Perfect & D. S. Lindsay (Eds.), *The SAGE Handbook of applied memory* (pp. 633–653). SAGE Publications. https://doi.org/10.4135/9781446294703.n35

Principe, G. F., Kanaya, T., Ceci, S. J., & Singh, M. (2006). Believing is seeing: How rumors can engender false memories in preschoolers. *Psychological Science*, *17*(3), 243–248. https://doi.org/10.1111/j.1467-9280.2006.01692.x

Quas, J. A., Wallin, A. R., Papini, S. P., Lench, H., & Scullin, M. H. (2005). Suggestibility, social support, and memory for a novel experience in young children. *Journal of Experimental Child Psychology*, *91*(4), 315–341. https://doi.org/10.1016/j.jecp.2005.03.008

Quesque, F., & Rossetti, Y. (2020). What do theory-of-mind tasks actually measure? Theory and practice. *Perspectives on Psychological Science*, *15*(2) 384–396. https://doi.org/10.1177/1745691619896607

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14-and 18-month-olds. *Developmental Psychology*, *33*(1), 12–21. https://doi.org/10.1037/0012-1649.33.1.12

Sauerland, M., Brackmann, N., & Otgaar, H. (2019). Rapport: Little effect on children's, adolescents', and adults' statement quantity, accuracy, and suggestibility. *Journal of Child Custody*, *15*(4), 268–285. https://doi.org/10.1080/15379418.2018.1509759

Scullin, M. H., & Bonner, K. (2006). Theory of mind, inhibitory control, and preschool-age children's suggestibility in different interviewing contexts. *Journal of Experimental Child Psychology*, *93*(2), 120–138. https://doi.org/10.1016/j.jecp.2005.09.005

Smith, T. D., Raman, S. R., Madigan, S., Waldman, J., & Shouldice, M. (2018). Anogenital findings in 3,569 pediatric examinations for sexual abuse/assault. *Journal of Pediatric and Adolescent Gynecology*, *31*, 79–83. https://doi.org/10.1016/j.jpag.2017.10.006

Thomas, M. S. C., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language, and Hearing Research*, *52*(2), 336–358. https://doi.org/10.1044/1092-4388(2009/07-0144)

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory & Cognition*, *14*(4), 329–338. https://doi.org/10.3758/BF03202511

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne*, *26*(1), 1–12. https://doi.org/10.1037/h0080017

Welch-Ross, M. (2000). A mental-state reasoning model of suggestibility and memory source monitoring. In K. P. Roberts & M. Blades (Eds.), *Children's source monitoring* (pp. 227–256). Lawrence Erlbaum Associates.

Wellman, H. M. (2012). Theory of mind: Better methods, clearer findings, more development. *European Journal of Developmental Psychology*, *9*(3), 313–330. https://doi.org/10.1080/17405629.2012.680297

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Westra, E., & Carruthers, P. (2017). Pragmatic development explains the Theory-of-Mind Scale. *Cognition*, *158*, 165–176. https://doi.org/10.1016/j.cognition.2016.10.021

Wood, J. M., & Garven, S. (2000). How sexual abuse interviews go astray: Implications for prosecutors, police, and child protection services. *Child Maltreatment*, *5*(2), 109–118. https://doi.org/10.1177/1077559500005002003

https://doi.org/10.1017/SJP.2022.22 Published online by Cambridge University Press