The role of declarative and procedural metamemory in event-based prospective memory in school-aged children

Milvia Cottini, Demis Basso & Paola Palladino

ABSTRACT

Prospective memory (PM) develops considerably during primary school years (7-8 years). Developmental changes have been mainly related to executive functions, although it has been recently suggested that PM would potentially benefit also from metamemory (MM). To date, only procedural MM, operationalized as performance predictions, has been investigated in relation to PM, while declarative MM remained unexplored. Adult's performance has shown to improve with predictions, but only in a resource-demanding (i.e., categorical) and not in a more automatic (i.e., specific) PM

task. The aim of the present investigation was to study whether PM performance of 7-year-old children (N = 59) would benefit from performance predictions. Thus, half of the children had to predict their performance and half of them received standard instructions for two PM tasks: one including categorical and one specific PM targets. To investigate the processes underlying the retrieval of PM targets and the effect of predictions, we further obtained measures for declarative MM, inhibitory control and working memory (WM). Results revealed that children benefitted from performance predictions in the categorical but not in the specific PM task. This advantage caused slower ongoing task response times, suggesting that strategic monitoring processes were enhanced. Moreover, PM performances were related to WM capacity and declarative MM. However, declarative MM mainly predicted the PM advantage in the prediction-group, showing that children with high MM knowledge especially benefitted from performance predictions. These evidences are the first showing the important relation between procedural MM, declarative MM and PM in school-aged children.

Keywords: prospective memory; declarative metamemory; procedural metamemory; performance predictions; school-aged children.

Introduction

In everyday life we frequently have to remember to carry out a previously planned action at the appropriate moment (e.g., buying bread when passing by a bakery, taking medicine at 8 am, or asking a colleague something after a meeting). This ability is defined as prospective memory (PM; Einstein & McDaniel, 1990), which is crucial for our autonomy and independence in daily life, as adults, but especially during childhood and adolescence. For example, PM develops considerably during childhood (Kvavilashvili, Kyle, & Messer, 2008), allowing children to become more

and more independent from adult help in daily activities. Particularly when entering school, children are expected to be able to remember (and fulfill) at least some of their self-planned intentions, as well as future tasks assigned from others (Mahy, Moses, & Kliegel, 2014a). Developmental changes in PM during the primary school years have shown to be related to development of executive processes (see Mahy & Munakata, 2015). However, recently it has been suggested that PM would also potentially benefit from metamemory, although there is little evidence thus far confirming this hypothesis (see Kvavilashvili & Ford, 2014). Our study's aim was to fill this gap, and to investigate the role of both procedural and declarative metamemory in children's PM.

Prospective memory in school-aged children and its underlying processes

In the last few years, interest in PM development has increased substantially (see Mahy, Kliegel, & Marcovitch, 2014). Research has shown that PM develops from pre-school age, throughout the school years, until late adolescence (Zimmermann & Meier, 2006), with important developmental advances identified between the age of 7 and 8 years. In particular, from this age, children have been shown to become increasingly accurate in remembering to execute delayed intentions (Kerns, 2000; Smith, Bayen, & Martin, 2010; Yang, Chan, & Shum, 2011). Besides the importance of retrospective memory (RM) processes for PM development, age-related improvements have been linked mainly to development of executive processes, such as inhibitory control, working memory (WM), set shifting and monitoring (e.g., Kliegel et al., 2013; Spiess, Meier, & Roebers, 2016; Yang et al., 2011).

Mahy and colleagues (2014a) proposed an *Executive Framework* to explain PM

development, falling clearly within the developmental research domain, and based on the *preparatory attention and memory* (PAM) theory (Smith, 2003; Smith & Bayen, 2004) and the *Multiprocess* view (McDaniel & Einstein, 2000). Accordingly, developmental advances in executive processes should support PM more effectively, particularly when executive demands of the task are high. Furthermore, the authors claimed that different executive functions would influence PM development at different ages and during different phases of PM (i.e., formation, retention, retrieval, execution and evaluation of an intention): WM may play an important role in early childhood, whereas inhibitory control, monitoring and shifting may be crucial later in school-age. Moreover, inhibitory control and set shifting are predicted to influence OT performance and cue detection, whereas WM and planning would have a greater effect during intention formation and retention. Besides executive processes, the authors also suggested that PM development would benfit from development of metamemory abilities, which also improve over childhood (especially during the primary school years) and play an important role in RM (see Schneider & Lockl, 2008). However, the study by Kvavilashvili and Ford (2014) remains the only confirmation of this hypothesis in children.

Metamemory and its relation to prospective memory

Metamemory (MM) is defined as the verbalizable knowledge and awareness of various memory or memory-related phenomena (Kreutzer, Leonard, Flavell, & Hagen, 1975). It can be distinguished as either declarative or procedural MM (Flavell & Wellman, 1977). Declarative MM includes all explicit and conscious knowledge and beliefs about memory, whereas procedural MM is related to application of this knowledge, i.e., using strategies in addition to controlling, regulating and monitoring

personal memory performances (Flavell & Wellman, 1977). The latter is usually assessed in conjunction with a memory task, before or after which subjects are asked to predict or judge their performance (see Schneider, 2015).

Studies on MM development have shown that this increases especially during the primary school years (e.g. Fritz, Howie, & Kleitman, 2010; Schneider, 1986). Whereas age-related improvements in declarative MM are promoted by language development and reasoning ability (Schneider, Kérkel, & Weinert, 1978), advances in procedural MM have shown to involve monitoring and cognitive control processes (e.g., Isingrini, Perrotin, & Souchay, 2008). MM has shown to play an important role in the development of RM (e.g., Cornoldi et al., 1991; DeMarie & Ferron, 2003; Geurten, Catale, & Meulemans, 2015). Specifically, around the age of 7-8 years, the relation between declarative and procedural MM, that is, MM knowledge and strategy-use, becomes stronger and more effective, being increasingly related to children's memory performance.

While there are several studies investigating the relationship between MM and RM, few have considered the role of MM in PM (e.g., Meeks, Hicks, & Marsh, 2007; Schnitzspahn, Zeintl, Jäger, & Kliegel, 2011). The majority of these have included adults, and only one has investigated the relation between procedural MM and PM performance in preschool children (Kvavilashvili & Ford, 2014). The authors asked 5-year-old children to predict their performances in both a PM and RM task. In line with MM-RM studies (e.g., Fritz et al., 2010), results showed that performance predictions on a memory-recall task were generally overestimated, and that children were overconfident about their performance. In contrast, when predicting performance on an event-based PM task, children's forecasts were relatively accurate compared to subsequent achievement. The authors argued that similar processes might underlie performance predictions and PM abilities, such as episodic future thinking. In fact,

some studies have suggested that projecting oneself into the future, while encoding a prospective intention, can enhance PM performance probably by increasing cue saliency (e.g., Brewer & Marsh, 2010; Kretschmer-Trendowicz et al., 2016). On the other hand, adults' outcomes have been rather inconsistent, showing predictions that were not always confirmed by actual PM performance, given that adults often underestimated their PM performance (e.g., Meeks et al., 2007; Schnitzspahn et al., 2011). Interestingly, some of these studies have shown that performance predictions improved PM performance (Meier et al., 2011; Rummel et al., 2013). In these, PM performance was higher in a group of participants who had to make predictions about their PM performance, compared to a control group. Moreover, predictions only improved performance on a more resource-demanding PM task (i.e., categorical PM task), but not on a more automatic PM task (i.e., specific PM task). This "prediction" advantage was accompanied by a cost, expressed as slower response times (RTs) on the ongoing task, suggesting that performance predictions enhanced the engagement in strategic monitoring (Rummel et al., 2013).

The present study

Given the importance of children's PM abilities when entering school, and the lack of research concerning its various underlying mechanisms (such as MM and executive processes), in the present study we decided to focus on 7-year-old children. This age is supposed to be critical for the development of PM (e.g., Kerns, 2000; Smith et al., 2010; Yang et al., 2011), MM (Schneider, 2015) and executive processes (Anderson, 2002; Lee, Bull, & Ho, 2013) as well as for their relationship (Spiess et al., 2015). Moreover, as demonstrated with adults, we were interested in investigating whether performance predictions (i.e., procedural MM) would influence

children's performance in a resource-demanding PM task.

In the present study, we manipulated PM task difficulty (within participants) and the presence/absence of PM performance predictions (between participants) in an event-based PM task. Thus, half of the children were asked to predict their performance, and the other half received standard instructions before performing two different PM tasks: one categorical (i.e., more resource demanding) and one specific (i.e., more automatic) PM task. First of all, following the studies by Hicks and colleagues (2005), children's performance was expected to be lower in the categorical rather than in the specific PM task. In order to understand the mechanisms involved in the retrieval of categorical versus specific PM targets more effectively, we further measured individual differences in declarative MM, inhibitory control and WM. In both tasks, individual differences in WM and declarative MM were expected to predict PM performance. Alternatively, inhibitory control is likely to predict PM performance in the more resource-demanding task as well as OT performance (see Mahy et al., 2014a).

According to previous studies with adults (Meier et al., 2011; Rummel et al., 2013), performance predictions are expected to improve children's PM performance by enhancing strategic monitoring processes. This was predicted to have differential impact on the tasks, favoring the categorical, rather than the specific PM task. Furthermore, we were interested in investigating the accuracy of children's PM predictions, and how they might be related to performance in the two PM tasks. Following outcomes of the single previous study with preschoolers, children's performance predictions should be generally accurate with respect to actual PM performance in both tasks (Kvavilashvili & Ford, 2014). In accordance, we would expect that children, who predicted they will remember, would actually remember to perform the PM task, while those who predicted they will forget, would actually forget to undertake the PM task. If, in addition to accurate predictions, we also find that

predictions benefit PM performance, it would mean that children were able to translate their MM judgments into adequate self-regulation strategies needed to successfully perform the task. This would be evidence that the relation between procedural MM and PM is well developed and that it affects PM performance itself (see Schneider, 2015). Moreover, this pattern of results is more likely to occur on the categorical rather than the specific PM task, given that these require different levels of strategic monitoring.

The literature on adult participants, however, indicates that adults are somewhat inaccurate in making predictions and mainly underestimate their PM performance (e.g., Meeks et al., 2007; Schnitzspahn et al., 2011). Accordingly, children who underestimate their future PM performance would probably adopt similar resource-allocation strategies, permitting them to perform the PM task correctly. In other words, if a child is not completely confident of his/her performance, but is able to perform it successfully, it is more likely that s/he would have monitored for the PM targets strategically. If this was the case, while the advantage of making explicit predictions would remain, it would also suggest that children's strategic monitoring abilities were well developed, but equally, that children were not yet aware of their actual skills. This would be true for the categorical, but not the specific PM task, given that strategic monitoring is needed less to perform the latter successfully. Consequently, this would still imply that the two processes are related, but that their relation is not yet fully developed. Conversely, if these predictions were incorrect and there was no beneficial effect on PM performance, then we could not conclude anything regarding the relationship between procedural MM and PM, and the paradigm would prove insufficiently sensitive for our goals.

Methods

Participants

Fifty-nine children (33 girls and 26 boys) participated in the study; all regularly attended the second grade in the same public school in a city of Northern Italy. Their ages were between 7 years, 0 months and 7 years, 10 months (mean age: 7 years, 5 months), and all were either native Italian speakers or sufficiently fluent in Italian. Parents and children gave written and oral consent (respectively) for participation.

Materials and procedures

Prospective memory paradigm

    Picture classification task. Event-based PM was measured using a semi-ecological computerized task, based on the classic experimental paradigm proposed by Einstein & McDaniel (1990). The OT consisted of a picture classification task, and was created using the SuperLab software. Pictures were real-object photographs taken from the Bank of Standardized Stimuli (BOSS; Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010), with the database consisting of 480 pictures in total. Because the database was standardized with adults, we conducted a preliminary study to adapt the set for 7-year-old children on the bases of familiarity, pleasantness and category agreement. From the original set, we selected 152 pictures on the basis of high familiarity ($M = 4.24$; $SD = 0.34$; min = 0; max = 5) and high category agreement ($M = 83\%$; $SD = 14$), subsequently presenting them to a group of 41 7-year-old children attending a public primary school in a city of Northern Italy. Participants were asked to name pictures one by one, evaluating each one for its pleasantness (1 = I don't like it; 2 = neutral; 3 = I like it), and categorizing them on the basis of five categories represented by the rooms in which they are usually found: e.g., kitchen (for food and kitchen utensils), bathroom, kids room (for toys), study

room (for school materials) and wardrobe (for clothes). Finally, we chose the most well-known ($M = 98\%$; $SD = 6$), most preferred ($M = 2.49$; $SD = 0.25$), and most easily classifiable ($M = 95\%$; $SD = 6$) pictures to be used in the PM task here. In addition we calculated mean response times (RTs) for picture classification ($M = 5,670$ ms; $SD = 1,391$ ms), in order to fix presentation length for each stimulus.

The resulting two versions of the picture classification task consisted of 75 pictures each. One version included three specific PM cues (i.e., a sandwich, a candy and a ball), whereas the other comprised three PM cues belonging to a specific category (i.e., fruit). In each picture classification task, participants were required to classify each object on the basis of five categories (kitchen, bathroom, kids room, study room and wardrobe), which were organized in blocks. Each block was preceded by the name and the image of the category (category pictures were downloaded from the Internet http://www.midisegni.it/disegni/casa.shtml) following the stimuli in sequence (see Figure 1 for a schematic example). Stimuli were presented successively on a white background. They remained visible for 5000 ms, being preceded by a fixation cross (500 ms), and followed by a blank screen (250 ms). Each block consisted of 15 trials, of which seven were target trials. The three PM cues were embedded in each OT and always presented on the same position across participants (i.e., every 23 trials).

-Figure 1 about here-

The task was presented as a game, and instructions were given telling a story, supported by images appearing on the computer screen. Our aim was to make the task as pleasant and ecologically valid as possible, using real-object photographs and reproducing an everyday situation, such as tidying the house and packing the

backpack for a school trip. Thus, we told a story about a boy named Karl and his dog Bubu. Every time Bubu enters Karl's house, he turns it into chaos and Karl always has to tidy it up. Karl asks the participant to help him to put everything in the right place as fast as possible, before his mother comes back home. To do this, they have to respond by pressing the "Yes"-key (S-key) whenever an object was part of the current category, and the "No"-key (L-key) when it was not. After a practice trial, the story continued with Karl recounting that he had also to finish packing his backpack for the school trip, and asking for help to find the missing objects. These objects (PM cues) did not have to be classified by pressing the Yes- or No-key, but instead had to be put in the backpack by pressing the spacebar. Participants were asked to repeat the instructions for the ongoing and the PM tasks in their own words, and after ensuring they understood the procedure, instructions for the filler-task were presented. Children were told that Karl had to finish his homework before tidying up the rooms and to be faster, he asks the participant to help him. The filler-task consisted in the spatial reasoning sub-scale of the Primary Mental Abilities (Thurstone & Thurstone, 1981; see related section below) and lasted approximately 5 minutes. In this instance, participants began the task without repeating the instructions. The participants performed the two versions of the task with an interval of one month between the first (categorical PM task) and the second version (specific PM task). Participants who failed to recognize PM cues were asked the following questions in order to probe whether failure was due to misunderstanding or inability to remember PM instructions (Kvavilashvili & Ford, 2014): First, "Was there something else to do during the task besides tidying up the rooms?" Second, "Was there something else to do whenever specific pictures appeared? Third, "Was there something to do when the picture of a fruit appeared?" Fourth, "Didn't you have to press the spacebar whenever the picture of a fruit appeared?" Participants who were

unable to answer the latter question were excluded from analysis, since their failure was likely due to forgetting the PM instruction, rather than being a pure PM failure (Kavilashvili et al., 2008).

Procedural Metamemory

Performance prediction paradigm. Procedural MM was evaluated using the same performance prediction paradigm as in Kvavilashvili & Ford (2014). Since this is the only study using the paradigm for PM in children, we followed their method in order to enable comparisons of results. To evaluate the influence of predicting personal performance, we divided our sample into two groups (Meier et al., 2011; Rummel et al., 2013). After giving instructions for the PM task, and after being sure these were understood, half of our participants were asked whether they thought they would remember the PM task. The following question appeared on the screen and was read to the child by the experimenter: "Do you think you will remember to press the spacebar whenever a fruit appears on the screen in order to put it in Karl's backpack?" After giving their responses, a confidence rating scale appeared on the computer screen. Participants were asked: "How sure are you that you will remember/forget?" Afterwards they had to rate their predictions on the scale by pressing either the Keys 1 (= not sure), 2 (= sure) or 3 (= very sure). The group of children who had to predict performances was the same in the categorical and the specific PM task condition.

Declarative Metamemory

Story task. Declarative MM was measured using a modified version of "The Captive Princess" (Cornoldi et al., 1991). A task in a narrative format was chosen in preference to a questionnaire, in order to offer children a better opportunity to

understand concepts of memory and reflection on mental states (Dyer, Shatz, & Wellaman, 2000). The original task was slightly modified in order to add some MM measures related to prospective remembering. The task is a suitable measure of MM knowledge which has been widely used with preschool- and school-aged children (e.g. Cornoldi et al., 1999; Lecce, Demicheli, Zocchi, & Palladino, 2014). The story is about a prince, who wants to save a princess captured in a castle because of witchcraft. The prince meets a farmer near the castle, who tells him that, to undo the spell, he has to ask a wise man who lives far away in a cave on the top of a mountain. The farmer also begs the prince to ask the sage to give him medicine for his sick son. The prince promises the farmer to bring him the medicine and departs for his long journey. The first section ends with the prince being in the wise man's cave, who reveals to him the antidote for the spell (a sequence of actions).

This section comprised three questions: one assessing *PM* ("Is there something else the prince has to remember to do?") and two others, assessing children's *knowledge of forgetting* ("Do you think the prince remembered to ask for the medicine?" and "The prince didn't remember to ask for it, so why do you think he didn't remember?"). Afterwards, children were told that the prince rode the whole way back to the castle. The second section ends with the prince standing in front of the castle's gate, followed by four questions: Again two evaluating children's *knowledge of forgetting*, ("Will the prince remember what to do to save the princess?" and "Unfortunately the prince didn't remember, so why does he not remember?") and one assessing *knowledge of retrieval* ("What can the prince do in order to recall the antidote?"). At this point of the story, the final section starts and children are told that the prince decides to ride back to the wise man to ask him again what to do to break the witch's spell. Before departing, the prince remembers that he has forgotten to ask the sage for the medicine. Consequently, this third section comprises questions

about children's *knowledge of storage*, that is, knowledge about memory maintenance strategies (i.e. "What can the prince do to be sure to remember to ask for the medicine this time?"). Afterwards, the children are told that the prince remembered to ask for the medicine and after the sage repeats the antidote for the spell to him, they are asked: "What can the prince do to be sure to recall the antidote once he arrives in front of the castle?"

Responses to these questions were coded according to the parameters used by the authors: questions examining *knowledge of forgetting* were evaluated on a scale of 0 (e.g. "I don't know") to 7 (e.g. "The prince may has forgotten because too much time has passed, during which he had to remember too many things"). The aim of this scale was to evaluate children's knowledge about the decay of information from memory, and the sensitivity of information to time delay between coding and retrieval as well as how this time is spent to apply memory strategies. Questions concerning *knowledge of retrieval* were evaluated on a scale from 0 (e.g. "I don't know") to 5 (e.g. "He has to recover information he used when he learned from the sage what to do"). This measure aimed to examine children's knowledge about rehearsal processes and mental activities able to contrast information's decay from memory. Questions referred to children's *knowledge of storage* were coded on a scale ranging from 1 (e.g. magic retrieval; or simply paying attention to) to 3 (e.g. rehearse information in the head).

Internal consistency of the revised story task was calculated, using Cronbach's alpha (α). The total measure of declarative MM showed good internal consistency, resulting in an overall α of .61.

Verbal and non-verbal abilities

Verbal meaning sub-scale, Primary Mental Abilities (PMA; Thurstone &

Thurstone, 1981). Vocabulary knowledge was assessed using the PMA verbal meaning sub-scale. This is a written task, which requires choosing one picture out of four matching the instructions given (e.g. "Mark the apple"). The test consisted of 30 items; one point was given for every correct answer.

Spatial reasoning sub-scale, Primary Mental Abilities (PMA; Thurstone & Thurstone, 1981). Non-verbal abilities were assessed using the PMA spatial reasoning sub-scale. It is a written test with 27 geometrical figures. Each figure has to be completed with the missing piece, choosing one from four possible options. One point was given for each correct answer.

Verbal working memory

Digit Span Forward. In order to measure verbal short-term memory, that is passive storage, the forward digit span task was used. Children were required to recall verbally presented digits in the same presentation order. Digits were presented at a rate of one per second, starting from the shortest series (three items) increasing the number of items if the sequence was correctly reported (at least two of the same length). No time limit was given for digit recall; scoring represented the number of correctly repeated digits.

Digit Span Backward. Verbal WM, that is active storage, was measured by means of the backward digit span. This time, children were required to recall verbally presented digits in a reversed order of presentation. As for the previous task, digits were presented at a rate of one item per second beginning with the shortest series to the longest (see forward digit span).

Inhibitory control

Go-No/Go task. In order to measure children's ability to inhibit prepotent

responses, we used a computerized task based on the Go-No/Go paradigm (Brocki &

Bohlin, 2004; Iaboni, Douglas, & Baker, 1995; Shue & Douglas, 1992). The task

consists in the presentation of a series of Go and No-Go stimuli, presented randomly

in the center of a computer screen. The task requires a response to Go stimuli, and

its inhibition in response to No-Go stimuli. The task used in this study comprises a Go

and a No-Go stimulus, represented by a yellow and a blue sphere respectively.

These stimuli appeared sequentially in the center of the screen. Each stimulus was

preceded by a fixation cross (250 ms) and lasted 500 ms, with a random inter-

stimulus interval ranging from 2,550 ms to 2,783 ms (Brocki & Bohlin, 2004).

Participants had to press the spacebar as fast as possible, only when the Go

stimulus appeared, and they needed to withhold response when a No-Go stimulus

appeared. The task consisted of a total of 50 trials, and in order to develop a habitual

response, the majority of the trials (75%) were Go-targets. Performance was

evaluated via the number of commission errors (i.e., giving a response to a No-Go

stimulus) and omission errors (i.e., failing to respond to a Go stimulus). Commission

errors are considered a direct measure of inhibitory control, while omission errors

represent inattention to the task (see Brocki & Bohlin, 2004).

Data analysis

All statistical analyses were performed by means of the free statistical software

R (R Core Team, 2016). First, descriptive statistics relating to children's mean age,

gender and mean scores of the various measures included here have been

presented separately for the two groups (i.e., with and without performance

predictions). In addition, children's performance prediction was compared to their

actual PM performance via ordinal logistic regression. Finally, the effects of

performance predictions and the role of WM, declarative MM and inhibitory control on PM and OT performances were analyzed as follows: A series of four mixed effects regression models (Pinheiro & Bates, 2000) were run while considering accuracy and RTs of the ongoing and PM tasks as dependent variables. The mixed models fitted on data had the following structure: one dependent variable, several variables included as either fixed or random effects.

Compared to traditional regressions, mixed effect regressions allow consideration of the whole structure of data in terms of fixed and random effects, thus ensuring better statistical power. Through this analysis, we could include the same variables of interest as in a common ANOVA, but also several other predictors, in the same model of estimation. In the present analysis, we employed mixed models to include group and task as fixed factors and other measures (such as Declarative MM, WM span forward and backward, inhibition) that, otherwise, could have been evaluated in a separate regression model. Moreover, standard regression includes only fixed effects, therefore it cannot remove part of the variance due to random variables. Mixed models were fit by means of the lme4 package (Bates, Maechler, Bolker, & Walker, 2014).

The two models fitted for accuracy included accuracy for ongoing and PM tasks (one for each model) as dependent variables. Since accuracy for the OT was codified as a dichotomous variable, a generalized mixed model with logit transformation was fit on the data (Jaeger, 2008). The eleven fixed effects considered were: Group as a two-level factor (Group0= without predictions, Group1= with predictions), Task as a two-level factor (Task1= categorical, Task2= specific), the Group × Task interaction; then Declarative MM, WM span forward, WM span backward, the Inhibition errors in the go/no-go task. The interaction between Group and these four variables were considered as continuous fixed effects predictors. The random effects considered in

the model were the random effect of Stimuli (that is, evaluating the contribution of the various stimuli presented) and Participants. These variables were considered as random effects, thus some sources of data variance were taken into account in the model, and this led to improvement of the statistical power of the analysis. Starting with an initial model including the three effects (Group, Task, Group × Task interaction) and random effects (Stimuli and Participants), and after ensuring that its convergence was obtained, a forward-fitting procedure was employed (as typically done with the "glmer" function). Fixed effects predictors were included in the model one at a time: the predictor was maintained if the model converged, otherwise it was discarded. The final model emerged after all predictors were tested for significant effect.

The two models fitted for RTs included log-transformed RTs (to reduce data skewness) as dependent variables (both for ongoing and PM tasks, by using the correct responses only). In addition to the eleven fixed effects considered for accuracy (Group, Task, and the Group × Task interaction, Declarative MM, WM span forward, WM span backward, Inhibition errors in the go/no-go task, together with the interaction between Group and the last four variables), Trial Number (that is, the ordinal position of each trial within the whole experiment, regardless of task type) and the Preceding Trial (that is, the RT to the stimulus presented before the current one) were included as covariates. Given that here, the specific PM task was always administered after the categorical PM task, comparison between the two tasks may be affected by practice effects or age-related intellectual maturation. One possible way to disentangle this potential effect is including trial number as a covariate in the analyses, as reported and explained in Cona and colleagues (2012). Finally, a correlation between the observations was taken into account by specifying an additional variable, included as an additional predictor; the RTs to the (log-

transformed) preceding trial (Baayen & Milin, 2010). The random effects considered were the same used for the two models for accuracy. Each initial model started by including all the variables, and the automatic backfitting function "step" (lmerTest package version 2.0-33: Kuznetsova, Brockhoff, & Christensen, 2015) was employed: Non-significant variables were eliminated from the model one at a time, starting with the variable with the lowest |t|, and resulting in a model containing significant effects only.

The four final models yielded by this procedure are shown in Tables 3 and 4. Fixed effect parameters are interpreted as the effects of traditional regressions. The influence of every fixed effect is calculated, partialling out the influence of the other significant fixed effects. Following standard procedure in regressions, a main effect was kept in the analysis, regardless of its significance, if it was part of significant interaction.

Results

Descriptive statistics

Descriptive statistics concerning children's mean age, gender and mean scores of the various cognitive measures are presented in Table 1. The comparison between the non-prediction and prediction groups was performed through a Multivariate Analysis of variance: approx. $F(11, 44) = 0.413$, $p.=.942$; Pillai's trace $= 0.094$, η^2_p $= .02$. Children's data were equivalent, and did not show any significant differences in respect of any of the variables (age, $p = .410$, $\eta^2_p = .02$; verbal abilities $p = .613$, η^2_p $<. 01$; non-verbal abilities, $p = .301$, $\eta^2_p = .02$; inhibitory control, $p = .639$, $\eta^2_p < .01$; digit span forward, $p = .363$, $\eta^2_p = .02$; digit span backward, $p = .934$, $\eta^2_p < .01$;

declarative MM relative to knowledge of forgetting, $p = .753$, $\eta^2_p < .01$; knowledge of

storage, $p = .536$, $\eta^2_p = .01$; knowledge of retrieval, $p = .851$, $\eta^2_p < .01$; and the total

MM score, $p = .994$, $\eta^2_p < .01$.

-Table 1 about here-

Performance predictions and prospective memory performances

Out of the 59 children, two were excluded from the analysis because they failed

to remember instructions for the categorical PM task, and one was excluded for the

same reason for the specific PM task. This indicated that their PM errors were not

due to PM difficulties, but rather to RM or comprehension difficulties (Kvavilashvili et

al., 2008). Of the final sample, 27 children predicted their performances and 29 did

not (see Table 2).

Within the PM performance prediction group, in the categorical task, all the

children predicted remembering the PM task, twelve of which were "sure" and 15

were "very sure" to remember. Of these 27 children, 21 (78 %) remembered the PM

task effectively, indicating that children are able to predict their actual performance

with reasonable accuracy. An ordinal logistic regression was performed on the group

with prediction only, with the two variables Confidence (as predictor, values: "sure"

and "very sure") and Remembered (as predicted, range: 0 - 3). Results indicated a

value of -0.500 ($SE = .709$), with t= -.705 (*ns*). This shows that, for the categorical

task, although children are able to predict the direction of their performance, their

precision is not indicative.

Within the specific PM performance task, 26 children in the prediction group

predicted remembering, with 1 predicting s/he would forget. In this instance, 21 out of

27 made a congruent prediction (82 %), remembering (or not) the PM task effectively. The ordinal logistic regression showed a value of -2.638 (SE = 1.154), with t= -2.286 ($p < .05$): This indicates that for each unit increase in Confidence (e.g., passing from "sure" to "very sure"), we would expect around 2.5 units decrease in the expected value of items remembered in the log odds scale. That is, the most optimistic children performed worse than those children who were more conservative in predicting their performance.

-Table 2 about here-

Effects of performance predictions on the prospective memory and ongoing task performances

Ongoing task performance. For the OT (Table 3), accuracy did not differ across groups or tasks (see Figure 2a). The only significant predictor resulted from the covariate Inhibition Error: for high values of this variable, OT accuracy was low. All other predictors were introduced one by one, but the model failed to converge for all of them.

Response times for the OT were slower for the group with predictions, compared to the group without predictions, but in the categorical task only (see Figure 2b). Moreover, a significant effect of the preceding trial was found (and removed from the main effect across both tasks), given the high correlation of the latency of current trial with that of the preceding trial (Baayen & Milin, 2015). The significance of random effect for Stimuli and Participants indicates significant variability in the overall performance of each participant, and between the various stimuli. Moreover, it indicates that, taking these sources of variance into account, the

goodness of fit of the model improves (the effect size of the model is $r^2 = .27$).

Importantly, since the four variables of Inhibition error, Declarative MM, WM Span

Forward and Backward were excluded during the backfitting procedure, they should

not be considered as relevant predictors in the analysis presented.

-Table 3 about here-

-Figure 2 about here-

Prospective memory performance. Considering PM (see Table 4), analyses on

accuracies revealed that the interaction between Task and Group was significant.

When PM cues were categorical, children who predicted their performance were

more accurate than children who did not, whereas there were no differences between

the groups when PM cues were specific (see Figure 3a). Moreover, a significant

contribution of Declarative MM and WM span (forward) was found. Both measures

were direct predictors of the accuracy for PM trials, so that the higher these values

were, the higher the accuracy for PM trials.

Response times for the PM tasks were slower for the group without predictions,

compared to the group with predictions (see Figure 3b); both groups also differed

with respect to Declarative MM. A non-significant main effect of Declarative MM and

the significant interaction of Group= 1 x Declarative MM indicates that metamemory

knowledge was a significant predictor for the performance prediction group only. In

this group, higher Declarative MM scores were associated with faster RTs. The

predictor variables of Inhibition error, WM Span Forward and Backward, and their

interaction with Group were excluded during the backfitting procedure. This suggests

that these predictors were not able to account for variability in RTs. The effect-size

for this model was $r^2 = .52$.

-Table 4 about here-

-Figure 3 about here-

Effects of confidence judgment of predictions on prospective memory performances

In order to evaluate the effects of differential prediction precision, a further analysis added the factor "Prediction Confidence" (being 1= "very sure", 2= "sure") to the models previously used, by including participants with predictions only. Analyses were conducted separately for each task, given that a prediction could be reliably connected to the current task, but not for both. Results indicated that, in the categorical condition, models for RTs converged for the OT only (Intercept, Pred1= "very sure": β= 7.334 (0.173), t= 42.49, *p.* < .001; Pred2= "Sure": β= 0.121 (0.035), t= 3.43, *p.* < 0.01; Log(Preceding Trial): β= 0.046 (0.022), t= 2.07, *p.* < .05), suggesting that participants who predicted "sure" showed slower RTs than those who predicted "very sure".

In the specific condition, the models converged for PM accuracy (Intercept, Pred1= "very sure": β= −13.353 (*SE* = 5.124), z= −2.61, *p* < .01; Pred2= "Sure": β= 6.865 (2.556), z= 2.69, *p* < .01), showing that participants who predicted "sure" showed higher accuracy rates than those who predicted "very sure". This also occurred for RTs (Intercept, Pred1= "very sure": β= 7.739 (0.290), t= 26.64, *p* < .001; Pred2= "Sure": β= −1.285 (0.564), t= −2.28, *p* < .05; Declarative MM: β= −0.034 (0.014), t= −2.34, *p* < .05; Inhibition errors: β= 0.143 (0.037), t= 3.90, *p* < .001;

Pred2= "Sure" × Declarative MM: β=0.046 (0.022) , t= 2.09, p < .05). This shows that RTs for participants who predicted "sure", were faster than for those "very sure" of their prediction, and the interaction indicated the latter group ("very sure") was faster with higher Declarative MM, while the former had no impact of Declarative MM. In addition, a general effect of Inhibition errors was also reported, indicating that higher number of errors led to slower RTs in both groups.

Discussion

The aim of the present investigation was to study the relationship between MM and PM in 7-year-old children. First, we were interested in examining whether children's PM performance would benefit from procedural MM, operationalized as performance predictions. Thus, half of our participants were asked to predict their performances and half received standard instructions for two different PM tasks; one being more resource demanding (i.e., categorical PM task) and one being less resource demanding (i.e., specific PM task). Moreover, we were interested in evaluating children's accuracy in predicting their PM performance in the two different tasks. Finally, to investigate the processes underlying the retrieval of different PM tasks and the effect of predictions, we also evaluated inhibitory control, WM and declarative MM.

Consistent with previous studies (e.g., Hicks et al., 2005), our results have shown that accuracy was generally lower in the categorical than in the specific PM task, indicating that responding to a categorical PM task is more resource demanding than to a specific PM task. In both tasks, inhibitory control abilities predicted accuracy in the OT, supporting the Executive Framework of PM development (Mahy et al., 2014a). However, contrary to what we expected, inhibitory control was not implicated in PM

target detection. That said, accuracy in both PM tasks (but not RTs) was related to WM capacity (i.e., consistent with Smith & Bayen, 2005; Yang et al., 2011). Thus, as predicted by the Executive Framework (Mahy et al., 2014a), children with greater WM span had also higher PM accuracy. An interesting and new finding was that declarative MM was an important predictor of PM performance too, showing that children with higher MM knowledge were better at prospective remembering. So far, the link between declarative MM, strategy use and memory performance has been investigated only in relation to RM. These studies have shown that relation between these processes becomes stronger around the age of 7 (Schneider, 2015) and is linked to advances in executive functions (Roebers & Feurer, 2016).

With respect to PM, to date, only procedural MM has been considered, using the performance prediction paradigm (e.g., Meeks et al., 2007; Schnitzspahn et al., 2011). Recent studies have also investigated the direct effect of predictions on adults' PM performance, by comparing groups with and without performance predictions (Meier et al., 2011; Rummel et al., 2013). Our results replicated these findings, showing that even 7-year-old children can benefit from performance predictions. Moreover, as in the adult studies, making predictions improved children's PM performance in the categorical, but not the specific PM task, when compared to the non-prediction group. This PM advantage was accompanied by slower OT RTs, indicating that these children monitored strategically for detection of PM targets (Smith 2003; Smith & Bayen, 2004). Moreover, the prediction group was faster than the non-prediction group in detecting the PM targets, further indicating that they monitored strategically in detecting categorical PM targets. Interestingly, faster RTs to PM targets in the prediction-group were mediated by declarative MM. Specifically, those children in the prediction group with high declarative MM were also those who monitored strategically the most for PM targets.

In order to investigate how the effect of performance predictions might be related to performance in both PM tasks, we further analyzed children's PM prediction accuracy (i.e., procedural metamemory). Similar to Kvavilashvili and Ford's study with preschoolers (2014), the percentage of children who predicted remembering and actually remembered the PM task was relatively high (>70%), showing that they were able to predict the direction of their future performance. However, when considering confidence judgments of predictions, precision was not always optimal. In the categorical PM task, children's confidence judgements were not directly related to actual PM performance, whereas in the specific PM task they were. Children who showed some caution (i.e., were only "sure" of their prediction), were more accurate and faster in detecting specific PM cues than children who were "very sure" of their prediction. Again, this indicates that they may have monitored strategically for detection of PM cues. However, given that successfully performing a specific PM task relies more on spontaneous processes, it follows that engaging in strategic monitoring is not functional (Einstein et al., 2005). That said, in the categorical PM task confidence judgments were related to OT RTs. Similarly to the specific condition, less overconfident children also had slower RTs, suggesting that they have monitored strategically for detection of categorical targets. These children may have judged the task as being more difficult, which resulted in a change of attention allocation policies, as argued by Hicks and colleagues (2005; see also Rummel & Meiser, 2013). Although reasonable, this interpretation needs to be corroborated further.

Studies using the performance prediction paradigm in relation to RM have frequently revealed inconsistent results (see Schneider, 2015), and adults' predictions have also not always been accurate, compared to their PM performance (e.g., Meeks et al., 2007; Rummel et al., 2013; Schnitzspahn et al., 2011). However, in Kvavilashvili and Ford's study (2014) 5-year-old children seemed to be highly accurate in their PM

predictions. Some researchers have argued that making accurate predictions may be influenced by a variety of factors (e.g., motivation, task familiarity, mode of assessment or training); factors which seem to be independent from metacognitive development. For example, Schneider (1998) reported that children often respond inconsistently when they have to predict performance, especially when a task is unfamiliar. Consequently, their predictions are not necessarily related to metacognitive deficits, but rather to motivational factors such as wishful thinking. The majority of children in our study (100% in the categorical and 93% in the specific condition) predicted they would remember the task, indicating that they may have been highly motivated to succeed. This may have increased not only children's motivation, but also their perception of importance of the PM task, in turn boosting their cognitive resources in order to detect the PM targets correctly (i.e., engaging strategic monitoring similarly to above). Similarly, manipulating importance of the PM task has shown to increase monitoring and improve PM cue detection, while interfering with the OT performance (see Walter & Meier, 2014 for a review).

However, the fact that in the present study almost all participants stated that they would remember rather than forget the PM task, may represent a limitation. Future studies should attempt to balance evaluations, providing a comparable number of children who predict to remember or forget. It would be reasonable to use more confidence levels, as well as to use different question types to the "yes/no-format". This would make it possible to compare PM performance between children who predicted successful performance with those who predicted the reverse , giving us a better insight in children's procedural MM in relation to PM. However, we suggest that delayed predictions, as well as post-dictions (that is, respectively, evaluations of the own performance while performing or after having performed the task; see Schneider, 2015), may be better indicators of procedural MM. Indeed, task experience may

enhance participants' insight into their likelihood of successful task performance, thus increasing precision. Alternatively, immediate performance predictions may be not a pure procedural MM measure, since they are likely to be distorted by other factors, as seen in both the present study and previous studies with adults (Meier et al., 2011; Rummel et al., 2011).

Besides the similar patterns of our results and the adult literature, one may question whether the differences between categorical and specific tasks were due to their sequence of administration, given that this variable was not counterbalanced. The mixed effects model approach has been adopted to eliminate confounding factors from the main effect, however, in the 30 days' interval between the first and second administration, children may have developed cognitively, thus, attaining better performance. Our data seemed to contradict this critique, given that it is unlikely that development occurred in the non-prediction group only. Moreover, other results were not affected by this potential confound, given that various effects (e.g., WM, declarative MM and inhibitory control) have been found in both categorical and specific tasks. Another critical point can be represented by accuracy rates on the specific PM task, which were high and nearly at ceiling. To clarify these points, a similar study has recently been conducted (Cottini, Basso, & Palladino, in preparation) in which 7-year-old children performed a comparable OT, but with specific PM cues only. In that study, task difficulty was higher, given that five PM targets were included instead of the three used in the present study. In the same way, only half of the children were asked to predict their performance. Preliminary results seem to corroborate our data, showing that performance predictions did not influence children's performance on a specific and more automatic PM task.

Summary and conclusions

The evidence emerging from our study is the first to demonstrate that performance predictions can be used with school-aged children as an effective strategy to improve performance on cognitively-demanding PM tasks. To our knowledge, it is also the first study exploring the relationship between PM, declarative MM and executive processes in school-aged children. Declarative MM has shown to play an important role, not only in the ability to 'remember the future', but also in the engagement of strategic monitoring processes. Since WM is another important factor for successful prospective remembering, in future research, it would be interesting to explore the effect of interventions on these two processes. Unlike the effect of training on cognitive processes, which may be time- and resource-demanding for children, and which may not always show transfer effects, providing simple strategies (such as thinking of possible future performance) may be effective in enhancing PM abilities. Similarly, this has been shown in a study including older adults, who benefitted substantially from an implementation intention strategy, in contrast to cognitive processing training (Brom & Kliegel, 2014). In future studies, it would be useful to compare the effect of different strategies, which can be used during intention formation for different PM tasks. Strategies such as implementation intentions (e.g. Basso & Olivetti Belardinelli, 2006; Gollwitzer, 1999) or episodic future thinking (e.g. Brewer, Knight, Meeks, & Marsh, 2011; Kretschmer-Trendowicz, et al., 2016) may be useful in more automatic PM tasks, since they seem to strengthen cue-action association. Alternatively, emphasizing PM task importance or making performance predictions can be used in more resource-demanding PM tasks. It would be also informative to study the processes underlying the effect of various strategies on different PM tasks. Future research should also include different age groups, as well as clinical populations with PM difficulties, such as children with autism spectrum disorder (e.g. Henry, Terrett,

Altgassen, Raponi-Saunders, Ballhausen, Schnitzspahn, & Rendell, 2014) or ADHD (e.g. Kliegel, Ropeter, & Mackinlay, 2006). Since executive and metacognitive processes in those populations are often insufficient, it may be useful to see whether these children's PM would benefit from performance predictions or other strategies.

In conclusion, the present study allowed us to define the contribution of explicit performance prediction, which indicated a positive effect on a resource-demanding PM task, as well as important roles of declarative MM and WM. This evidence may, in turn, allow us to determine the most important factors in implementing pragmatic educational procedures. Moreover, the present study may encourage future research to study development of PM in relation to declarative and procedural MM.

Acknowledgments