Frontline Learning Research
An official journal of EARLI

Frontline Learning Research Vol.3 No. 4 (2015) 37- 55
ISSN 2295-3159

Effects of a Short Strategy Training on Metacognitive Monitoring across the Life-span

Nicole von der Linden[1], Elisabeth Löffler, Wolfgang Schneider

University of Würzburg, Germany

Article received 29 July / revised 20 November / accepted 21 November / available online 18 January

Abstract

The present study was conducted to explore the potential positive influence of a short strategy training on metacognitive monitoring competencies covering a life-span approach. Participants of four age groups (3rd-grade children, adolescents, younger and older adults) concluded a paired-associate learning task. Additionally, they gave delayed Judgments-of-Learning (JOLs), that is, they rated their certainty that they would later be able to recall specific details correctly, and Confidence Judgements (CJs), that is, they rated their certainty that the provided answers in the recall test were correct. Half of the participants underwent a short strategy training in order to enhance their recollection of contextual details thus providing a diagnostic basis for forming metacognitive judgements. Results revealed significant gains in memory performance after completing the strategy training. Moreover, a positive effect of the strategy training on JOLs and CJs differentiation and accuracy could be detected. Effects were most pronounced for children and older adults. Participants who had completed the strategy training also reported a decrease of familiarity-based metacognitive judgments and were able to identify memories for which no reliable cues existed more easily than participants in the control condition. Accordingly, improvements in monitoring performance seemed to be due to a shift in underlying cues. In sum, this study integrates traditional aims from the relatively separately existing lines of metacognitive research in the developmental and cognitive literature and adds to understanding and improving monitoring judgments in a lifetime sample.

Keywords: metamemory, judgments-of-learning, confidence judgments, monitoring, life-span, strategy instruction

[1] *Corresponding author:* Nicole von der Linden, Department of Psychology, University of Würzburg, Röntgenring 10, 97070 Würzburg, Germany. Phone: +49(0)931 / 31 89067, Fax: +49(0)931 / 31 2763, Email: linden@psychologie.uni-wuerzburg.de DOI: http://dx.doi.org/10.14786/flr.v3i4.196

1. Introduction

Accurate metacognitive monitoring plays an important role in many everyday situations as well as in learning contexts (Schneider, 2010; Son & Metcalfe, 2000). In daily routines, metacognitive monitoring is for instance relevant when one has to decide about whether one has memorized the departure time of one's train, or whether one has taken appropriate notes of a lecture. Moreover, subjective monitoring judgments influence learning behaviors, especially the selection of to-be-studied items and the allocation of study time (see Son & Metcalfe, 2000, for a review). Structured learning situations are not only important to children and young adults but also for life-long learning which has gained importance in recent years. Yet, life-span perspective is still rare in the metacognitive monitoring literature.

Metacognitive research has traditionally been conducted within two main but separate lines of research: a developmental perspective (Flavell, 1999) which focuses on the changes of metacognitive abilities during life and a cognitive tradition (for an overview cf. Dunlosky & Metcalfe, 2009; Koriat & Levy-Sardot, 1999) which tries to explore mechanisms underlying metacognitive monitoring processes and its consequences for regulation of learning typically in an adult sample only. In the following, we present a study which is one of few existing attempts to combine both perspectives concerning metacognitive monitoring processes. Firstly, the present study aimed at exploring developmental trajectories of monitoring abilities in a life-span sample from early school age to older adulthood. A second purpose was to improve participants' monitoring competencies by reinforcing highly diagnostic cues in paired-associate learning situations (McCabe & Soderstrom, 2011; Robinson, Hertzog, & Dunlosky, 2006) and to investigate the role of familiarity and recollection-based cues for monitoring processes.

In learning situations two aspects of monitoring are of special interest: Judgments of Learning (JOL) and Confidence Judgments (CJ). According to Nelson and Narens' (1990) seminal model of procedural metamemory, JOLs provide subjective information about the degree to which encoded information has been mastered and can be potentially recalled during a future memory test (Nelson & Narens, 1990). Findings from studies using different age groups suggest that even young children can effectively monitor their learning progress under certain circumstances. On the one hand, the results indicate that immediate JOLs are typically inaccurate and also represent overestimations of one's actual performance. Remarkably, this is true not only for children of different ages but also for adults. Immediately after studying new information, judgments about its future recall seem severely biased by the false belief that information currently in short-term memory can be easily recalled some minutes later. Obviously, this bias operates similarly in participants of different ages. On the other hand, however, even young children can make rather accurate assessments of the subsequent recallability of items when this judgment is somewhat delayed, that is, when it takes place a minute or two after studying the item. In other words, even young children seem to have a good feeling for which items will be recallable and which will not when long-term memory information has to be accessed for the JOL (Schneider, 2015).

Confidence judgments (CJs) concern retrieval monitoring and are typically made after a response is given to indicate how sure participants are about the correctness of an answer. CJs are thought to reflect a substantive sense of certainty that arises from the strength of the memory that is being retrieved, and this sense of certainty has been interpreted as an indicator of memory accuracy (Ghetti, Lyons, Lazzarin, & Cornoldi, 2008; Roebers, 2002).

Metacognitive monitoring judgments are commonly believed to be based on multiple cues. This accessibility view has been proposed for immediate and delayed JOLs (Koriat, 1997; Metcalfe & Finn, 2008; Toth, Daniels, & Solinger, 2011) as well as for CJs (Kelley & Jacoby, 1996; Kelley & Sahakyan, 2003). The accuracy of those judgments depends on whether accessible cues are diagnostic of memory performance or not (Dunlosky & Metcalfe, 2009). Cues are highly diagnostic if they influence metacognitive judgments and recall performance in a similar way. Thus far multiple sources involved in the construction of monitoring judgments have been postulated. For example, research investigating immediate JOLs emphasizes encoding fluency as a major base, which in turn depends on different factors such as the concreteness and the frequency of the items (Begg, Duft, Lalonde, Melnick, & Sanvito, 1989) as well as familiarity of items

(Nelson & Narens, 1990). Delayed JOLs have been linked to retrieval fluency (Benjamin & Bjork, 1996) as well as success and ease of item retrieval (Nelson, Narens, & Dunlosky, 2004). Similar to JOLs, for CJs a number of different cues have been discussed to influence accuracy, among them perceived ease (Zakay & Tuvia, 1998) and vividness of retrieval (Robinson, Johnson, & Robertson, 2000).

Different attempts have been made to categorize these various types of cues (Koriat, 1997; Kelley & Jacoby, 1996). In recent years the literature has begun to discuss the distinction between familiarity-based and recollection-based cues for different metacognitive judgments (Daniels, Toth, & Hertzog, 2009; McCabe & Soderstrom, 2011; Metcalfe & Finn, 2008; Toth et al., 2011). Recollection is typically defined as the consciously controlled intentional use of memory that allows for the retrieval of qualitative details of a past event. This process is frequently associated with the subjective experience of vivid remembering. Familiarity, by contrast, usually refers to experiences of prior events that may arise from activated semantic representations. The relative contribution of recollection and familiarity may differ from task to task. Memory tasks that require participants to recall or recognize details about the target item rely heavily on recollection. So far, the literature lacks a systematic examination of the role of recollection and familiarity cues across different monitoring indicators and across the life-span.

Evidence suggests that both immediate (Daniels et al., 2009; Toth et al., 2011) and delayed JOLs (Metcalfe & Finn, 2008) are influenced by recollection and familiarity processes. Yet, delayed JOLs are mainly based on cues related to recollection such as target retrievability (for an overview see Metcalfe & Finn, 2008), whereas familiarity processes, such as processing fluency, have been identified as primary cues for immediate JOLs in younger adults (Matvey, Dunlosky, & Guttentag, 2001; Rhodes & Castel, 2008). To our knowledge, the role of familiarity and recollection processes underlying JOLs in children has not been examined to date. Concerning older adults, first evidence suggests that they have more problems with monitoring recollection processes than younger adults (Daniels et al., 2009). Several findings support the idea that recollection processes increase the accuracy of monitoring processes. First, as noted above, delayed JOLs have been shown to be more accurate than immediate JOLs for children, younger, and older adults (Connor, Dunlosky, & Hertzog, 1997; Koriat & Shitzer-Reichert, 2002; Nelson & Dunlosky, 1991; Schneider,Visé, Lockl, & Nelson, 2000). This can be explained by the fact that for delayed JOLs participants actively assess long-term memory (recollection processes), which is more predictive of recall than short-term memory used for immediate JOLs. Additionally, with delayed JOLs participants seem to rely more on idiosyncratic cues of encoding and remembering than with immediate JOLs (Koriat, 1997). Idiosyncratic cues refer to personal, item-specific details, for instance, images or associations. Providing a rich basis of idiosyncratic cues (e.g. through a strategy training) should facilitate the identification of recollection-based memories. Further evidence for the importance of recollection processes for accurate monitoring in younger adults is provided by the following fact: Focusing participants' attention to cues connected with target retrievability enhanced JOL accuracy compared to immediate JOLs (McCabe & Soderstrom, 2011). Furthermore, for recollection-based memories higher (Daniels et al., 2009; Toth et al., 2011) and more accurate immediate JOLs (Toth et al., 2011) could be found than for familiarity-based memories. In sum, more research on the role of familiarity and recollection processes for JOLs is needed, not only for children but also in terms of comparisons of broader age ranges. Yet, existing evidence points to the fact that the retrieval of contextual information seems to provide a reliable cue for later memory performance in different age groups.

Similarly, CJs seem to be based on recollection (analytic) or familiarity (non-analytic) components (Kelley & Jacoby, 1996; Kelley & Sahakyan, 2003). Recollection processes have been identified as playing a major role for CJ accuracy (Kelley & Sahakyan, 2003) and accuracy losses with increasing old age have been linked to recollection impairments (Kelley & Sahakyan, 2003; Wong, Cramer, & Gallo, 2012). For children the role of recollection and familiarity processes underlying CJs has not been addressed yet. Remember-know-judgments which are positively correlated with CJs (Holmes & Weaver, 2010) also emphasize the role of recollection and familiarity processes for metacognitive judgments. Remember-know-judgments distinguish whether a memory content is associated with specific contextual information

(recollection) or is familiarity-based, with participants unable to retrieve the personal encounter with a memory detail.

Although there is consensus that participants base their monitoring judgments on various cues, they do not always take into consideration the factors which are most predictive of memory performance (Koriat, 1997; Touron, Hertzog, & Speagle, 2010). As discussed above, recollection-based cues are considered to be important and also highly diagnostic for both JOLs and CJs. Therefore, a training program that aims at strengthening the accessibility of those cues should enhance monitoring accuracy (McCabe & Soderstrom, 2011). Yet, to our knowledge no systematic training in this area has been carried out. A training should be beneficial for all age groups but especially for older adults, among whom deficits in monitoring processes have been linked to problems with recollection processes for JOLs (Toth et al., 2011), CJs (Shing, Werkle-Bergner, Li, & Lindenberger, 2009; Wong et al., 2012) and Feeling-of-knowing judgments (Souchay, Bacon, & Danion, 2006). Children should also particularly profit from such a training procedure as developmental progression in monitoring skills seems to be influenced by improved retrieval processes for both JOLs (Koriat & Shitzer-Reichert, 2002) and CJs (Roderer & Roebers, 2011).

The study presented here was designed to fill this gap in the literature and to explore the effect of recollection and familiarity based cues on metacognitive monitoring. Although the role of recollection and familiarity-based processes for metacognitive processes has received more attention in recent years, available studies have focused on one type of metacognitive judgment only and involve one or at most two age groups (Daniels et al., 2009; Souchay et al., 2006). Especially empirical studies with children are scarce. Therefore the design included a life-span perspective to account for the life-long significance of learning. Moreover, two different types of metacognitive judgments (JOLs and CJs) were included in the present study in order to allow for direct comparison within one sample. Multiple but partly different cues seem to underlie delayed JOLs and CJs (Koriat, 1997, 2012) for they occur in different stages of the learning process. Compared to JOLs, strengthening recollection processes may have a somewhat greater effect on CJs as the longer interval to the learning stage might otherwise foster the reliance on familiarity, especially in older adults (Shing et al., 2009). Consequently, it is of interest to compare different monitoring indicators yet such studies are very rare in the literature (Leonesio & Nelson, 1990).

Specifically, participants of four age groups (early school age to later adulthood) were asked to complete a paired-associate learning task and to give delayed JOLs (which are more accurate than immediate JOLs across all of the different age groups; see above) and CJs. A paired-associate task was chosen in order to ensure comparability with related studies and because for this stimulus material ample evidence exists for the effectivity of strategy trainings across all included age groups (see below). Half of the participants underwent a strategy training in order to enhance their recollection of contextual details during JOL, CJ and test collection, thus providing a diagnostic basis for forming metacognitive judgments. To ensure that the training was transferable to rehearsing processes in everyday life and for different age groups a short instruction in mental imagery was chosen. In paired-associate learning, mental imagery has proven to be the most efficient way of processing (Richardson, 1998), and it effectively improves recall performance in different age groups from first grade on to older age (Richardson, 1998; Verhaeghen, Marcoen, & Goossens, 1992; Willoughby, Porter, Belsito, & Yearsley, 1999).

Even more relevant to our study, some recent studies provide first evidence for the fact that strategy use successfully improves monitoring processes for both JOLs and CJs although this research does not specifically explore the cues underlying monitoring judgments and hardly ever an explicit strategy training was done. Hertzog, Sinclair, and Dunlosky (2010) have shown that spontaneous strategy use (e.g. mental imagery), which was not induced but only accessed after JOL collection, substantially influenced JOLs and JOL resolution, that is, the accuracy with which a person can monitor the relative recallability of different items, in adults aged 18 to 81. Robinson et al. (2006) instructed but not trained younger and older adults to use a mental imagery strategy when memorizing pairs of items. They found that the size of the JOLs and recall performance were positively correlated with strategy use in both age groups and mental imagery was identified as a diagnostic cue for JOLs. To our knowledge no comparable studies exist for children. Thus in

JOLs, so far no attempts have been made to directly train subjects of a broad age range to apply an imagery strategy.

Concerning CJs, Nietfeld and Schraw (2002) trained college students to use various strategies for probability tests. As a result, subjects benefitted from the instructions both in terms of performance and monitoring accuracy (CJs). Besides Shing et al. (2009) showed that participants from 10 to 75 years of age benefitted from strategy training in terms of their CJs by enlarging the difference in CJs provided after hits compared to false alarms.

In accordance with the literature we expected a positive effect of strategy training on both recall processes and metacognitive processes (JOLs and CJs) in all age groups. As our study is the first to include a life-span approach to investigate the influence of recollection processes on metacognitive monitoring, developmental effects were of special interest. We proposed that children and older adults would benefit most from the strategy training: production deficits concerning strategy use are most pronounced in children and older adults (Naveh-Benjamin, Brav, & Levy, 2007; Pressley & Levin, 1977), and recall performance increases during childhood declines in older adulthood (Weinert & Schneider, 1996). Although generally little developmental progression is found for JOLs, recent evidence suggests that under certain circumstances deficits in recollection processes may play a role in lower JOL accuracy in older adults (Daniels et al., 2009; Toth et al., 2011). As for CJs, their accuracy has been shown to improve over the primary school years (Roebers, von der Linden, Howie, & Schneider, 2007), and they seem to be influenced by retrieval processes (Roderer & Roebers, 2010). Deficits in older adults' CJs have also been linked to deteriorated recollection processes (Kelley & Sahakyan, 2003). Additionally, we aimed to compare the effects of our training on different monitoring indicators as the importance of recollection processes might vary in different stages of the learning process.

Since we proposed that a strategy training should be effective mainly due to enhanced accessibility of recollection-based cues, we additionally asked participants to classify the basis of their recall as Recollection, Familiarity or No Memory (RFN-judgments). These classifications were successfully introduced for JOLs by Daniels et al. (2009) and Toth et al. (2011). We extended the use of RFN-judgments to CJs.

2. Method

2.1 Sample

A total of 160 (85 male, 75 female) participants of four age groups (40 children in 3rd grade, 40 adolescents in 7th and 8th grade, 40 younger adults between 19 and 26 years of age and 40 older adults between 60 and 75 years) took part in our study. This sample size surpasses the required number of $N = 132$ participants as determined by an a-priori power-analysis which was conducted with the premise to detect medium-sized effects according to Cohen ($d = .25$). They were recruited via contacting their schools directly and via newspaper and internet advertisements. Children and adolescents received small gifts, whereas the other participants got 10-Euro vouchers or were paid in cash. Subjects' mean ages were 8.38 ($SD = 0.49$) for the children, 12.73 ($SD = 0.72$) for the adolescents, 22.75 ($SD = 2.02$) for the younger adults and 68.40 ($SD = 4.08$) for the older adults.

2.2 Materials

The learning items consisted of pairs of concrete German nouns from different semantic categories (e.g. zoo animals, furniture, clothing etc.). To vary the difficulty, one half of the pairs represented two words

from the same category and the other half of the pairs comprised words from two different categories. The item list for children and older adults consisted of 45 word pairs in the study phase and 60 word pairs in the recognition phase (the latter including 30 pairs identical to the study phase, 15 newly matched pairs and 15 completely new pairs). Item pairs for adolescents included 54 word pairs in the study phase and 72 items in the recognition phase. The numbers for younger adults were 60 and 80 item pairs respectively. Four practice pairs not included in the analysis preceded each single phase of the experiment.

The appearance of word pairs as identical or recombined was counterbalanced among the subjects. The order of presentation was randomized as well.

2.3 Procedure

The consent of the parents and of the school was obtained for the children and adolescents before the beginning of the study. Participants were tested individually in quiet rooms in the school or in the laboratory.

Half of the subjects of each age group were randomly assigned to the strategy-instruction condition (experimental group). They received instructions on visual imagery. The test administrator first explained the advantages of memorizing word pairs as one interconnected image emphasizing the importance of integration of the two images. The explanation was facilitated by two drawings, one of a frog carrying a banana and one of a candle burning a letter. Then the subjects in the experimental condition had to practice this visual imagery strategy by means of ten word pairs which were different from those used in the experiment. They were given feedback on the quality of their imagery and were asked to imagine another combined image if necessary. The instruction was standardized for all participants with the restriction that the wording in children's was slightly simplified. Participants of all age groups reported that they easily understood the strategy. The participants in the experimental group were instructed to use the visual imagery strategy while memorizing the items. In the control group no strategy instruction was given.

The word pairs were presented on a computer screen with presentation rates of 8 seconds per item pair for children and older adults, 6 seconds for adolescents, and 2.5 seconds for younger adults. The presentation rates were adapted in order to control for baseline difficulty between the age groups. Subjects were instructed to concentrate on the pairs because they later would have to recognize them and to indicate whether the word pair had appeared in the study phase or not.

In the JOL phase, each left noun of the item pair (stimulus) was presented on the screen in the same order as in the learning phase. To avoid relearning only the stimulus was shown. Subjects were asked to indicate the likelihood of recognizing the word pair in about 30 minutes. JOLs were rated on a thermometer scale from 0 (very unsure) to 100 (very sure) successfully used in previous studies (Koriat, Ackerman, Lockl, & Schneider, 2009; Koriat & Shitzer-Reichert, 2002).

In the recognition phase, word pairs of each type (i.e., either identical, recombined, or new) were presented. Participants had to indicate by opting for "yes" or "no" whether they thought that the item had appeared in exactly this combination in the studying phase. After that, the item disappeared from the screen in order to avoid distraction from the presented item pair, and participants were asked to indicate how they generated the yes-or-no-decision. They had to decide between three options: a) "I can remember the word pair very well" (Recollection), b) "The word pair seems familiar to me" (Familiarity), c) "I cannot remember the word pair at all" (No memory). Finally, subjects had to indicate on a hot-cold-scale equivalent to that used for JOLs how sure they were that the given answer was correct (CJ).

At the end of the session, the test administrator asked whether the subjects in the experimental condition had applied the strategy while memorizing the items. Participants in the control group were asked whether they had employed any strategy, and if so, to specify the strategies and to indicate how often they were used.

3. Results

A preliminary analysis assessing the effect of gender did not reveal any systematic differences between male and female participants. Thus data were collapsed across this variable. Scheffé tests were used as a post-hoc follow-up on main effects. The level of significance was set to $p < 0.05$.

In a first step of analysis, we assessed memory performance in terms of the percentage of correctly recognized items, that is, either identical items correctly recognized as "old" or recombined or completely new items correctly classified as "new" items. Next, we analyzed JOLs and CJs as indicators of metacognitive monitoring. Finally, we will report changes in the RFN judgments as cues for monitoring processes. Results are reported as a function of age group and experimental condition in order to examine the influence of cognitive development and strategy instruction on recognition rates and metacognitive monitoring.

3.1 Recognition rates

The first column of Table 1 shows the mean proportion of correctly recognized items as a function of age group and experimental condition, that is overall recognition rates. An ANOVA with age group and experimental condition as between-subject factors revealed a main effect of age group ($F(3,152) = 5.34$; $p < .01$; $\eta^2 = .10$). A post-hoc analysis indicated that younger adults performed significantly better than children (.80 vs. .71 correct, respectively). Furthermore, a main effect of the experimental condition was found ($F(1,152) = 21.82$; $p < .001$; $\eta^2 = .13$), indicating that those participants who had received the strategy instruction recognized significantly more word pairs correctly than participants who had not received such an instruction (.80 vs. .72, respectively).

The second to fourth column of Table 1 splits the recognition rates into percentages depending on the type of word pair: that is, whether the word pair in the recognition phase was identical to that in the study phase or whether it was recombined or a completely new word pair. Inferential statistics were conducted separately for each word pair in order to facilitate the interpretation. For the identical word pairs, an ANOVA with age group and experimental condition as between-subject factor revealed a significant main effect of age group ($F(3,152) = 5.14$; $p < .01$; $\eta^2 = .09$). A post-hoc analysis showed that children (.65) recognized fewer of the identical items correctly than younger (.77) and older adults (.76). Furthermore, the main effect of the experimental condition reached significance ($F(1,152) = 5.24$; $p < .05$; $\eta^2 = .03$) with subjects in the experimental group (.74) recognizing more items correctly than subjects in the control group (.69). For the recombined word pairs, only the main effect of strategy instruction reached the significance level ($F(1,152) = 5.24$; $p < .001$; $\eta^2 = .11$), with subjects in the strategy instruction group recognizing more of the recombined word pairs correctly (.75) than subjects who had not received the strategy instruction (.61). As for the new word pairs, a significant main effect of strategy instruction was found as well ($F(1,152) = 7.04$; $p < .01$; $\eta^2 = .04$). Those participants who had received the strategy instruction recognized more of the new word pairs correctly (.94) than those who had not been thus instructed (.88).

Table 1

Recognition rates as a function of age group, experimental condition, and type of word pair

Age group	Type of word pair			
	Overall	Identic word pairs	Recombined word pairs	New word pairs
Children				
Control group	.65 (.06)	.60 (.16)	.54 (.19)	.88 (.13)
Experimental group	.77 (.11)	.71 (.18)	.72 (20)	.91 (.10)
Adolescents				
Control group	.70 (.11)	.64 (.17)	.63 (.20)	.89 (.11)
Experimental group	.81 (.09)	.76 (.12)	.76 (.15)	.94 (.09)
Younger Adults				
Control group	.79 (.09)	.75 (.10)	.73 (.19)	.92 (.10)
Experimental group	.81 (.12)	.79 (.15)	.75 (.23)	.93 (.10)
Older Adults				
Control group	.74 (.12)	.79 (.14)	.53 (.28)	.85 (.21)
Experimental group	.80 (.11)	.73 (.18)	.77 (.22)	.95 (.08)

Standard deviations are in parentheses.

3.2 Metacognitive Monitoring

3.2.1 Mean JOLs before correct vs. incorrect responses

Figure 1 shows participants' mean JOL ratings as a function of the correctness of the subsequent response, age group, and experimental condition. An ANOVA with correctness of response as within-subject factor and age group and experimental condition as between-subject factors revealed a significant main effect of correctness of response ($F(1,151) = 135.38$; $p < .001$; $\eta^2 = .47$): Subjects gave higher JOLs before correct (57.98) than before incorrect responses (46.78). In addition, a significant interaction between the factors correctness of response and experimental condition was found ($F(1,151) = 6.22$; $p < .05$; $\eta^2 = .04$). Furthermore, the triple interaction between correctness of response, experimental condition, and age group attained a significant level ($F(3,151) = 3.43$; $p < .05$; $\eta^2 = .06$). In order to examine the direction of the interactions post hoc, we analyzed the experimental and the control group data separately. For subjects in the experimental condition, an ANOVA with correctness of response as within-subject factor and age group as between-subject factor revealed a main effect of correctness of response ($F(1,75) = 89.07$; $p < .001$; $\eta^2 = .54$) with mean JOLs being higher before correct (59.58) than before incorrect responses (45.98). For the participants in the control condition the main effect of correctness of response was also significant ($F(1,76) = 47.36$; $p < .001$; $\eta^2 = .38$). Furthermore, for the subjects in the control condition a significant interaction between correctness of response and age group was found ($F(3,76) = 5.93$; $p < .01$; $\eta^2 = .19$). Subsequent analyses revealed that only the adolescents and the younger adults distinguished between correct and incorrect responses given that it was only for these two age groups that the factor correctness of response turned out to be significant (children: $F(1,19) = 0.66$; $p = .426$; $\eta^2 = .03$; adolescents: $F(1,19) = 19.82$; $p < .001$; $\eta^2 = .51$; younger adults: $F(1,19) = 66.06$; $p < .001$; $\eta^2 = .78$; older adults: $F(1,19) = 4.43$; $p = .05$; $\eta^2 = .19$).

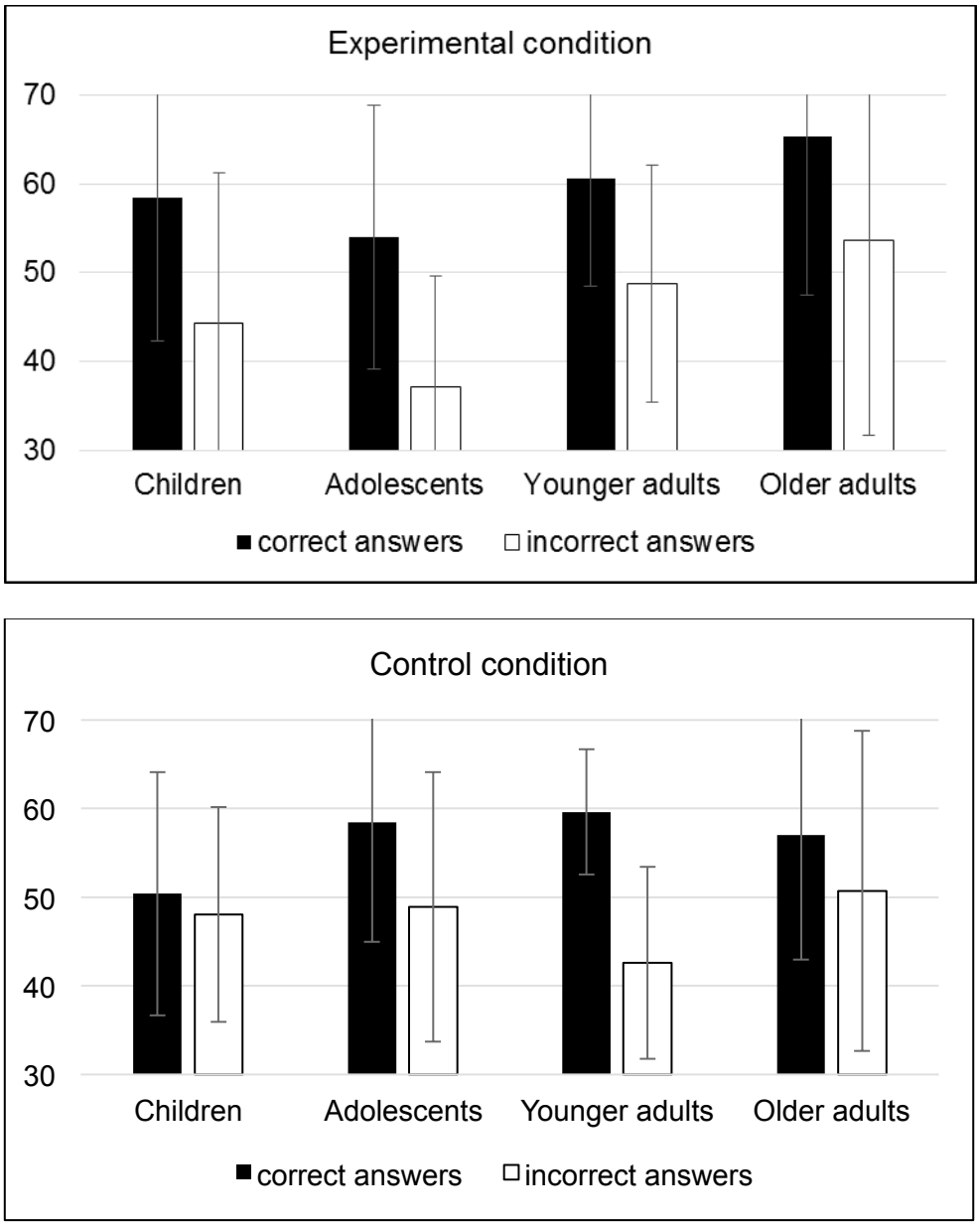

Figure 1. Mean JOLs preceding correct vs. incorrect answers as a function of age group and experimental condition.

3.2.2 JOL accuracy

In order to assess JOL accuracy as a function of age group, question format and experimental condition, Goodman-Kruskal gamma correlations between JOLs and recall performance were computed for each participant, and then averaged for each single cell in the experimental design. Gamma correlations are considered to be the most appropriate measure of metacognitive accuracy (Nelson, 1984) and are commonly used in the contemporary literature (Nelson & Dunlosky, 1991; Schneider et al., 2000). A positive correlation indicates that higher JOLs were given for items that were recalled correctly than for those recalled incorrectly.

Table 2 shows mean gamma correlations for JOLs as a function of age group and experimental condition. One-tailed t-tests revealed that all gamma correlations were different from zero for almost all

groups. The only exception concerned the children in the control group whose mean gamma correlations for JOLs were not significantly different from zero.

Table 2

Mean gamma correlations as a function of age group and experimental condition

Age group	JOLs	CJs
Children		
Control group	.03 (.24)	.18 (.20)
Experimental group	.32 (.28)	.38 (.29)
Adolescents		
Control group	.20 (.19)	.33 (.21)
Experimental group	.30 (.19)	.42 (.21)
Younger Adults		
Control group	.32 (.15)	.42 (.20)
Experimental group	.27 (.23)	.35 (.23)
Older Adults		
Control group	.14 (.26)	.38 (.26)
Experimental group	.29 (.25)	.49 (.26)

Standard deviations are in parentheses.

An ANOVA with age group and experimental condition as between-subject factors revealed a significant main effect of strategy instruction ($F(1,151) = 11.36$; $p < .001$; $\eta^2 = .07$). In addition, the ANOVA showed a significant interaction between age group and experimental condition ($F(3,151) = 3.58$; $p < .05$; $\eta^2 = .07$). In order to examine the direction of this effect, the mean gamma correlations of each age group were tested individually using univariate ANOVAs. In children, the main effect of experimental condition was significant ($F(1,38) = 12.25$; $p < .01$; $\eta^2 = .24$) with children in strategy-instruction group having higher gamma correlations (.32) than those in the control group (.03). In older adults, the results pointed into the same direction (experimental group: 29; control group: .14): the main effect of experimental condition was just short of being significant ($F(1,38) = 3.29$; $p = .077$; $\eta^2 = .08$).

3.2.3 Mean CJs after correct vs. incorrect responses

Differentiation in CJs was analyzed in the same way as for JOLs (cf. figure 2): mean CJ ratings after correct and incorrect responses respectively were calculated for each subject. As for JOLs, an ANOVA with correctness of response as within-subject factor, and age group and experimental condition as between-subject factors was conducted. The ANOVA revealed a main effect of age group ($F(3,151) = 6.26$; $p < .001$; $\eta^2 = .11$). Post-hoc tests according to Scheffé's procedure showed that younger adults gave significantly lower mean CJs (65.56) than the other age groups (children: 76.10, adolescents: 74.59, older adults: 75.00), regardless whether the answer was correct or not. In addition, a main effect of correctness of response was found ($F(3,151) = 188.75$; $p < .001$; $\eta^2 = .56$). Subjects of all age groups gave higher ratings after correct (79.02) than after incorrect responses (66.60). Finally, the interaction between correctness of response and age group reached the significance level ($F(3,151) = 5.61$; $p < .001$; $\eta^2 = .10$). To further explore the direction of the effect, separate ANOVAs for CJs after correct and incorrect answers were conducted. For correct answers, the main effect of age group reached significance ($F(3,156) = 3.42$; $p < .05$; $\eta^2 = .06$). Subsequent post-hoc analyses showed that older adults (82.57) were more confident after correct responses than younger adults (74.25). For the CJs after incorrect answers, a significant main effect of age group was found as well ($F(3,155) = 7.89$; $p < .001$; $\eta^2 = .13$). Here, post-hoc analyses showed that younger adults gave lower CJs (57.33) after incorrect responses than the other three age groups (children: 72.63; adolescents: 69.00; older adults: 67.42).

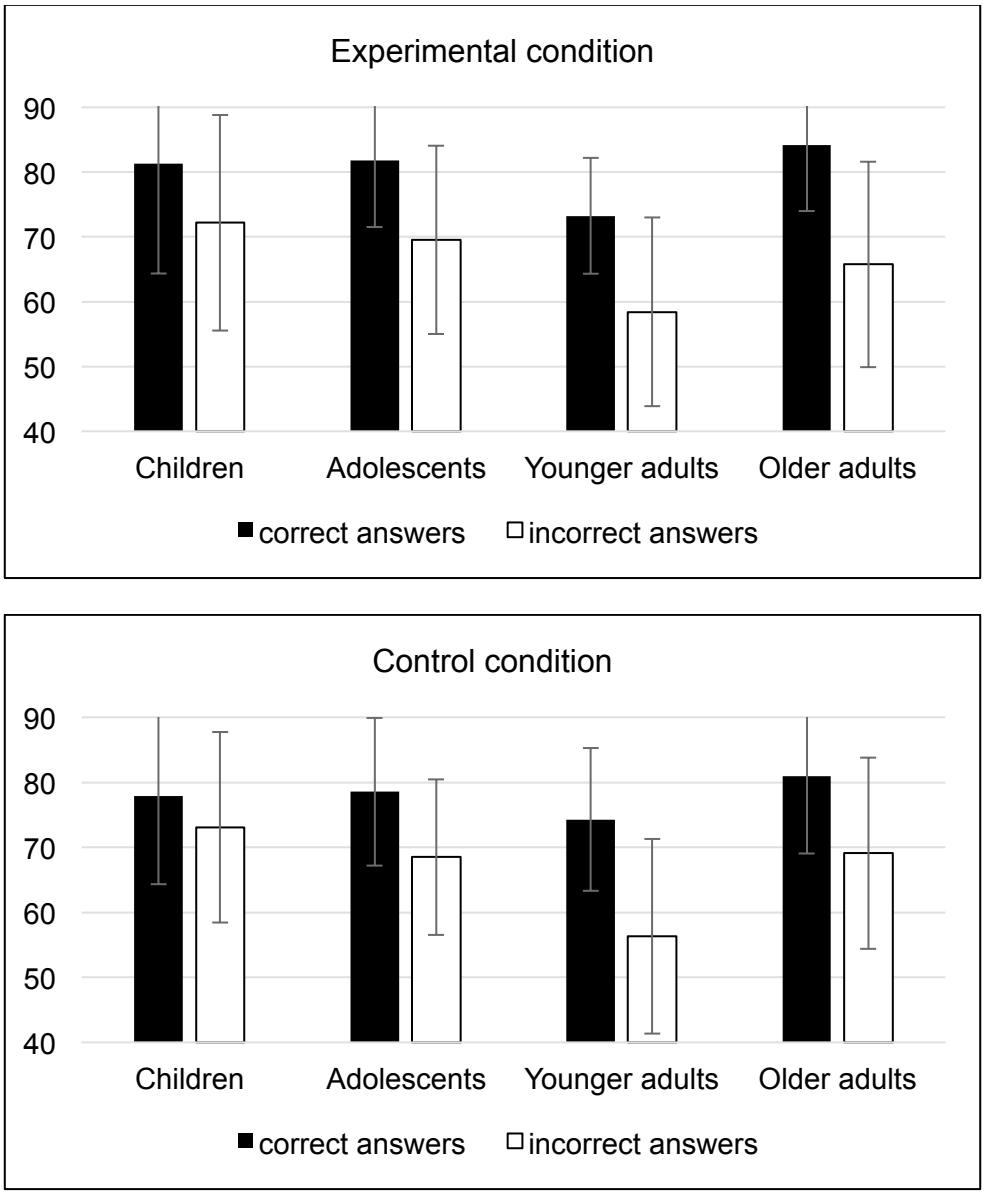

Figure 2. Mean CJs after correct vs. incorrect answers as a function of age group and experimental condition

3.2.4 CJ accuracy

CJ accuracy was assessed in the same way as JOL accuracy. Mean gamma correlations are displayed in Table 2. All gamma correlations were different from zero (using one-tailed t-tests). An ANOVA with age group and experimental condition as between-subject factors revealed a main effect of age group ($F(3,151) = 2.93$; $p < .05$; $\eta^2 = .06$). Post-hoc tests according to Scheffé showed a significant difference between children (.28) and older adults (.44). Furthermore, a main effect of strategy instruction was found ($F(1,151) = 4.75$; $p < .05$; $\eta^2 = .03$). Subjects in the strategy instruction condition had higher mean gamma correlations (.41) than participants without such instruction (.33).

3.3 RFN-judgments

In addition, the RFN-judgments subjects made were contrasted. We compared the percentage of how often each option was picked out because age groups differed in quantity of items. Table 3 shows the percentage of each RFN-judgment as a function of age group and experimental condition. An ANOVA was

conducted with RFN-judgment as within-subject factor and age group and experimental condition as between-subject factor. We found a significant main effect of RFN-judgment ($F(1,152) = 29.99$; $p < .001$; $\eta^2 = .17$). Paired contrasts revealed that subjects chose "familiarity" (.22) less often than "recollection" (.40) and "no memory" (.38). Furthermore a significant interaction between RFN-judgment and experimental condition was found ($F(1,152) = 4.12$; $p < .05$; $\eta^2 = .03$). Separate ANOVAs for each judgment showed that the strategy instruction had a significant effect only for the answer "familiarity" ($F(1,158) = 12.63$; $p < .01$; $\eta^2 = .07$), and "no memory" ($F(1,158) = 5.54$; $p < .05$; $\eta^2 = .03$): subjects in the experimental group chose the "familiarity" option less often (.19) than subjects in the control condition (.26), and the "no memory" option more often (.41) than the control group (.34).

Table 3

RFN-judgment in percentage of chosen option as a function of age group and experimental condition

Age group	Recollection	Familiarity	No memory
Children			
Control group	.41 (.24)	.27 (.15)	.32 (.25)
Experimental group	.35 (.23)	.19 (.11)	.45 (.21)
Adolescents			
Control group	.41 (.19)	.30 (.16)	.29 (.18)
Experimental group	.42 (.18)	.20 (.11)	.38 (.20)
Younger Adults			
Control group	.35 (.16)	.24 (.10)	.40 (.17)
Experimental group	.41 (.22)	.22 (.11)	.38 (.17)
Older Adults			
Control group	.42 (.21)	.22 (.11)	.36 (.16)
Experimental group	.40 (.20)	.15 (.09)	.45 (.22)

Standard deviations are in parentheses.

3.4 Spontaneous and instructed strategy use

In a last step of analysis, we assessed the outcomes for the Strategy Use Questionnaire. Table 4 shows how many participants of each age group and experimental condition reported having used the visual imagery strategy. Participants' open responses were categorized as visual imagery by two independent raters (kappa = .95).

We compared the use of the visual imagery strategy in the experimental and the control group. An ANOVA with age group and experimental condition as between-subject factors revealed a significant main effect of age group ($F(3,154) = 10.97$; $p < .001$; $\eta^2 = .18$) with post-hoc analysis showing that younger adults (.83) applied mental imagery more often than the other age groups (children: .48; adolescents: .53; older adults: .55). Furthermore, we found a significant main effect of experimental condition ($F(1,154) = 221.46$; $p < .001$; $\eta^2 = .63$): Participants who received the strategy instruction used mental imagery much more often than participants in the control condition (.28 vs. .95). The interaction between age group and experimental condition was also significant ($F(3,154) = 9.38$; $p < .001$; $\eta^2 = .16$). Separate analyses carried out for participants in the experimental group on the one hand and participants in the control group on the other hand showed there was no main effect of age group for participants who received the strategy instruction, indicating that the participants were able to transfer the training to the learning process. For subjects in the control group, the ANOVA revealed a main effect of age group ($F(3,74) = 12.30$; $p < .001$; $\eta^2 = .35$). A post-hoc analysis showed that the percentage of younger adults (.65) spontaneously applying a mental imagery strategy was significantly higher than that of the other age groups. More specifically, none of the

children, only 5% of the adolescents, and about 25% of the older adults applied such a strategy spontaneously.

Table 4

Percentage of subjects reporting the use of visual imagery as a function of age group and experimental condition

Age group	Children	Adolescents	Younger Adults	Older Adults
Control group	.00	.05	.65	.25
Experimental group	.95	1.00	1.00	.85

4. Discussion

The present study is among the first to explore metacognitive monitoring skills across the life-span, and also to investigate the effects of a memory strategy training principally suited to improve skills in this domain. Thus our study combined traditional interests of the developmental and cognitive literature on metacognition. In particular, we focused on possible positive effects of strengthening highly diagnostic cues (Koriat, 1997). This was achieved by training half of our participants in visual imagery before memorizing item pairs, and by assessing the training effects on monitoring quality, that is, JOL and CJ differentiation and accuracy. We postulated that instructing subjects to connect idiosyncratic content to items should lead them to rely less on familiarity but to focus on recollection processes when monitoring their performance. An important innovative aspect of our study was its life-span perspective in that four age groups (children in third grade, adolescents in seventh and eighth grade, younger and older adults) were included in the sample. Especially for children only very few studies exist that have explored the cues underlying monitoring judgments but for all included age-groups more research on the effects of familiarity and recollection-based cues is needed.

First, the results show that our manipulation of task difficulty across the age groups was successful: recognition rates in both the experimental and the control condition were of comparable height across the four age groups. Thus, the task was suitable for participants from primary school to older adulthood. Secondly, we found significant gains in recognition performance in subjects who underwent the strategy training. This effect was most pronounced for recombined and new item pairs but still substantial for the overall data. These results thus point to the fact that our experimental manipulation was successful and are in line with many previous findings: It has been shown for various age groups from primary school to older adulthood that visual imagery is an efficient strategy in paired-associate learning (Richardson, 1998; Verhaegen et al., 1992), and that its instruction leads to superior recall performance compared to spontaneous use (Shing et al., 2009). Also in accordance with the literature, self-reported use of visual imagery was higher in the experimental than in the control group, and substantial spontaneous use of visual imagery was only reported by young adults.

We acknowledge that we cannot rule out the possibility that pre-training differences had an effect on memory performance. Yet, participants in our study were randomly assigned to experimental and control group which substantially reduces the risk of imbalance in potentially confounding factors. Although our sample was at the lower end of recommended sample sizes for randomization (Bortz & Döring, 2009), the expected positive results of the strategy training on recognition performance and the reported low level of

strategy use in the control group speak for a true effect of the strategy training. However, in future research the inclusion of a pre-training measurement would further substantiate the results.

In accord with our hypothesis, we found evidence that both JOL differentiation between later correct and incorrect answers and JOL accuracy as measured by Goodman-Kruskal gamma correlations were enhanced by our strategy training in certain age groups. Concerning JOL differentiation, in the experimental group subjects of all age groups differentiated between correct and incorrect answers compared to the control group where only adolescents and younger adults gave higher JOLs before correct than before incorrect responses. Additionally, adolescents in the experimental condition descriptively showed a more pronounced discrimination between later correct and incorrect answers, as compared to those in the control condition. Concerning JOL accuracy, the analysis revealed that only children's accuracy improved significantly by the strategy instruction. Although there was a similar tendency in the group of older adults, only a marginal effect of strategy training was found. Similarly, adolescents' and younger adults' gamma correlations were not enhanced by the training.

We also detected the expected positive effect of visual imagery on CJ quality. For CJ accuracy, we found a significant main effect of strategy instruction, implying that all age groups benefitted from the strategy instruction. Concerning differentiation, training effects were found for children and older adults; here, the difference between correct and incorrect answers was about twice as high in the experimental condition than in the control condition. In contrast, adolescents and younger adults showed about the same amount of differentiation in both conditions.

The strategy training had positive effects on both JOLs and CJs which were most pronounced for children and older adults for both monitoring indicators. Yet, impact on CJs could be detected in a broader age range than for JOLs. This points to the fact that both JOLs and CJs rely on cues involved in our strategy instruction but at the same time draw onto different sources. For CJs information from the retrieval process might be most significant and it is possible that cues based on recollection processes are even more important for accurate judgments than for JOLs. Further research is needed to clarify this point.

In sum, the results confirm our predictions that a strategy training can improve the quality of metamemory monitoring judgments. This finding is in line with outcomes of other studies that have shown a positive influence of strategy use on prospective and retrospective monitoring judgments for children, younger and older adults (Hertzog et al., 2010; Nietfield & Schraw, 2002; Robinson et al., 2006; Shing et al., 2009) and expands the existing literature by training strategy use explicitly and by the inclusion of two monitoring judgments.

The developmental trends found in our study were also as expected: That is, children and older adults benefitted most from strategy instruction in terms of enhancing both their JOL and CJ quality, followed by the adolescents. These results are in accordance with developmental trends in regard to production deficits concerning strategy use (Naveh-Benjamin et al., 2007), and of recollection processes in general (Ghetti & Angelini, 2008). This outcome also emphasizes the practicability of our training. It proved to be effective yet was simple enough to be understood by elementary school children, and could also be successfully acquired by older adults in very short time.

Still, the strategy training did not account for much monitoring improvements in younger adults, with the exception of CJ accuracy. One possible explanation for this finding is the high level of spontaneous use of visual imagery reported by young adults in the control group. It appears likely that the short strategy training did not greatly improve the already high level of strategy use in young adults. Obviously, young adults showed high competence to memorize and to monitor their recall performance in paired-associate learning without further instruction. It is possible and should be investigated in further research that more pronounced effects of a strategy training would be found on more complex tasks. Support for this assumption comes from studies where gains from a strategy training in CJ accuracy could be shown for a comprehensive problem-solving task (Nietfield & Schraw, 2002).

Presumably further reasons were responsible for the fact that we were not able to confirm the positive effects of the strategy instruction for all age groups and for all indicators of metacognitive monitoring. One possible cause is the influence of the memory paradigm. A recognition task was chosen in this study in order to explore the basis of memory and monitoring processes by collecting RFN-judgments. Specifically, participants were asked to rate the quality of their recognition memory as recollection, familiarity or no memory as an indication of the mode of action of our strategy training. Yet, it seems possible that recognition processes make differentiation of recollection and familiarity-based cues more difficult than recall, given that no active memory retrieval was necessary. Recall processes seem to offer more cues to increase the accuracy of monitoring judgments, as compared to recognition (Buratti & Allwood, 2012). Other research showed a positive effect of strategy use on monitoring accuracy required active memory recall (Hertzog et al., 2010; Robinson et al., 2006). Although Shing and colleagues (2009) found positive effects of a strategy training on CJs in a recognition task, they used more complicated stimuli (Malay word pairs) than those used in this study. Furthermore they collected metacognitive measures of calibration and not resolution as done here.

Another explanation for this unexpected outcome could be that we collected delayed JOLs which have been shown to be more accurate than immediate JOLs (Nelson & Dunlosky, 1991). This seems to be due to the fact that delayed JOLs in all age groups are based on active assessment of long-term memory (recollection processes) instead of short-term and long-term memory as in immediate JOLs. Yet, we wanted to explore a possible add-on effect to maximize the quality of monitoring judgments. This in turn makes it more difficult to show an effect than with immediate JOLs which are commonly used in many studies (Daniels et al., 2009; Hertzog, Fulton, Mandiwala, & Dunlosky, 2013; Robinson et al., 2006).

A third possibility is that the effects of a short strategy intervention as used here are generally limited. Possibly results of a longer intervention could exceed the promising findings of our training in all age groups and for both monitoring indicators. This issue would be worth to be explored in a follow-up study.

Yet in general, a strategy instruction proved to be a promising starting point to influence monitoring processes in different age groups and across various monitoring indicators. As a mode of operation we proposed that the strategy instruction should be effective due to a shift in accessible cues. Specifically, we assumed that improvement should be due to the fact that now sources of monitoring judgments should be less familiarity-based cues and increasingly recollection-based cues. RFN-judgments confirm that - as predicted - the number of familiarity based judgments was significantly reduced in the experimental group. This trend was accompanied by more “no memory” responses. We assume that subjects in the experimental group profited from the instruction in that they were able to decide for which memories no reliable cues existed. In such cases, the answer “no memory” was correctly given. Participants could have used recall of interactive imagery to discriminate recollection states: If they can recall something about the image created to memorize a word pair, they are more confident that they base their delayed-JOL on a diagnostic cue. Strategy recall seems to have a similar effect on Feeling-of-knowing judgments (Hertzog, Fulton, Sinclair, & Dunlosky, 2014). Thus, although the number of recollection-based memories as perceived by the participants could not be enhanced, the strategy training seems to have increased participants’ awareness of possible cues and enabled them to distinguish more securely between real memories and no memories. This contrast of correct recall and no retrieval at the time of the JOL has been shown to be the most important source for high accuracy of delayed-JOLs (Nelson et al., 2004). It appears likely that the number of guesses which probably fell into the familiarity category could be successfully reduced by our training. Given that no age effects were found, the instruction seems to be effective in a similar way across all age groups in that it reduces the impact of familiarity.

In sum, the present study yields evidence that a strategy training is a suitable means to improve prospective and retrospective monitoring processes throughout the life-span, especially in children and older adults. The instruction used in this study proved to be an economic procedure that could be successfully applied in different age groups. The training was simple enough to be easily mastered by both elementary school children and older adults. Therefore a transfer to everyday-life situations seems possible.

The findings of the present study emphasize the significance of recollection-based cues as well as its distinction from other cues for metacognitive monitoring processes and encourage further research in this direction. Especially an expansion of our findings to more complex stimuli like texts or films and the investigation of the effect of a more elaborate strategy training are of interest. This would allow to further test relevance for every-day life. Additionally, it would be interesting to explore the effects of a strategy training on a larger samples as our study included relative small sample sizes per group and to follow up long-term effects of the training. We do not exclude the possibility that multiple cues underlie and can significantly influence monitoring processing. Yet, along with recent research (Ghetti et al., 2008; McCabe & Soderstrom, 2011; Metcalfe & Finn, 2008; Toth et al., 2011) we assume that exploring the role of recollection and familiarity processes in mediating the accuracy of monitoring judgments is a promising issue for future research.

Improving monitoring processes is of great importance as it is very closely linked to memory performance. Thus successful monitoring represents a very valuable competence in different learning contexts. Our results demonstrate that monitoring occurs from childhood on, but that there is still room for improvement at every age level. At the same time, the findings also illustrate that there are very economical ways to improve metacognitive monitoring in different age-groups. They thus indicate a direction which is worth to be pursued in future research.

Keypoints

- Approaches to improve metacognitive monitoring in a broad age range covering the life-span are still very rare.
- Integration of traditional developmental and cognitive research questions, as in this study, are scarce in the metacognitive literature.
- Our results show that a short training in visual imagery enhances both memory and metamemory performance, especially in children and older adults.
- Improvements in monitoring seem to be associated to a use of more reliable cues after the strategy training.

Acknowledgments

This research was conducted as part of a research project on the development of procedural metacognitive knowledge across the life-span and financed by the German Research Foundation (DFG-Gz. SCHN 315/45-1). We wish to thank all participating children, adolescents and adults as well as teachers, principals and parents for their cooperation.

References

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language, 28*, 610-632. DOI: 10.1016/0749-596X(89)90016-8

Benjamin, A. S., & Bjork, R. A. (1996). Retrieval fluency as a metacognitive index. In L. Reder (Ed.), *Implicit memory and metacognition* (pp. 309-338). Hillsdale, NJ: Erlbaum. DOI: 10.1037/0096-3445.127.1.55

Bortz, J., & Döring, N. (2009). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer.

Buratti, S., & Allwood, C. M. (2012). The accuracy of meta-metacognitive judgments: Regulating the realism of confidence. *Cognitive Processing, 13*, 243-253. DOI: 10.1007/s10339-012-0440-5

Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging, 12* (1), 50-71. DOI: 10.1037/0882-7974.12.1.50

Daniels, K. A., Toth, J. P., & Hertzog, C. (2009). Aging and recollection in the accuracy of judgments of learning. *Psychology and Aging, 24*, 494-500. DOI: 10.1037/a0015269

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Thousand Oaks, CA: Sage Publications, Inc.

Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology, 50,* 21-45. DOI: 10.1146/annurev.psych.50.1.21

Ghetti, S., & Angelini, L. (2008). The development of recollection and familiarity in childhood and adolescence: Evidence from the dual-process signal detection model. *Child Development, 79*, 339-358. DOI: 10.1111/j.1467-8624.2007.01129.x

Ghetti, S., Lyons, K. E., Lazzarin, F., & Cornoldi, C. (2008). The development of metamemory monitoring during retrieval: The case of memory strength and memory absence. *Journal of Experimental Child Psychology, 99*, 157-181. DOI: 10.1016/j.jecp.2007.11.001

Hertzog, C., Fulton, E. K., Mandviwala, L., & Dunlosky, J. (2013). Older adults show deficits in retrieving and decoding associative mediators generated at study. *Developmental Psychology, 49*, 1127-1131. DOI: 10.1037/a0029414

Hertzog, C., Fulton, E. K., Sinclair, S. M., & Dunlosky, J. (2014). Recalled aspects of original encoding strategies influence episodic feelings of knowing. *Memory and Cognition, 42,* 126-140. DOI: 10.3758/s13421-013-0348-z

Hertzog, C., Sinclair, S. M., & Dunlosky, J. (2010). Age differences in the monitoring of learning: Cross-sectional evidence of spared resolution across the adult life span. *Developmental Psychology, 46*, 939-948. DOI: 10.1037/a0019812

Holmes, A. E., & Weaver III, C. A. (2010). Eyewitness memory and misinformation: Are remember/know judgments more reliable than subjective confidence? *Applied Psychology in Criminal Justice*, *6*, 47-61.

Kelley, C. M. & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language*, 35, 157-175. DOI: 10.1006/jmla.1996.0009

Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language, 48*, 704-721. DOI: 10.1016/S0749-596X(02)00504-1

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*, 349-370. DOI: 10.1037/0096-3445.126.4.349

Koriat, A. (2012). The Self-Consistency Model of Subjective Confidence. *Psychological Review, 119*, 80-113. DOI: 10.1037/a0025648

Koriat, A., Ackerman, R., Lockl, K., & Schneider, W. (2009). The memorizing effort heuristic in judgments of learning: A developmental perspective. *Journal of Experimental Child Psychology, 102*, 265-279. DOI: 10.1016/j.jecp.2008.10.005

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken, & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 483-502). New York, NY: Guilford Press.

Koriat, A., & Shitzer-Reichert, R. (2002). Metacognitive judgments and their accuracy: Insights from the processes underlying judgments of learning in children. In M. Izaute, P. Chambres, P.-J. Marescaux (Eds.), *Metacognition: Process, function, and use* (pp. 1-17). New York: Kluwer.

Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology, 16*, 464-470. DOI: 10.1037/0278-7393.16.3.464

Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition, 29,* 222-233. DOI: 10.3758/bf03194916

McCabe, D. P., & Soderstrom, N. C. (2011). Recollection-based prospective metamemory judgments are more accurate than those based on confidence: Judgments of remembering and knowing (JORKS). *Journal of Experimental Psychology: General, 140*, 605-621. DOI: 10.1037/a0024014

Metcalfe, J., & Finn, B. (2008). Familiarity and retrieval processes in delayed judgments of learning. *Journal of Experimental Psychology, Learning, Memory & Cognition, 34*, 1084-1097. DOI: 10.1037/a0012580

Naveh-Benjamin, M., Brav, T. K., & Levy, O. (2007). The associative memory deficit of older adults: The role of strategy utilization. *Psychology and Aging, 22*, 202-208. DOI: 10.1037/0882-7974.22.1.202

Nelson, T. O. (1984). A comparison of current measures of the accuracy of Feeling-of-Knowing predictions. *Psychological Bulletin, 95*, 109-133. DOI: 10.1037/0033-2909.95.1.109

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "Delayed-JOL Effect". *Psychological Science, 2*, 267-270. DOI: 10.1111/j.1467-9280.1991.tb00147.x

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125-141). New York: Academic Press.

Nelson, T. O., Narens, L., & Dunlosky, J. (2004). A revised methodology for research on metamemory: Pre-judgment recall and monitoring (PRAM). *Psychological Methods, 9*, 53-69. DOI: 10.1037/1082-989X.9.1.53

Nietfeld, J. L., & Schraw, G. (2002). The effect of knowledge and strategy training on monitoring accuracy. *The Journal of Educational Research, 95*(3), 131-142. DOI: 10.1080/00220670209596583

Pressley, M., & Levin, J. R. (1977). Developmental differences in subjects' associative-learning strategies and performance: Assessing a hypothesis. *Journal of Experimental Child Psychology, 24*, 431-439. DOI: 10.1016/0022-0965(77)90089-3

Rhodes, M. G., & Castel, A. D. (2008). Metacognition and part-set cuing: Can inference be predicted at retrieval. *Memory and Cognition, 36*, 1429-1438. DOI: 10.3758/MC.36.8.1429

Richardson, J. T. E. (1998). The availability and effectiveness of reported mediators in associative learning: A historical review and an experimental investigation. *Psychonomic Bulletin & Review, 5*, 597-614. DOI: 10.3758/BF03208837

Robinson, A. E., Hertzog, C., & Dunlosky, J. (2006). Aging, encoding fluency, and metacognitive monitoring. *Aging, Neuropsychology, and Cognition, 13*(3-4), 458-478. DOI: 10.1080/13825580600572983

Robinson, M. D., Johnson, J. T., & Robertson, D. A. (2000). Process versus content in eyewitness metamemory monitoring. *Journal of Experimental Psychology: Applied, 6*(3), 207-221. DOI: 10.1037/1076-898X.6.3.207

Roderer, T., & Roebers, C. M. (2010). Explicit and implicit confidence judgments and developmental differences in metamemory: An eye-tracking approach. *Metacognition and Learning, 5*, 229-250. DOI: 10.1007/s11409-010-9059-z

Roebers, C. M. (2002). Confidence judgments in children's and adults' event recall and suggestibility. *Developmental Psychology, 38,* 1052-1067. DOI: 10.1037/0012-1649.38.6.1052

Roebers, C. M., von der Linden, N., Schneider, W., & Howie, P. (2007). Children's metamemorial judgments in an event recall task. *Journal of Experimental Child Psychology, 97*, 117 - 137. DOI: 10.1016/j.jecp.2006.12.006

Schneider, W. (2010). Metacognition and memory development in childhood and adolescence. In H. S. Waters & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (pp. 54-81). New York: Guilford Press.

Schneider, W. (2015). *Memory development from early childhood through emerging adulthood.* New York: Springer.

Schneider, W., Visé, M., Lockl, K., & Nelson, T. O. (2000). Developmental trends in children's memory monitoring: Evidence from a judgment-of-learning (JOL) task. *Cognitive Development, 15*, 115-134. DOI: 10.1016/S0885-2014(00)00024-1

Shing, Y. L., Werkle-Bergner, M., Li, S.-C., & Lindenberger, U. (2009). Committing memory errors with high confidence: Older adults do but children don't. *Memory, 17*, 169-179. DOI: 10.1080/09658210802190596

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 204-221. DOI: 10.1037/0278-7393.26.1.204

Souchay, C., Bacon, E., & Danion, J. (2006). Metamemory in schizophrenia: An exploration of the feeling-of-knowing state. *Journal of Clinical and Experimental Neuropsychology, 28,* 828-840. DOI: 10.1080/13803390591000846

Toth, J. P., Daniels, K. A., & Solinger, L. A. (2011). What you know can hurt you: Effects of age and prior knowledge on JOL accuracy. *Psychology and Aging, 26*, 919-931. DOI: 10.1037/a0023379

Touron, D. R., Hertzog, C., & Speagle, J. Z. (2010). Subjective learning discounts test type: Evidence from an associative learning and transfer task. Experimental Psychology, 327-337. DOI: 10.1027/1618-3169/a000039

Verhaeghen, P., Marcoen, A., & Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: A meta-analytic study. *Psychology and Aging, 7*, 242-251. DOI: 10.1037/0882-7974.7.2.242

Weinert, F. E., & Schneider, W. (1996). Entwicklung des Gedächtnisses . In D. Albert & K.-H. Stapf (Hrsg.), *Enzyklopädie der Psychologie, Themenbereich C, Serie II, Band 4* (S. 433-487). Göttingen: Hogrefe.

Willoughby, T., Porter, L., Belsito, L., & Yearsley, T. (1999). Use of elaboration strategies by students in grades two, four and six. *The Elementary School Journal, 99*(3), 221-231.

Wong, J. T., Cramer, S. J, & Gallo, D. A. (2012). Age-related reduction of the confidence-accuracy relationship in episodic memory: effects of recollection quality and retrieval monitoring. *Psychology and Aging, 27*(4), 1053-1065. DOI: 10.1037/a0027686

Zakay, D., & Tuvia, R. (1998). Choice latency times as determinants of post-decisional confidence. *Acta Psychologica, 98*, 103-115. DOI: 10.1016/S0001-6918(97)00037-1