British Journal of Developmental Psychology (2010), 28, 767–784
© 2010 The British Psychological Society
The British Psychological Society
www.bpsjournals.co.uk

Self-paced study time as a cue for recall predictions across school age

Anja Hoffmann-Biencourt[1]*, Kathrin Lockl[2], Wolfgang Schneider[1], Rakefet Ackerman[3] and Asher Koriat[3]
[1]University of Wuerzburg, Germany
[2]University of Bamberg, Germany
[3]University of Haifa, Israel

Recent work on metacognition indicates that monitoring is sometimes based itself on the feedback from control operations. Evidence for this pattern has not only been shown in adults but also in elementary schoolchildren. To explore whether this finding can be generalized to a wide range of age groups, 160 participants from first to eighth grade participated in a study based on a self-paced study time (ST) allocation paradigm. In contrast to previous studies, picture pairs instead of word pairs were used as stimuli to compensate for reduced reading skills in younger participants. Actual ST and judgments of learning (JOLs) made at the end of each study trial were used as core variables. The results are in line with previous findings, in that children's JOLs decreased with increasing ST, suggesting that JOLs were based on the memorizing effort heuristic that easily learned items are more likely to be remembered. Weaker inverse relationship between JOLs and ST was found for the younger children. Overall, these results underline the importance of mnemonic cues in shaping metacognitive feelings not only in adults but also in older children and expose a developmental trend in their use along childhood.

The investigation of metacognitive processes has been carried out by two different lines of research, an experimental-cognitive line and a developmental line (Koriat & Shitzer-Reichert, 2002). Whereas researchers in the field of experimental-cognitive psychology have paid special attention to several basic issues in relation to metacognitive monitoring and control processes, researchers in developmental psychology initially focused on children's declarative metacognitive knowledge about memory strategies, tasks demands, and personal characteristics (Schneider & Bjorklund, 1998, 2003). It was only in recent years that developmental psychologists began to investigate age-related progression in procedural metacognitive competencies, that is children's ability to monitor and control their learning and memory activities

*Correspondence should be addressed to Anja Hoffmann-Biencourt, University of Wuerzburg, Roentgenring 10, 97070 Wuerzburg, Germany (e-mail: a.hoffmann@psychologie.uni-wuerzburg.de).

DOI:10.1348/026151009X479042

(see Schneider & Lockl, 2008, for a recent review). Today, the concept of metacognition is not anymore limited to the field of developmental psychology. Because of its relevance for a large number of concepts related to thinking, it is also used in the field of clinical or educational psychology or even cognitive neuroscience (cf. Schneider, 2010).

Recent models of metacognition also add relevant components such as self-regulation skills (e.g., Efklides, 2001). Efklides (2001, 2008) distinguishes metacognitive knowledge, i.e., knowledge which is retrieved from memory, metacognitive experiences, i.e., a person's experiences during a cognitive endeavour (e.g., on-line task-specific knowledge, feelings, goals, judgments) and metacognitive skills, i.e., strategies like planning and evaluating. Metacognitive knowledge and metacognitive experiences are both aspects, which are supposed to be manifestations of the monitoring function, while metacognitive skills are involved in strategy use for the control of cognition. The control function is reflected by self-regulatory processes. While on-line task-specific knowledge is characterized as conscious and analytic, the other metacognitive experiences are supposed to be products of non-conscious, non-analytic inferential processes (Efklides, 2006). On this basis, metacognitive experiences act as a trigger for rapid, non-conscious control decisions but also conscious analytic ones. In the same way, metacognitive experiences can make use of both the affective and the cognitive regulatory loops.

An assumption that is held by researchers in both fields, i.e., experimental cognitive as well as developmental psychology, is that monitoring processes are important because monitoring guides control (Nelson & Narens, 1990). For instance, monitoring processes are supposed to play a central role in directing how people study. This view has been supported by numerous studies including mainly adult participants that showed that individuals use memory monitoring, especially judgments of learning (JOLs), to decide which items to study, and how long to spend on them (e.g., Metcalfe, 2002; Nelson, Dunlosky, Graf, & Narens, 1994; Nelson & Narens, 1990; Son & Metcalfe, 2000). Concerning the accuracy of monitoring processes, the adult literature demonstrated that the degree of correspondence between JOLs and recall performance is moderate, and numerous factors were shown to influence this correspondence (Thiede & Anderson, 2003). For example, the delay between studying the items and making JOLs proved to be critical, with longer delays leading to increased JOL accuracy ('the delayed JOL-effect', Dunlosky & Nelson, 1992).

Regarding monitoring processes in a developmental context, only few studies examined children's performance in JOL tasks. For instance, Schneider, Visé, Lockl, and Nelson (2000) reported that 6- to 10-year-old children's JOL accuracy improved considerably when JOLs were elicited at some delay rather than immediately after study, thus replicating research with adults. In a 'delayed' condition even young children were able to effectively predict their memory performance, and no significant age trends in JOL accuracy emerged. In other studies, elementary schoolchildren were shown to differentiate between related (high associative) and unrelated (low associative) item pairs by assigning higher JOLs for related item pairs than for unrelated item pairs (Koriat & Shitzer-Reichert, 2002; Lockl & Schneider, 2003). With regard to the correspondence between JOLs and actual memory performance, the results reported by Koriat and Shitzer-Reichert (2002) were not entirely consistent with those reported by Schneider *et al.* (2000) as only the former found an age-related increase in monitoring accuracy. Roebers, von der Linden, Howie, and Schneider (2007) examined 8-, 10-year-olds' and adults' monitoring abilities in the context of a complex, everyday memory task and reported that even children in the youngest age group were able to appropriately differentiate in their JOLs between correct and incorrect answers as well as between answerable and unanswerable questions (about which no information was available). JOL accuracy was comparable across age groups, with moderate to high Gamma correlations between JOLs and recall performance. Taken together, recent studies assessing monitoring abilities in JOL tasks yielded inconsistent results. Whereas some studies demonstrated age-related increase in children's monitoring skills others did not.

Looking at the interplay between monitoring and control processes, there is more consistent evidence suggesting a developmental progression across the elementary school years. In one of the earliest studies, Masur, McIntyre, and Flavell (1973) showed that 9-year-olds and college students tended to select items for additional study that were not recalled correctly on a previous study trial, whereas 7-year-olds did not seem to consider previous-trial performance in selecting items for additional processing. Under self-paced learning instructions, only children from the age of about 10 years on spent more time studying unrelated paired associates than they did on related paired associates (Dufresne & Kobasigawa, 1989). Younger children spent approximately the same amount of time on related pairs as they spent on unrelated pairs. Although many of the younger children in the Dufresne and Kobasigawa (1989) study were able to distinguish between unrelated and related paired associates, they apparently found it difficult to translate this metacognitive knowledge into adequate self-regulation (for similar findings, see Lockl & Schneider, 2002a). To investigate the relation between monitoring processes and self-regulation processes more precisely, Lockl and Schneider (2003) recorded 7- and 9-year-old's JOLs after a first short learning phase and presented the same items again for self-paced study. Although children of both age groups studied items with lower JOLs for a longer period of time than item pairs with higher JOLs, the relation between monitoring and control was significantly stronger for 9-year-olds than for 7-year-olds. Furthermore, recall performance was significantly better for children with high self-regulation (in terms of JOL–study time (ST) correlation) than for those with low self-regulation, regardless of age. In a recent study, Son (2005) found that first graders' JOLs did not influence their subsequent study strategy, that is, their decision to study the item immediately or after a delay. Only when the task was changed and children could decide to either read or to generate an item in the short term, were these decisions guided by their metacognitive judgments. Overall, the available evidence suggests that there are clear age-related increases in the ability to use the output from monitoring for further self-regulated study behaviour. Depending on the difficulty or complexity of the task, adequate self-regulation may be observed in early elementary schoolchildren (Kobasigawa & Metcalf-Haggert, 1993) or in secondary schoolchildren (Brown, Smiley, & Lawton, 1978).

The underlying assumption of the studies mentioned above is that self-regulated learning behaviour is based on monitoring processes. In this sense, item difficulty is monitored before actually investing ST. The outcome of monitoring processes then serves as a basis for subsequent ST allocation. Contrary to this view, Koriat, Ma'ayan, and Nussinson (2006) were the first to demonstrate that the sequence of monitoring and control may also be reversed. Koriat *et al.* (2006) investigated what they labelled the 'data-driven function' of ST. According to this approach, learners use self-paced ST as a basis for subsequent JOLs. Whereas in the first case increased effort is associated with higher JOLs and better recall, the latter case is characterized by greater effort being associated with lower JOLs and lower recall.

So far, the recently advanced 'control–affects–monitoring' hypothesis has been predominantly studied in adult participants (Koriat & Ma'ayan, 2005; Koriat *et al.*, 2006).

It was shown that both JOLs and recall decreased with self-paced ST, whereas JOLs and recall generally increased with experimenter-determined presentation duration. With regard to immediate JOLs, the results suggest that learners use study effort as measured by self-paced ST as a cue for JOLs, relying on the so-called memorizing effort heuristic. This heuristic implies that the more effort is invested in studying an item, the less likely it is to be recalled. Thus with regard to self-paced ST, it is the invested amount of effort which affects subsequent JOLs. The described data-driven character of ST is characterized as an unconsciously applied cue. This is in line with the cue-utilization view (Koriat, 1997), according to which JOLs are based on intrinsic, extrinsic, and mnemonic cues which are predictive of recall. The latter category refers to internal, subjective indicators, signalling the extent to which an item has been mastered. The claim that metacognitive feelings and judgments are products of non-conscious, heuristic processes (Koriat & Levy-Sadot, 1999) is also supported by recent conceptualizations of metacognition (e.g., Efklides, 2001; Schunk & Zimmerman, 1998).

The first study which examined the memorizing effort heuristic as a basis for JOLs from its developmental aspect was conducted by Koriat, Ackerman, Lockl, and Schneider (2009b). They had children from first to third grade and from fifth to sixth grade study a list of word pairs under self-paced instructions, and observed a developmental shift in the development of the memorizing effort heuristic. While first and second graders did not differentiate between short- and long-studied items in their JOLs, the results of the older age groups disclosed the heuristic that longer self-paced STs resulted in lower recall predictions. The negative correlation between ST and JOLs reflecting the cue-utilization aspect could also be shown for cue validity, which is expressed by a negative correlation between ST and recall performance. Remarkably, recall decreased with ST in both age groups under study. This finding strengthens the validity of the memorizing effort heuristic. Furthermore, children from all grades showed evidence for accurate monitoring accuracy, which is reflected by the positive relationship between JOLs and recall performance.

However, a limitation of the study by Koriat *et al.* (2009b) that constrains the interpretation of the results might be that only word pairs were used as stimuli. Younger children, especially first graders who just have began to learn reading might have taken more time for decoding the word pairs than older children and this may have confounded the variables reading time and ST. Furthermore, there is evidence that there are age differences concerning the way in which children and adults process pictures and words (e.g., Ackerman, 1981; Cramer, 1972, 1973; Means & Rohwer, 1976). B. Ackerman (1981) suggested that young children typically encode stimuli in a fashion that stresses the sensory aspects of the stimuli and that there are age differences in the efficiency with which the semantic information in stimuli is processed. Thus, using word pairs may be especially demanding for younger children because learning word pairs to a great extent requires a focus on the semantic information.

In order not to place younger children at a disadvantage relative to older participants and in order to provide a common ground across age groups, in the present study, pairs of pictures were used as stimuli. In doing so, we adopted the procedure developed by Koriat *et al.* (2009b) to test the 'control–affects–monitoring' hypothesis. According to the memorizing effort heuristic, learners should base their JOLs on the previously invested ST, i.e., study or memorizing effort. In this way, ST acts as an implicit mnemonic cue (Koriat, 1997). Based on the results by Koriat *et al.* (2009b), we hypothesized that inverse ST–JOL and ST–recall correlations should emerge during primary school age, reflecting reliance on the memorizing effort heuristic. This correlation should consolidate with increasing age. Moreover, as suggested by previous work (e.g., Koriat & Shitzer-Reichert, 2002; Schneider *et al.*, 2000), we expected a positive correlation between JOLs and recall. The mentioned relationships should also hold true when controlling for a possible mediating effect of item difficulty. A particular aim of the present study was to explore cue utilization, cue validity, and monitoring accuracy under a developmental perspective. Given that the study by Koriat *et al.* (2009b) so far is the only one to suggest developmental trends in children's reliance on memorizing effort as a basis for their JOLs, further work seems important. Compared to the study by Koriat *et al.* (2009b), we used a more comprehensive approach to assess developmental trends by recruiting children from Grade 1 to Grade 8 (7- to 14-year-olds). This extension over eight grades is particularly important for examining the later stage of the developmental trend (Koriat *et al.*, 2009b) and for a comparison of results found with young adults (Koriat *et al.*, 2006).

Method

Overview

The study involved a 4 (age groups) × 2 (item difficulty) design with item difficulty (highly related item-pairs versus unrelated pairs) as a within-subject factor. Self-paced ST, JOLs and recall for each item pair were recorded as dependent variables.

Participants

One hundred and sixty children from first to eighth grade participated in the study, with 20 children in each grade. Whereas the first–fourth grade participants attended primary school, the fifth–eighth grade participants attended secondary school. Participants were mostly of middle-class or upper middle-class socio-economic background. In order to facilitate the presentation of the following analyses, two grade levels were combined into one age group. Following this procedure, we obtained four age groups with 40 participants in each of them: Age Group 1 (Grades 1 and 2) with a mean age of 7.5 years ($SD = 0.7$ months), Age Group 2 (Grades 3 and 4) with a mean age of 9.3 years ($SD = 0.6$ months), Age Group 3 (Grades 5 and 6) with a mean age of 11.5 years ($SD = 0.8$ months), and Age Group 4 (Grades 7 and 8) with a mean age of 13.7 years ($SD = 0.7$ months). In each age group, an approximately equal number of boys and girls were included. Overall, 89 girls and 71 boys participated in the study. Children were recruited from different schools of a city situated in southern Germany. All schools were situated in the same district with predominantly middle-class socio-economic background. All pupils of randomly selected classes and their parents received a short information letter with an invitation to participate in a study exploring children's learning behaviour. Parents' consent was given for the majority of the children in each class. These children were tested individually for about 30 min.

Materials and procedure

The procedure was adopted from a recent study by Koriat *et al.* (2009b), who used word pairs of different difficulty levels as stimulus materials. Related word pairs were semantically or associatively related, while in hard pairs the two members of each pair were unrelated. In the present study, the word pairs were replaced by pairs of pictures, which had been developed by Lockl and Schneider (2002a, 2003). Each pair consisted of two coloured line drawings, presented on a computer screen (1.5 GHz PC-compatible laptop, screen resolution: $1,280 \times 1,024$ pixels). All pictures (3×4 cm) represented familiar objects well-known to all children in the sample (see Appendix). They were presented in the middle of the screen and in one of four orders, counterbalanced across participants. The final item set of 26 pairs of pictures (2 practice items and 24 test items) was selected on the basis of a norming study from an existing item pool. In the corresponding norming study, a group of second ($N = 27$) and fourth graders ($N = 24$) had been asked to indicate how many people (out of 100) would be able to recall each of the shown pairs of pictures. All of the finally selected pairs of pictures represented different degrees of memorability ratings. Each of the related pairs had higher memorability ratings than each of the unrelated pairs of pictures. In the following, we will refer to the difficulty ratings which were obtained in the norming study as normative item difficulty. The use of this kind of stimulus material allowed us to control for influences of reading skills, especially in the younger children.

Participants were told that they would be shown 24 pairs of pictures included in the following tasks. They were told to study each pair as long as necessary in order to be able to recall the response picture, when presented with the stimulus picture in a later test phase. The picture pairs appeared one after another on the computer screen. On each trial, the participants pressed a button when ready to study the next picture pair. The time between the presentation of a pair for study and the following button press was recorded automatically, indicating the ST for that item. Following each study trial, the relevant picture pair disappeared and the child was immediately asked to indicate the likelihood of recalling the target in the later cued-recall test. Each participant indicated his or her JOLs by sliding a pointer, which was initially placed in the middle of the scale, on a coloured thermometer, with deep blue representing 'very cold' and deep red representing 'very hot' (see Koriat & Shitzer-Reichert, 2002; Koriat *et al.*, 2009b). The procedure was explained in detail by referring to a familiar children's game and using the two practice items mentioned above. Even the younger participants knew the cold–hot game very well. Finally, the children's judgment was transformed into a JOL percentage score ranging from 0 to 100% (i.e., from 'very cold' to 'very hot'). After having studied one pair and having indicated the corresponding JOL, the child could continue to the next study trial by navigating with a mouse. Once the study phase was completed, each child made a free line drawing during 1 min. In the following test phase, the cue pictures were presented again in a random order. Responses were given orally and the answers were entered by the experimenter on a keyboard. The procedure was explained by using the two practice items, which had already been employed in the study phase. If a child was not able to give an answer, it was allowed to continue with the next cue picture.

Results

Preliminary analyses examining the effect of gender on the dependent measures showed no systematic differences between male and female participants. Consequently, all data were collapsed across gender.

In the following, we first present results concerning the three dependent variables, that is, ST, JOLs and recall, as a function of age group (between-person factor) and item difficulty (within-person factor). All corresponding means are reported in Table 1.

Table 1. Mean STs, JOLs, and recall as a function of age group and item difficulty (*SD* in parentheses)

		Age group				
		1 (first and second graders)	2 (third and fourth graders)	3 (fifth and sixth graders)	4 (seventh and eighth graders)	All age groups
STs (s)	Related	6.9 (3.5)	8.0 (4.8)	7.2 (4.4)	9.5 (7.4)	7.9 (5.3)
	Unrelated	7.3 (4.2)	8.7 (5.4)	9.6 (6.5)	12.9 (8.9)	9.6 (6.7)
	Overall	7.1 (3.7)	8.4 (5.0)	8.4 (5.3)	11.2 (8.0)	8.8 (5.8)
JOLs (%)	Related	76.7 (13.7)	67.3 (13.8)	70.8 (10.1)	73.8 (10.3)	72.1 (12.5)
	Unrelated	64.4 (21.9)	57.5 (13.9)	46.2 (13.9)	47.1 (12.5)	53.8 (17.5)
	Overall	70.5 (16.9)	62.4 (12.5)	58.5 (9.2)	60.4 (9.5)	62.9 (13.1)
Recall (%)	Related	75.6 (17.5)	82.5 (15.7)	89.4 (12.9)	91.9 (8.5)	84.8 (15.3)
	Unrelated	21.7 (15.7)	31.3 (20.9)	45.8 (27.3)	54.0 (26.5)	38.2 (26.1)
	Overall	48.6 (15.0)	56.9 (16.1)	67.4 (17.7)	72.9 (15.2)	61.5 (18.5)

Then we report results concerning the effect of ST on individual JOLs and recall with corresponding within-person correlations between ST and JOLs (cue utilization), ST and recall (cue validity), and JOLs and recall (monitoring accuracy). If not specified otherwise, the level of significance was set to $p < .05$. In order to determine whether the observed statistically significant differences are meaningful, we also report the corresponding standardized effect sizes (η^2).

Study time

An analysis of variance with ST as a dependent variable revealed a main effect of Age Group on individual mean ST, $F(3, 156) = 3.67$, $p < .05$, $\eta^2 = .07$. Subsequent *post hoc* tests (Scheffé) showed a significant difference in ST between Age Groups 1 and 4 with seventh and eighth graders having significantly longer overall STs compared to first and second graders ($p < .01$). Furthermore, there was a main effect for item difficulty, $F(1, 156) = 54.76$, $p < .001$, $\eta^2 = .26$. Participants allocated more time to hard items than they allocated to easy items. This main effect of item difficulty was modified by a significant age group × item difficulty interaction. Simple *t* tests indicated significant differences between the ST for related and unrelated pairs for the older Age Groups (Age Group 2: $p < .01$, Age Group 3 and 4: all p's $< .001$) but not for the youngest Age Group. Whereas first and second graders spent about the same time studying the hard pairs as they spent studying the easy pairs, older children spent more time studying the hard pairs.

Judgments of learning

An analysis of variance with JOL as a dependent variable yielded a main effect of Age Group (4) on individual mean overall JOLs, $F(3, 156) = 7.31$, $p < .001$, $\eta^2 = .12$, with Age Group 1 showing significantly higher JOLs than Age Groups 2, 3, and 4 (all p's $< .05$). Furthermore, the analysis revealed a main effect of item difficulty with $F(1, 156) = 284.38$, $p < .001$, $\eta^2 = .65$, and a significant interaction, $F(3, 156) = 15.36$, $p < .001$, $\eta^2 = .23$. Although children in all age groups gave higher JOLs to related than to unrelated picture pairs the difference in JOLs between the two types of pairs was smaller for the younger than for the older children.

Recall

The corresponding analysis with recall as a dependent variable revealed a main effect of Age Group, $F(3, 156) = 18.13$, $p < .001$, $\eta^2 = .26$. Subsequent t tests revealed that Age Group 4 showed significantly higher overall recall than Age Groups 1 and 2 (all p's $< .001$). The same pattern emerged for Age Group 3 with regard to Age Groups 1 and 2 (all p's $< .001$). Furthermore, there was a main effect of item difficulty, $F(1, 156) = 832.75$, $p < .001$, $\eta^2 = .84$, and a significant interaction, $F(3, 156) = 5.12$, $p < .01$, $\eta^2 = .09$. The interaction reflects the observation that the effect of item difficulty on recall decreased with age. Simple t tests indicated significantly higher recall for related items in all Age Groups (all p's $< .001$). Corresponding effect sizes (Cohen's d) were .35 (Age Group 1), .41 (Age Group 2), .73 (Age Group 3), and .82 (Age Group 4).

Cue utilization: Memorizing effort as a cue for JOLs

To examine the participants' reliance on memorizing effort as a basis for subsequent prediction of recall, we adopted a procedure used by Koriat *et al.* (2006, 2009b), reflecting the within-person relationship between the variables mentioned before. After splitting all STs at the median for each participant, mean JOLs for below-median and above-median STs were calculated for each participant (see Figure 1a). An Age Group × Study Time (below- vs. above-median ST) analysis of variance with ST as an within-subject factor yielded a significant main effect of Age Group, $F(3, 156) = 7.18$, $p < .001$, $\eta^2 = .12$, a significant main effect of ST, $F(1, 156) = 111.08$, $p < .001$, $\eta^2 = .42$, and a significant interaction, $F(3, 156) = 7.61$, $p < .001$, $\eta^2 = .13$.

An inspection of Figure 1a reveals that JOLs were higher for items with below-median STs than for items with above-median STs. The interaction between ST and Age Group is due to the fact that this difference in JOLs between below-median STs and above-median STs was more pronounced for older students than for younger students. Simple t tests indicated significantly higher JOLs for items with short ST in comparison to items with long ST across all Age Groups (Age Group 1: $p < .01$, Age Groups 2, 3, and 4: all p's $< .001$).

The results described so far were substantiated by within-participant gamma correlations which were calculated for the variables ST and JOLs (see Table 2). Within-participant gamma correlations for Age Groups 1, 2, 3, and 4 averaged $-.11$, $-.13$, $-.28$, and $-.34$, respectively. Each of these correlations was significantly different from zero ($p < .01$ for Age Group 1; $p < .001$ for Age Groups 2 to 4). A one-way analysis of variance for Age Group 4 on the within-person gamma correlations yielded a significant effect of Age Group, $F(3, 156) = 9.02$, $p < .001$, $\eta^2 = .15$. *Post hoc* tests showed that ST – JOL gamma correlations for Age Groups 1 and 2 were significantly lower than those obtained for the two older age groups ($p < .05$). None of the other pair differences reached significance.

As ST is strongly correlated with normative item difficulty (Son & Metcalfe, 2000), the observed ST–JOL correlation is also consistent with the dominant view in research on metacognition, which stresses the link from monitoring (JOL) to control (ST). In order to explore the extent to which the ST – JOL relationship is mediated by normative item difficulty, within-person Pearson correlations were computed with normative item difficulty partialled out (Table 2). The weak ST – JOL relationship disappeared in both younger age groups when item difficulty was partialled out. In the older age groups, a significant relationship between ST and JOLs remained when controlling for item difficulty.

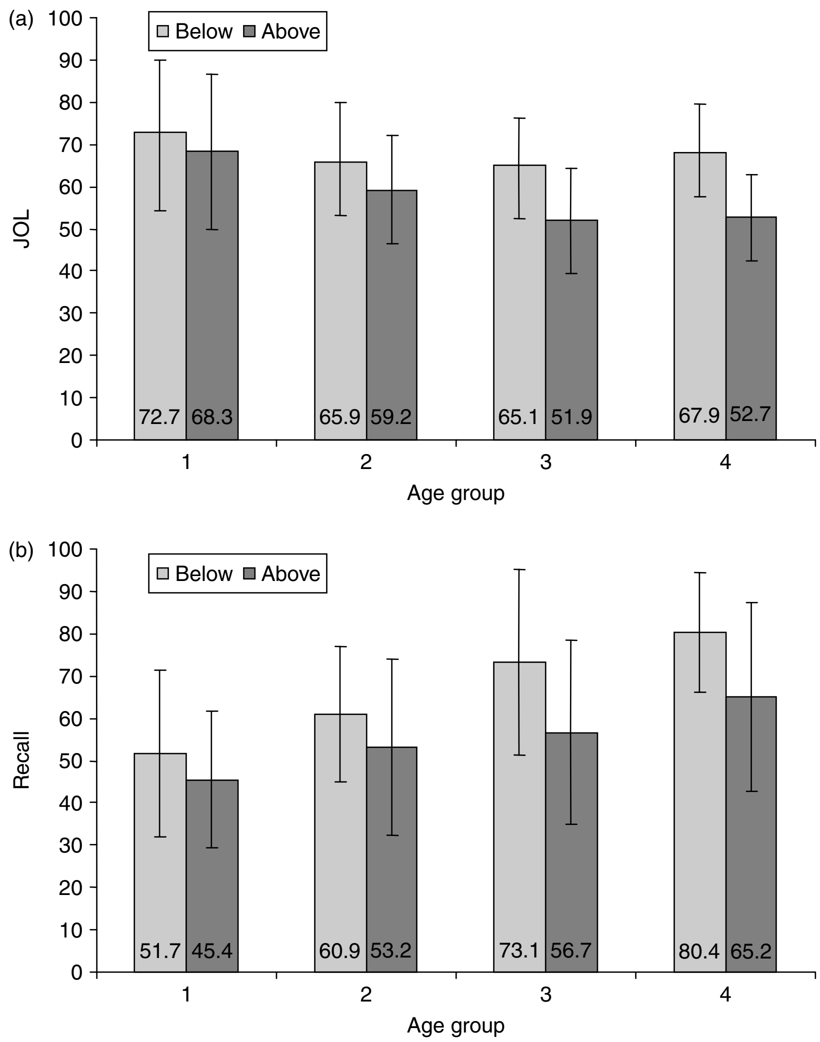

Figure 1. Mean JOL (a) and mean recall (b) for below-median and above-median ST for each of the four age groups (error bars = *SD*).

Cue validity: The validity of ST as a predictor of recall

The same procedure as described above was used to calculate recall for items with below-median and above-median STs in order to examine the within-person relationship between ST and recall (see Figure 1b). An Age Group × Study Time (below vs. above median ST) analysis of variance with ST as an within-subject factor yielded a significant main effect of Age Group, $F(3, 156) = 15.51$, $p < .001$, $\eta^2 = .23$, a significant main effect of ST, $F(1, 156) = 52.07$, $p < .001$, $\eta^2 = .25$, and a significant interaction, $F(3, 156) = 2.69$, $p < .05$, $\eta^2 = .05$. Overall, recall was higher for items with below-median STs and older participants recalled more items than younger participants. Independent samples *t* tests indicated significantly higher recall for items with short ST in comparison to items with long ST for the older Age Groups (Age Group 2: $p < .05$, Age Group 3 and 4: all p's $< .001$), but not for Age Group 1.

Table 2. Mean within-person ST–JOL gamma, Pearson and partial (controlling for normative item difficulty) correlations as a function of age group

	Age group				
	1 (first and second graders)	2 (third and fourth graders)	3 (fifth and sixth graders)	4 (seventh and eighth graders)	All age groups
Gamma	−.11**	−.13***	−.28***	−.34***	−.22***
Pearson	−.13*	−.17***	−.29***	−.38***	−.25***
Pearson/partial	−.09	−.07	−.15**	−.19***	−.12***

*$p < .05$; **$p < .01$; ***$p < .001$.

Furthermore, within-person gamma correlations between ST and recall were calculated for all Age Groups. The mean correlations for Age Groups 1, 2, 3, and 4 averaged −.10, −.15, −.32, and −.29, respectively. All correlations were significantly different from zero ($p < .05$ for Age Group 1; $p < .001$ for Age Group 2–4). A one-way analysis of variance for Age Group on the reported mean gamma correlations yielded a significant effect, $F(3, 154) = 5.89$, $p < .01$, $\eta^2 = .10$. The youngest age group showed significantly lower within-person gamma correlations than Age Groups 3 and 4 (all p's $< .001$). Also, within-gamma correlations obtained for Age Group 2 were significantly lower than those found for the older Age Groups (all p's $< .05$).

Relations between achievement and the accuracy of JOLs in predicting recall

Within-person gamma correlations for JOLs and recall were computed as an indicator for achievement, that is, the extent to which JOLs are valid predictors of subsequent recall. These correlations averaged .36, .34, .64, and .62 for Age Groups 1, 2, 3, and 4, respectively. All correlations were significantly different from zero (all p's $< .001$). A one-way analysis of variance for Age Group on the mean ST-recall gamma correlations yielded a significant main effect, $F(3, 154) = 10.82$, $p < .001$, $\eta^2 = .17$. *Post hoc* tests (Scheffé) revealed that the correlations for the younger Age Groups 1 and 2 were significantly lower than those found for the older Age Groups (all p's $< .01$).

Discussion

Although there has been a great deal of work in the last two decades on the on-line metacognitive processes of adults that occur during learning (see Koriat, 2007, for a review), only a few studies have explored the basis and accuracy of JOLs in children (see Schneider & Lockl, 2008, for a review). Even these few developmental studies have focused primarily on calibration, rather than on resolution. That is, they investigated the extent to which children's metacognitive judgments correspond on average to their memory performance (see Schneider & Lockl, 2008). The results of these studies have generally indicated that preschoolers tend to overestimate their future memory performance, whereas schoolchildren's predictions tend to be more realistic (see Schneider & Pressley, 1997). Consistent with this finding and the results obtained by Koriat *et al.* (2009b), in the present study JOL levels decreased with age, while recall levels increased with age. Moreover, the present study focused on the internal dynamics that are revealed by within-person correlations between JOL, ST, and recall, borrowing from experimental paradigms that have been used extensively in the study of adult metacognition. Using these paradigms to investigate developmental trends has already yielded important insights (see Schneider & Lockl, 2008).

The work on JOLs in children and adults has generally followed Nelson and Narens' (1990) distinction between two components – monitoring and control. Overall, several developmental results as well as those obtained with adults are in line with the assumption that monitoring affects control (Nelson & Leonesio, 1988; Son & Schwartz, 2002) which is claimed to apply to a variety of metacognitive judgments (Barnes, Nelson, Dunlosky, Mazzoni, & Narens, 1999; Nelson & Narens, 1990; Son & Schwartz, 2002), implying that metacognitive feelings are not mere epiphenomena but actually exert measurable effects on the regulation of learning and remembering (Nelson, 1996; Son & Schwartz, 2002). It is this assumption that has motivated in part the increased interest in metacognitive processes in children and adults (Koriat & Goldsmith, 1996; Nelson, 1996; Schneider & Lockl, 2008; Schneider & Pressley, 1997).

A control–affects–monitoring model has been put forward recently (Koriat, 2006; Koriat *et al.*, 2006). Consistent with this model, research with young adults (Koriat *et al.*, 2006) indicated that JOLs made at the end of a study trial decrease as the amount of self-paced ST increase. These finding led Koriat and his colleagues to conclude that although metacognitive monitoring often guides control operations, sometimes it follows control operations and is based on the feedback from them. Consistent with this view, in self-paced learning, JOLs made at the end of each study trial often *decrease* with the amount of time spent studying the item, suggesting that JOLs are based on the memorizing effort heuristic that easily learned items are more likely to be remembered. As noted above, this finding first established with adults was also confirmed in a recent developmental study carried out by Koriat *et al.* (2009b).

The main goal of the present study was to investigate the control–affects–monitoring hypothesis in children, and to test the validity of the findings by Koriat *et al.* (2009b) by using different study material, that is, picture pairs instead of word pairs. Furthermore, the sample included children from Grades 1 to 8 to explore the critical age period for the shift to occur and to bridge the gap between younger children on the one hand (Koriat *et al.*, 2009b) and young adults on the other hand (Koriat *et al.*, 2006).

Let us first review the findings concerning the effects of intrinsic item difficulty on ST, JOLs, and recall. As far as the distinction between related and unrelated paired associates is concerned, we found that children in all grades gave higher JOLs to the related than to the unrelated items, and in parallel, recall was also better for the easier pairs than for the harder pairs. These results replicate previous findings with children and adults. Unlike the pattern obtained with adults, however, the magnitude of the effect of item difficulty was much larger for recall than it was for JOLs. The failure of children to appreciate the effects of item difficulty on recall derived from their tendency to overestimate the recall of the unrelated pairs. This overestimation may be related to the general tendency of children to overestimate their memory performance (see Schneider, 1998) but it has to be clarified why that tendency was specifically observed for the hard pairs.

Of particular interest is the finding that the older children spent significantly more time studying the unrelated than the related pairs whereas the younger children's STs did not differentiate significantly between the two types of pairs. These results are consistent with previous findings (see Dufresne & Kobasigawa, 1989; Lockl & Schneider, 2002b, 2004). Dufresne and Kobasigawa observed that even the youngest children in their sample seemed to distinguish between related and unrelated items although this distinction was not reflected in their ST allocation. In fact, the same pattern emerges in our study, as can be seen by comparing the results for JOLs with those for ST (Table 1). The younger children did give significantly lower JOLs to the harder items. If the regulation of ST is data driven, then the age-related sensitivity of ST to intrinsic item difficulty would seem to imply that there is a developmental change in the very regulation of ST and not only in the application of metacognitive knowledge to the spontaneous regulation of ST. Older children's ST allocation would seem to be attuned to features of the studied materials (including differences in intrinsic item difficulty) that are critical for learning and remembering (see further below).

A second interesting finding concerns the developmental trend observed for monitoring accuracy as indexed by the JOL–recall correlation. The data obtained for the four age groups (first to eighth grade) provide a detailed insight into the development of monitoring accuracy. Although the younger age groups already showed a moderate correlation between JOLs and recall, this relationship is even more pronounced in the older age groups. This finding is consistent with the results reported by Koriat and Shitzer-Reichert (2002) and Koriat *et al.* (2009b). However, as mentioned above, there are studies that did not find an age-related increase in JOL accuracy (Roebers *et al.*, 2007; Schneider *et al.*, 2000). This discrepancy may be explained by procedural differences. Studies that showed an age-related trend in monitoring accuracy, including the present study, used immediate JOLs whereas studies that did not find such an age trend used (predominantly) delayed JOLs. Accordingly, a tentative explanation could be that immediate JOLs are particularly demanding for younger children. It can be assumed that immediately after study, information about the to-be-judged item is based on short-term memory (cf. Nelson & Dunlosky, 1991). If the information about the item at the time of the JOL is still present, young children may have more difficulties – compared to older children and adults – to predict that some forgetting will occur, and that the forgetting will be especially pronounced for harder items. On the other hand, when JOLs are delayed until the to-be-judged item is not anymore available in short-term memory, even young children may acknowledge that information that is not present at the moment probably will not come to mind at a later point in time. Clearly, more research is needed to resolve this issue.

We turn next to the main findings of interest for the 'control affects monitoring' model of memory monitoring. The first important finding is that older children show clear evidence for reliance on the memorizing effort heuristic in making JOLs. Thus, the present data support the findings by Koriat *et al.* (2009b) and provide an important extension of this study with regard to the study materials and the age groups under study. The present study showed that the recall predictions of fifth- to eighth-graders consistently confirm the implicit heuristic that the more time they spend studying an item the *less* likely they are to recall it in the future. Although use of this heuristic was also observed for both of the younger age groups (1 and 2), the association between JOLs and STs is somewhat weaker. The shift in cue utilization in this study occurs at about the same age as in the study by Koriat *et al.* (2009b), that is, during the middle of the elementary school years. Although the actual STs and recall scores differ somewhat between both studies (e.g., first graders' STs were longer in the study by Koriat *et al.* (2009b), overall recall was higher in the present study) the developmental pattern regarding cue utilization is very similar. Thus, the fact that in both studies different study materials were used (word vs. picture pairs) strengthens the validity and generalizability of the age-related trend concerning cue utilization.

Looking at the oldest age group (seventh and eighth graders) in the present study, the gamma correlation between ST and JOLs is $-.34$ and comes relatively close to the gamma correlation found for college students ($-.48$) in the study by Koriat *et al.* (2006). Despite this similarity, however, there still seems to be a developmental progression in cue utilization from eighth grade to the college years. The evidence found for the older age groups is in line with the proposition that JOLs are based on mnemonic cues that derive from task performance rather than on explicit metacognitive knowledge, because the actual correlation between JOLs and ST was in the opposite direction from that reflected in children's self-reports concerning the interrelationship between ST and expected recall (Koriat *et al.*, 2009b). Furthermore, the recall performance of these children also supports the validity of the memorizing effort heuristic. Longer self-paced STs were predictive of poorer recall performance. The correspondence between cue utilization and cue validity that was observed possibly contributed to the accuracy of the JOLs in predicting inter-item differences in recall performance.

Thus, these results present evidence for the data-driven function of ST. That means that it is the studied item or, more precisely, the learner–item interaction which determines the invested amount of ST in the way that learners spend as much ST as apparently necessary to master an item. Following the cue-utilization view (Koriat, 1997), self-paced ST for an item is then used as a basis for subsequent JOLs. This information, signalling the extent to which an item has been mastered, serves as an unconsciously applied cue, with little memorizing effort being associated with higher JOLs and better recall as well as greater effort being associated with lower JOLs and lower recall. However, it has to be acknowledged that it is not entirely clear whether it is really perceived memorizing effort or some other variable that covaries with ST (e.g., satisfaction with the attempt to understand the relation between the two pictures) that accounts for the impact of ST on JOL.

An alternative interpretation for the correlational pattern between ST and JOLs could be that it is item difficulty that affects both ST and recall predictions. To take this possibility into account, we computed partial correlations between ST and JOLs. It could be shown that in the younger age groups the correlations between ST and JOLs failed to reach significance when controlling for item difficulty. In the older age groups, however, the correlations between ST and JOLs remained significant after controlling for the effect of normative item difficulty. As already suggested by Koriat *et al.* (2009b; Koriat, Ackerman, Lockl, & Schneider, 2009a), these results suggest that young learners are not aware of the easily learned, easily remembered (ELER) heuristic and they rather expect that increased ST is associated with better recall.

Accordingly, the results for the younger children are compatible with two different interpretations. That is, younger children could have either relied on item difficulty or on the memorizing effort heuristic when making JOLs. In contrast, the findings for the older children are more clear-cut. The fact that the correlation between ST and JOLs remained significant shows that older children acted in accord with the heuristic indicating that the more time one spends studying an item the *less* likely one will be to recall it in the future. The correspondence between cue utilization and cue validity that was observed in the older age groups possibly contributed to the accuracy of their JOLs in predicting inter-item differences in recall performance.

However, it could also be argued that partialling out normative item difficulty from within-person correlation between item STs and item JOLs may not be an optimal means for deciding between the two different interpretations. It is possible that each individual's assessment of item difficulty deviates somewhat from normative item difficulty. Therefore, the possibility that the variance that is unique to each individual's assessment accounts for the partial correlation between item difficulty and JOL for that individual cannot be ruled out. Because normative item difficulty was collected in a sample of second and fourth graders, the older children's individual assessments of item difficulty may even deviate more from the normative item difficulty, resulting in somewhat higher partial correlations for the older than for the younger children. This possibility has to be taken into account.

Apart from that, the possible effects of normative item difficulty merit a further examination, as the results presented so far mainly rely on correlational data. A recent study by Koriat (2008) addressed this point by using study materials based on unrelated words that differed minimally in normative item difficulty. This manipulation yielded similar results for adults compared to Koriat *et al.* (2006). In order to assess effects of item difficulty in children, we propose that as a further step, study material differing only minimally in normative associative relatedness should also be used in a sample of school-aged children. This would allow for a conclusive evaluation of the possible effects of normative item difficulty on the examined aspects of cue utilization. A recent study by Koriat *et al.* (2009a) shows evidence for this argumentation. They extended results obtained with adults (Koriat, 2008) and showed that also second- and fourth-graders' recall and JOLs decreased with an increasing number of trials to acquisition, supporting both the validity of the ELER heuristic in young children and its utilization in monitoring one's own learning. These results were obtained with a list consisting of unrelated and related paired associates. When using only hard pairs as stimuli, however, fourth graders' but not second graders' JOLs evidenced reliance on this heuristic in making JOLs. Thus, this provides developmental support for the described negative correlations even when there are no normative differences in item difficulty.

So far, our argumentation focused on the idea that the outcome from control processes may serve as a basis for monitoring processes during learning. However, as mentioned by Koriat *et al.* (2006), this does not mean that the 'monitoring–affects–control' and 'control–affects–monitoring' hypotheses are mutually exclusive. Koriat *et al.* (2006) noticed that ST rather includes both, a monitoring and a control function. The control function is stressed in goal-driven learning contexts, i.e., when ST is used as a strategic tool for regulating memory performance towards the achievements of desired objectives given specific constraints. The monitoring function of ST emerges in self-paced learning. With the aim of assessing the future recallability of an item, ST is used as an index of the invested learning effort and thus acts as a mnemonic cue. The results of Koriat and colleagues (Koriat & Ma'ayan, 2005; Koriat *et al.*, 2006) so far provided evidence for both of the presented models within the same situation. In case of data-driven effort, greater effort was shown to be associated with lower JOLs and lower recall. In contrast, when effort was rather goal driven, it was associated with higher JOLs and better recall. The authors obtained their results by assigning a different number of points for the to-be-learned items. The comparison of mean ST and mean JOLs for one-point and three-point items showed a positive correlation between the two variables, thus reflecting the control function of ST. Nevertheless, within each incentive level, the more ST was invested, the lower was the corresponding JOL. Consequently, the proposed distinction must not be understood in a dogmatic way.

Based on the common assumption that monitoring guides control processes, we have contributed evidence for the recently advanced idea that control processes may also serve as a basis for metacognitive monitoring. These findings are also in account with recent conceptualizations of metacognition (e.g., Efklides, 2001, 2008). We strongly encourage further examination of this topic, especially with regard to the role of item difficulty in the described relationship between ST and JOLs. The so far presented evidence for monitoring being based on feedback from control operations, should be pursued by a detailed exploration of the possible combinations of monitoring and control processes. This presents a fruitful and promising area of research within the topic of procedural metacognition under a developmental aspect.

Acknowledgements

The presented study is part of the project 'The intricate relationships between metacognitive monitoring and metacognitive control during learning: a developmental perspective'. This research was supported by a grant from the G.I.F., the German–Israeli Foundation for Scientific Research and Development.

References

Ackerman, B. P. (1981). Encoding specificity in the recall of pictures and words in children and adults. *Journal of Experimental Child Psychology*, *31*, 193–221.

Barnes, A. E., Nelson, T. O., Dunlosky, J., Mazzoni, G., & Narens, L. (1999). An integrative system of metamemory components involved in retrieval. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII – cognitive regulation of performance: Interaction of theory and application* (pp. 287–313). Cambridge, MA: MIT Press.

Brown, A. L., Smiley, S. S., & Lawton, S. Q. C. (1978). The effects of experience on the selection of suitable retrieval cues for studying texts. *Child Development*, *49*, 829–835.

Cramer, P. (1972). A developmental study of errors in memory. *Developmental Psychology*, *7*, 204–209.

Cramer, P. (1973). Evidence for a developmental shift in the basis for memory organization. *Journal of Experimental Child Psychology*, *16*, 12–22.

Dufresne, A., & Kobasigawa, A. (1989). Children's spontaneous allocation of study time: Differential and sufficient aspects. *Journal of Experimental Child Psychology*, *47*, 274–296.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory and Cognition*, *20*, 374–380.

Efklides, A. (2001). Metacognitive experiences in problem solving: Metacognition, motivation, and self-regulation. In A. Efklides, J. Kuhl, & R. M. Sorrentino (Eds.), *Trends and prospects in motivation research* (pp. 297–323). Dordrecht, The Netherlands: Kluwer.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review*, *1*, 3–14.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, *13*, 277–287.

Kobasigawa, A., & Metcalf-Haggert, A. (1993). Spontaneous allocation of study time by first- and third-grade children in a simple memory task. *Journal of Genetic Psychology*, *154*, 223–235.

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.

Koriat, A. (2006). Are we frightened because we run away? Some evidence from metacognitive feelings. In B. Uttl, N. Ohta, & A. L. Siegenthaler (Eds.), *Memory and emotion: Interdisciplinary perspectives* (pp. 83–103). Malden, MA: Blackwell.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *Cambridge handbook of consciousness* (pp. 289–325). New York: Cambridge University Press.

Koriat, A. (2008). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory and Cognition*, *36*, 416–428.

Koriat, A., Ackerman, R., Lockl, K., & Schneider, W. (2009a). The easily learned, easily-remembered heuristic in children. *Cognitive Development*, *24*, 169–182.

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

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 409–517.

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. Troppe (Eds.), *Dual process theories in social psychology* (pp. 483–502). New York: Guilford.

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language*, *52*, 478–492.

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General*, *135*, 36–69.

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

Lockl, K., & Schneider, W. (2002a). Zur Entwicklung des selbstregulierten Lernens im Grundschulalter: Zusammenhänge zwischen Aufgabenschwierigkeit und Lernzeiteinteilung [The development of self-regulated learning in elementary school children: Associations between task difficulty and allocation of study time]. *Psychologie in Erziehung und Unterricht*, *49*, 3–16.

Lockl, K., & Schneider, W. (2002b). Developmental trends in children's feeling of knowing judgements. *International Journal of Behavioral Development*, *26*, 327–333.

Lockl, K., & Schneider, W. (2003). Metakognitive Überwachungs- und Selbstkontrollprozesse bei der Lernzeiteinteilung von Kindern. [Metacognitive monitoring and self-control processes for children's allocation of study time]. *Zeitschrift für Pädagogische Psychologie*, *17*, 173–183.

Lockl, K., & Schneider, W. (2004). The effects of incentives and instructions on children's allocation of study time. *European Journal of Developmental Psychology*, *1*, 153–169.

Masur, E. F., McIntyre, C. W., & Flavell, J. H. (1973). Developmental changes in apportionment of study time among items in multitrial free recall task. *Journal of Experimental Child Psychology*, *15*, 237–246.

Means, B. A., & Rohwer, W. D. (1976). Attribute dominance in memory development. *Developmental Psychology*, *12*, 411–417.

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Child Psychology: General*, *131*, 349–363.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, *51*, 102–116.

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.

Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science*, *5*, 207–213.

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the 'labor-in-vain-effect'. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 676–686.

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. 125–173). New York: Academic Press.

Roebers, C., von der Linden, N., Howie, P., & Schneider, W. (2007). Children's metamemorial judgments in an event recall task. *Journal of Experimental Child Psychology*, *97*, 117–137.

Schneider, W. (1998). The development of procedural memory in childhood and adolescence. In G. Mazzoni & T. O. Nelson (Eds.), *Monitoring and control processes in metacognition and cognitive neuropsychology* (pp. 1–21). Mahwah, NJ: Erlbaum.

Schneider, W. (2010). The development of metacognitive competencies. In B. Glatzeder, V. Goel & A. von Müller (Eds.) *Towards a theory of thinking: Building blocks for a conceptual framework* (pp. 203–214). Berlin, Germany: Springer.

Schneider, W., & Bjorklund, D. F. (1998). Memory. In W. Damon, D. Kuhn, & R. S. Siegler (Eds.), *Handbook of child psychology: Cognition, perception and language* (Vol. 2, pp. 467–521). New York: Wiley.

Schneider, W., & Bjorklund, D. F. (2003). Memory and knowledge development. In J. Valsiner & K. Conolly (Eds.), *Handbook of developmental psychology* (pp. 370–403). London: Sage.

Schneider, W., & Lockl, K. (2008). Procedural metacognition in children: Evidence for developmental trends. In J. Dunlosky & R. A. Bjork (Eds.), *A handbook of memory and metamemory* (pp. 391–409). Mahwah, NJ: Erlbaum.

Schneider, W., & Pressley, M. (1997). *Memory development between 2 and 20*. Hillsdale, NJ: Erlbaum.

Schneider, W., Visé, M., Lockl, K., & Nelson, T. O. (2000). Developmental trends in children's memory monitoring: Evidence from a judgment-of-learning task. *Cognitive Development*, *15*, 115–134.

Schunk, D. H., & Zimmerman, B. J. (Eds.), (1998). Self-regulated learning: From teaching to self-reflective practice. New York: Guilford.

Son, L. K. (2005). Metacognitive control: Children's short-term versus long-term study strategies. *Journal of General Psychology*, *132*, 347–363.

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.

Son, L. K., & Schwartz, B. L. (2002). The relation between metacognitive monitoring and control. In T. J. Perfect & B. L. Schwartz (Eds.), *Applied metacognition* (pp. 15–38). Cambridge: Cambridge University Press.

Thiede, K. W., & Anderson, M. C. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology*, *28*, 129–160.

Received 6 May 2009; revised version received 21 September 2009

Appendix 1

List of the picture pairs used as study material

		Picture pairs	
		Cue	Target
Practice trials	a	Tennis racket	Tennis ball
	b	Football	Candle
Test trials – related	1	Cup	Plate
	2	Shirt	Cravat
	3	Sun	Moon
	4	Eye	Ear
	5	Hammer	Nail
	6	Bird	Nest
	7	Apple	Pear
	8	Plug	Socket-outlet
	9	Sock	Shoe
	10	Tree	Leaf
	11	Broom	Dustpan
	12	Cat	Dog
Test trials – unrelated	13	Lamp	Frog
	14	Car	Fish
	15	Trumpet	Umbrella
	16	Carrot	Book
	17	Cow	Scissors
	18	Pizza	Violin
	19	Dandelion	Dice
	20	Butterfly	Bed
	21	Piano	Banana
	22	Cloud	Cupboard
	23	Chair	Clock
	24	Saw	Airplane