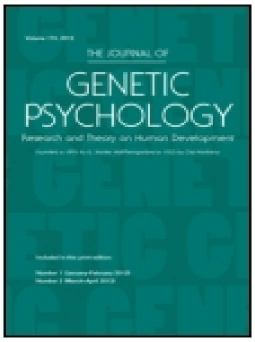
A Longitudinal Study of Semantic Grouping Strategy Use in 6-11-Year-Old Children: Investigating Developmental Phases, the Role of Working Memory, and Strategy Transfer



The Journal of Genetic Psychology

Research and Theory on Human Development

ISSN: 0022-1325 (Print) 1940-0896 (Online) Journal homepage: https://www.tandfonline.com/loi/vgnt20

A Longitudinal Study of Semantic Grouping Strategy Use in 6–11-Year-Old Children: Investigating Developmental Phases, the Role of Working Memory, and Strategy Transfer

Tamara M. J. Schleepen & Lisa M. Jonkman

THE JOURNAL OF GENETIC PSYCHOLOGY, *175*(6), 451–471, 2014

ISSN: 0022-1325 print / 1940-0896 online
DOI: 10.1080/00221325.2014.958126

Routledge
Taylor & Francis Group

A Longitudinal Study of Semantic Grouping Strategy Use in 6–11-Year-Old Children: Investigating Developmental Phases, the Role of Working Memory, and Strategy Transfer

Tamara M. J. Schleepen and Lisa M. Jonkman
Maastricht University, the Netherlands

ABSTRACT. This two-cohort longitudinal study on the development of the semantic grouping strategy had three goals. First, the authors examined if 6–7-year-olds are nonstrategic before becoming strategic after prompting at 8–9 years of age, and if 8–9-year-olds are prompted strategic before spontaneous strategy use at 10–11 years of age. Children 6–7 and 8–9 years old performed two sort-recall tasks (one without and one with a grouping prompt) at two time points separated 1.5 years from each other. Second, the authors investigated whether short-term or working memory capacity at time point 1 predicted recall in children who did or did not use the semantic grouping strategy 1.5 years later. Third, the authors investigated whether prompted strategic children and children who used the strategy spontaneously differed in strategy transfer to a new task. Developmental results confirmed previous cross-sectional results, but in a longitudinal two-cohort study 6–7-year-olds were nonstrategic, and became prompted strategic around 8–9 years of age, followed by spontaneous strategy use at age 10–11 years. The authors found that memory capacity was not predictive of later use of the strategy. New findings were that prompted strategic children were as equally able as spontaneously strategic children to transfer the strategy to a new task, albeit with smaller recall benefits.

Keywords *semantic grouping strategy, developmental phases, strategy transfer, working memory*

The use of memory strategies, such as grouping to-be-remembered information on similar features (e.g., on semantic category or color) is crucial for learning and remembering. In particular for children, it is important to use memory strategies efficiently because they are important for learning several academic skills (Dehn, 2008). Also, children who make less use of such strategies have been reported to be at risk for developing learning difficulties (Bauer, 1977).

Given the key role (semantic) memory strategies play in learning, it is of great importance to identify the factors involved in children's ability to intentionally and effectively apply them on semantically related material. In the present study we examined the role of age and working memory capacity within a two-cohort longitudinal design. Transfer of strategy use to a new

Received December 20, 2013; accepted August 18, 2014.
Address correspondence to Tamara M. J. Schleepen, Maastricht University, Cognitive Neuroscience, Oxfordlaan 55, Maastricht 6200, the Netherlands; tamara.schleepen@maastrichtuniversity.nl (e-mail).

task was also studied. A previous cross-sectional study (Schleepen & Jonkman, 2012) provided evidence for children passing through three developmental phases of strategy use (semantic grouping) between the ages of 6 and 12 years old. That is, the majority of 6–7-year-old children appeared to be nonstrategic, most of the 8–9-year-olds did not spontaneously initiate a semantic grouping strategy but could apply it after receiving a simple grouping prompt, and most children older than 10 years (i.e., 10–12-year-olds) initiated semantic grouping strategies themselves without prompting. In the same study, children's working memory capacity was identified as a mediator of the successful application of the semantic grouping strategy (that is when it leads to better memory). In cross-sectional designs, however, group differences and variance might have played a role in the reported developmental results. One aim of the present study was thus to confirm the above-mentioned developmental phases in 6–11-year-old children, using a two-cohort longitudinal design. Furthermore, longitudinal (repeated) measurement allows for the study of the predictiveness of working memory capacity at a younger age for strategy use at a later age, which has to our knowledge not been investigated before but is important for the identification of potential risk factors present at an early age. A third important aim of the present study, to our knowledge not studied before in children age 6–11 years old, was to examine whether children that did or did not need prompting to initiate the semantic grouping strategy differed in how well they could transfer the strategy to a new memory task. We present a short review of the previous literature on these three topics (development of semantic grouping, involvement of working memory, and transfer).

Development of Semantic Grouping Strategies

Two aspects distinguish immature from mature memory strategy development. The first is the level at which children are able to initiate strategies spontaneously/deliberately whenever the to-be-learned material asks for it. The second is the success of strategy application, that is whether it results in memory benefits or not. Three different developmental phases have been distinguished in the developmental literature: (a) the mediation-deficient phase (first reported by Reese, 1962) in which children are largely nonstrategic, meaning that they are not able to spontaneously initiate semantic grouping strategies (Bjorklund & de Marchena, 1984) and cannot be prompted or trained to do so (Schleepen & Jonkman, 2012); (b) a production-deficient phase in which children do also not engage in spontaneously initiated strategy use but do so successfully after prompting (Flavell, 1970); and (c) a utilization deficient phase in which children are able to produce/initiate a strategy but do not benefit from it (i.e., do not show improved recall performance; Bjorklund, Miller, Coyle, & Slawinski, 1997; Schwenck, Bjorklund, & Schneider, 2009). Whether there are specific age-ranges at which the majority of the children proceed through these stages is not completely clear. The utilization deficiency has received by far the most attention in the developmental literature and has mostly been reported in cross-sectional studies (Bjorklund & Coyle, 1995; Schwenck et al., 2009). In the few longitudinal studies that have been done by Schneider and colleagues, utilization deficiencies were however found in only a small minority of children (Kron-Sperl, Schneider, & Hasselhorn, 2008; Schneider, Kron, Hunnerkopf, & Krajewski, 2004; Schneider, Kron-Sperl, & Hunnerkopf, 2009) or were not found at all (Schlagmuller & Schneider, 2002). One aim of this study was to examine the development of mediation and production deficiency using a two-cohort repeated measure (longitudinal) design, which has to our knowledge not been done before. Utilization deficiencies were not studied because of their rare occurrence in longitudinal studies (see previous). Based on Schleepen and Jonkman, the two cohorts were a group of 6–7-year-old and a group of 8–9-year-old children that were measured twice, 1.5 years apart. The hypothesis was that the majority of 6–7-year-olds would be first mediation deficient before becoming production deficient 1.5 years later at 8–9 years old. The majority of children in the 8–9-year-old group were hypothesized to pass from production deficient to spontaneous strategy users 1.5 years later at 10–11 years old.

Longitudinal Relation Between Working Memory Capacity and Semantic Grouping Strategy Use

The processes of maintaining and at the same time processing (updating/manipulating) information are called working memory processes, whereas short-term memory processes only involve simple maintenance (Baddeley, 2000). Especially working memory capacity has been reported to be an important factor in explaining individual differences in a variety of cognitive skills, including reading and mathematics (e.g., De Smedt et al., 2009; Gathercole & Baddeley, 1993; Holmes & Adams, 2006). In the memory strategy literature, having sufficient working memory capacity has also been shown to be important for successfully using the semantic grouping strategy in both children (Schleepen & Jonkman, 2012) and adults (McNamara & Scott, 2001; Rosen & Engle, 1997). Although it is not exactly clear for which processes working memory is most needed during the encoding and retrieval of categorized (semantically related) information, the central executive component of working memory (Baddeley) is thought to play a role in mediating organizational strategies such as semantic grouping (Gershberg & Shimamura, 1995) and in maintaining constant conscious awareness of to-be-encoded targets. During the retrieval of categorized information, the central executive is needed during controlled search for earlier stored category names, which would then automatically trigger the category names (Cinan, 2003; Rosen & Engle, 1997).

As far as we know, three other previous studies have investigated the role of short-term memory/working memory capacity (inferred from digit-span forward and backward scores, respectively) in semantic strategy use in children, all from the group of Schneider and colleagues (Kron-Sperl et al., 2008; Schneider et al., 2009; Schneider et al., 2004). Only in one study (Schneider et al., 2004) did consistent strategy users have higher working memory capacity than utilization deficient children, but these groups only contained nine and seven subjects, respectively. In a later study differences in short-term memory instead of working memory capacity contributed significantly to recall performance (Kron-Sperl et al., 2008), but this finding was not replicated in a follow-up study (Schneider et al., 2009). The discrepancy in findings between these developmental studies may be explained by differences in the to-be-studied memory material and in the way strategy use was computed. Jonkman and Schleepen (2012) found a mediating role of working memory capacity (and not short-term memory) on strategy use—recall relations using pictures that did not have well-learned associations (e.g., dolphin-bird instead of cow-milk). This is assumed to make the strongest demands on mental resources (Bjorklund & de Marchena, 1984; Bjorklund & Jacobs, 1985) and probably explains higher working memory involvement.

Also this mediating role of working memory capacity was only found when strategy use was calculated on the basis of clustering during recall and not during sorting, the latter requiring less resources because pictures were in view of the children (see Schleepen and Jonkman, 2012). But as mentioned previously, all of these studies did not study longitudinal relations, which was the goal of the present study.

Strategy Transfer

One factor that plays a crucial role in learning is the ability to transfer new cognitive skills (Cox, 1997). Transfer refers to reusing knowledge learned earlier to aid task performance in another context, such as when a new task is performed that is distinct, but seemingly similar as the one performed during learning (Day & Goldstone, 2012). An interesting question investigated in the present study is whether prompted strategic (production deficient) children are able to transfer their recently learned semantic grouping skills to a new task when not prompted and if they are able to do this as well as consistent strategy users. Although relatively little is known about this issue, there are three studies that provide some information about this. Ringel and Springer (1980) examined strategy transfer of the sorting strategy in 7-, 9- and 11-year-old children and showed that 7- and 9-year-old children only showed strategy transfer to a sort-recall task containing new pictures after having received elaborate instructions, practice, and feedback on strategy use. The 11-year-old children only needed instructions and practice in strategy use (and not feedback) to show strategy transfer. In another study by Schwenck et al. (2009), 4–8-year-old children were trained in the use of a sorting strategy (i.e., group items during study), a clustering retrieval strategy (i.e., group items during recall), or no training. Transfer was assessed two weeks later on a sort-recall task containing new pictures and new semantic categories. By using multivariate cluster analysis, they found that children who were classified as production deficient (i.e., prompted strategic) were not able to generalize the semantic grouping strategy to the transfer task. Finally, in a recent study by Clerc and Miller (2013) it was investigated whether 4-, 4.5-, and 5-year-old children who were initially strategic or became strategic after prompting showed transfer to tasks that were superficially different but had the same underlying logic as the main task. A selective attention memory task was used as the common structure between tasks; in this task children are required to remember one category of items and not pay attention to items of another category during study. The authors reported that although both prompted strategic and initially strategic children transferred the selective strategy to the transfer tasks, recall performance decreased in both groups, pointing to a utilization deficiency. In our previous cross-sectional study we reported results suggesting that the critical age for being able to successfully apply a semantic grouping strategy after prompting is about 8–9 years (Schleepen & Jonkman, 2012), an age group not included in all of the above reviewed studies. Interestingly, this study also suggested that 8–9-year-old children do not always need extensive training to develop strategy use, because they demonstrated successful application of the semantic grouping strategy after only receiving a single general grouping prompt, in which no grouping categories were mentioned and no feedback was provided. However, it is as yet unclear how consistent this strategy acquisition is (i.e., if there is successful transfer to a new task), so we investigated it in this study.

METHOD

Participants

To answer the present research questions 51 of the 83 children that participated in the previous cross-sectional study (Schleepen & Jonkman, 2012) were invited for a second measurement taking place 1.5 years after the first measurement (for data from first measurement, see Jonkman & Schleepen, 2012). Only 51 of the originally participating 83 children were retested because 18 children moved to high school after the first measurement and 14 parents did not return the informed consent form for participation of their child at the second measurement. The 51 remaining children, from which longitudinal data were obtained were divided in two age cohorts of children being either 6–7 (age range: 6.6–7.8 years) or 8–9 (age range: 8.5–9.9 years) years-old at the time of the first measurement. The rationale for this cohort assignment was based on findings in the earlier cross-sectional study that the 6–7-year-old children were nonstrategic and the 8–9-year-old children were prompted strategic (Schleepen & Jonkman, 2012). The follow-up measurement of these two age cohorts enabled us to answer our research question of whether 6–7-year-old children are first nonstrategic before becoming prompted strategic at 8–9 years old, and if 8–9-year-old children first are prompted strategic before they are able to spontaneously and successfully use the semantic grouping strategy after age 10 years (i.e., at 10–11 years old). Because our prior cross-sectional study showed that the majority of children moved on to a next developmental phase every two years, we deliberately chose to use a 1.5-year time interval between the two measurement points to be able to capture these developmental changes.

Demographic characteristics (age, group size, gender, IQ, Child Behavior Checklist [CBCL] attention scores, and socioeconomic status) for both age groups are shown in Table 1. IQ scores were derived from a short form of the Wechsler Intelligence Scale for Children–III (WISC-III, Dutch version; Wechsler, 1991) administered at time point 1, including vocabulary and block design subtests, that correlates .9 with the full test (Jeyakumar, Warinner, Raval, & Ahmad, 2004; Spreen, 1998). Exclusion criteria were the presence of psychiatric or neurological disorders, medication use, an IQ score below 80 or a score above the clinical threshold (t scores > 70) on the attention subscale of the CBCL (Achenbach, 1991). Because attention problems are highly related to working memory deficits, effort was made to exclude undiagnosed attention deficits by allowing parents to fill out the CBCL at time point 1. None of the children met one or more of the exclusion criteria. The study was approved by a local ethical committee of the Maastricht University Faculty of Psychology and Neuroscience.

Procedure

The two test sessions (separated by 1.5 years) took place in a quiet room at the children's school, lasting about 2 hr each. In both sessions, tasks were administered in a fixed order. In addition to the tasks reported on here, children performed a computerized working memory (memory binding) task that was administered in two separate parts during the sessions. With this task we aimed to investigate a research question unrelated to those included in the current study and for

TABLE 1
Demographic Characteristics per Age Group: The Ages of the Children Groups Represent the Ages at the First Measurement Point

	6–7 (n = *26)*			*8–9* (n = *25)*		
Variable	%	M	SD	%	M	SD
Gender[a] (% female)	34.6			52		
Age (years)		7.00	0.40		9.20	0.49
Attention score[b]		54.80	4.50		54.60	4.30
Estimated IQ[c]		108.20	13.40		106.30	12.70
SES[d]		5.90	1.90		6.00	1.50

[a]Gender did not significantly differ between age groups, $\chi^2(1, N = 51) = 1.6, p = .21$.

[b]None of the children scored above the clinical threshold on the attention subscale of the Child Behavior Checklist (CBCL; Achenbach, 1991). For one child in the 6–7-year-old group there were no CBCL scores available. CBCL attention scores did not significantly differ between age groups, $t(48) = 0.09, p = .92$.

[c]IQ was not significantly different between age groups, $t(49) = 0.06, p = .60$.

[d]Socioeconomic status (SES) was determined by Hollingshead's (1975) occupational scale for the parent holding the higher status job (1 or 2 = unskilled or unemployed positions, 3 or 4 = skilled or semiskilled laborers, 5 or 6 = managerial professions, 8 or 9 = major professions). Parental occupation data was not available for one child in the 6–7-year-old group. There was no significant difference in SES between age groups, $t(48) = –0.33, p = .74$.

that reason the results of this task are not described in this article. Each session began with the first four blocks of the computerized memory task, followed by the first sort-recall task. In this task (i.e., the standard task), no grouping instruction was provided and the child's performance was intended to reflect spontaneous use of grouping strategies. After children completed the remaining three blocks of the computerized task, a second sort-recall task (i.e., the instruction task) was presented in which children were prompted to apply a grouping strategy. Children then completed the short form WISC-III and digit-span forward and backward tests. At the end of each session, children were rewarded with a small toy present. See Figure 1 for a time schedule showing which tasks were administered at time point 1 and time point 2.

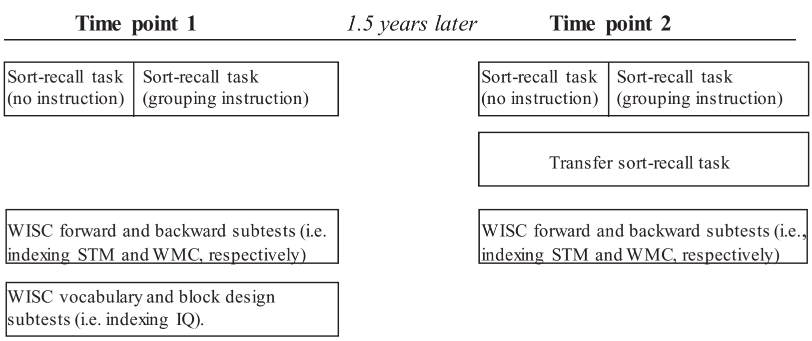


FIGURE 1 Time schedule showing which tasks were administered at time point 1 and time point 2. WISC = Wechsler Intelligence Scale; STM = short-term memory; WM = working memory.

TABLE 2
Stimulus Items per Category in the Two Sort-Recall Tasks and in the Transfer Sort-Recall Task

Semantic category	*Standard task*	*Instruction task*	*Semantic category*	*Transfer task*
Fruit	Peer (Pear)	Ananas (Pineapple)	Tools	Liniaal (Ruler)
	Aardbei (Strawberry)	Kers (Cherry)		Tang (Shrew)
	Citroen (Lemon)	Meloen (Melon)		Zaag (Saw)
	Druiven (Grapes)	Appel (Apple)		Schroevedraaier (Screwdriver)
Animals	Dolfijn (Dolphin)	Kat (Cat)	Vegetables	Paprika (Paprika)
	Hond (Dog)	Olifant (Elephant)		Ui (Union)
	Vogel (Bird)	Koe (Cow)		Mais (Corn)
	Aap (Monkey)	Zwaan (Swan)		Champignon (Agaric)
Clothes	Jas (Coat)	Hoed (Hat)	Vehicles	Fiets (Bicycle)
	Rok (Skirt)	Broek (Trousers)		Auto (Car)
	Das (Tie)	Want (Glove)		Helicopter (Helicopter)
	Sok (Sock)	Jurk (Dress)		Zeilboot (Sailboot)

Tasks

Sort-recall task

The material for the sort-recall task comprised two stimulus sets of 12 black-and-white line drawings of objects, printed on 6 × 5 cm cards. The borders of the cards (0.2 cm) were printed in one of four colors (i.e., green, blue, brown, grey). The objects in each set could be grouped according to semantic category (four objects from three distinct categories; see Table 2) or by color. The color manipulation was already included in our initial study to investigate if children group information on the basis of color (Schleepen & Jonkman, 2012). This was not the case, but the color manipulation was still included in the current study to keep the picture material and its perceptual characteristics exactly similar between measurements. For each stimulus set, different pictures of objects (from the same semantic categories) were used and the two stimulus sets were counterbalanced across the two sort-recall tasks and across age groups.

The choice of line drawings in the two stimulus sets was based on several considerations. First, semantic categories of stimuli were selected that were common to children within the age range of 6–12 years old. This was undertaken based on familiarity rates for 5- and 6-year-old-children collected by Cycowicz, Friedman, Rothstein, and Snodgrass (1997). This procedure yielded highly and equally familiar objects from the three semantic categories. Second, only object stimuli were included that consisted of 3–7-letter words. This restriction ensured that differences in recall performance could not be attributed to the fact that short-item words are better recalled than long-item words (Neath, Bireta, & Surprenant, 2003). Finally, the least-associative items within each category were selected to constitute the two stimulus sets. This selection was accomplished by means of a pilot study with adults in which all possible combinations of pairings of two pictures from the same semantic category were presented to a group of 13 adults. These adults rated the level of association between the items on a 10-point Likert-type scale ranging from 1 (*low associative*) to 10 (*high associative*). Mean association scores for all pairs per category fell between 2.3 and 2.7 (*SD*s = 1.5 and 1.9, respectively). Adults were instructed that the (semantic) associations could be based on different features and that the ratings should be based on their first impressions. The rationale behind including adults in this pilot was that if adults, with their elaborate knowledge bases, were to rate item pairs as having a low semantic association, it could be assumed this would be the case for children.

In the sort-recall tasks, the cards were placed on a 3 row × 4 column array on the table and covered with a cloth. The layout of the cards was similar across participants, with the restriction that no two items from the same semantic or color category laid adjacently. After uncovering the cards, children had to label the pictures to verify that they were familiar with each object. In case an object was unfamiliar (which rarely happened), the correct name was provided by the experimenter without mentioning the corresponding semantic category. Children were then instructed to study the pictures and were told that the pictures would be removed after a while, after which they would be asked to verbally report as many pictures as they still remembered. They were also told that they were allowed to move the pictures in any way that might help them remembering.

Following a 75-s study period, the experimenter made a photograph of the arrangement of the pictures and removed them from the child's view after which children had to count to 50 for a period of 30 s. This served as a buffer clearing task to control for any recency or primacy effects. Children then had to recall as many items as possible, which was voice recorded so that clustering at retrieval could be computed afterward. After 10 s of silence, the experimenter asked whether the child could remember any other pictures. After another continued silence of 15 s, the task was ended.

The second sort-recall task was identical with the exception that, before the start of this task, children were told that it might be easier to remember the pictures if they are organized in groups of pictures that belong together. This instruction was given after children had labeled the pictures. This instruction is considered nondirective because no hints were given as to the perceptual or semantic categories into which pictures could be grouped. At the end of the second sort-recall task, children were checked for color blindness by asking them to label the four colors that were used.

Because children had to remember the same two sets of pictures in the sort-recall tasks at the first and second time points, it was checked whether children perhaps still remembered some words by asking them to report any pictures they might remember from 1.5 years earlier. Twenty-three of the 26 children age 8–9 years old and 20 of the 25 children age 10–11 years old could not remember any pictures from time point 1. On the basis of these findings we concluded that practice effects could not explain possible longitudinal effects.

At the second measurement point, all children also performed a transfer-sort-recall task. In this task, children received the same neutral instructions as in the standard task (see previous). This was done to examine if prompted strategic and spontaneously strategic children were able to generalize strategy use to a sort-recall task with new categories and pictures.

Short-term memory and working memory tasks

The digit-span forward and digit-span backward tests, adapted from the WISC-III, were used as measures of short-term memory and working memory capacity, respectively. Backward digit recall has been shown to be a measure of working memory in children (Gathercole, Pickering, Ambridge, & Wearing, 2004; St Clair-Thompson, 2010).

Transfer sort-recall task

This task was similar as the sort-recall tasks described previously with the exception that this task comprised different pictures of different semantic categories (see Table 2). Thus, just as in prior studies near transfer was studied, that is, transfer of the semantic grouping strategy to a task that is structurally similar to the task performed before instruction or training. Also, instructions were similar as in the standard task (see previous) to study if children were able to spontaneously apply the semantic grouping strategy in the transfer task. Familiarity rates and word length were comparable between the pictures in the standard and instruction task. In addition, all pictures were also lowly associated with each other to study deliberate memory strategy use.

Scoring Strategy Use

The procedures for scoring strategy use were similar to those used in most prior research. Adjusted ratio of clustering (ARC) scores (Roenker, Thompson, & Brown, 1971) were used to assess the amount of color-or semantic sorting during study and clustering at retrieval. The ARC score does not vary systematically with amount recalled (Murphy, 1979) making it an appropriate measure of strategy use when different levels of recall are expected for different groups, as in the present case. In the present study two types of clustering (ARC) scores were calculated. First, the sorting ARC score that reflects strategy use during the study phase of the task and reflects the extent to which pictures from the same color or semantic category are sorted in successive order above a certain chance level. Second, we calculated an ARC retrieval clustering score by counting the words belonging to the same semantic category that were recalled in successive order above a certain chance level during the retrieval phase of the task. ARC scores vary between –1 and 1, with a score of 1 reflecting perfect sorting or clustering, a score of zero indicating sorting or clustering at chance level, and a score below zero representing sorting or clustering below-chance level. In calculating the ARC sorting score, a repetition was counted if two pictures of the same semantic category or color were laid above, below or next to each other. Intrusions (recall of items that were not part of the memory set) and perseverations (repetition of items already recalled) were not included in calculating the ARC retrieval scores. Following Coyle and Bjorklund (1997), sorting during study or clustering at retrieval was considered to be meaningful or above chance (i.e., children were considered strategic) when clustering scores were equal to or greater than .5. An ARC score of .5 reflects a value of slightly more than one standard deviation greater than sorting or retrieval clustering expected by chance.

RESULTS

Because of uninterpretable ARC sorting and retrieval clustering scores for the color dimension in all age groups (ARC score $<.5$), only semantic sorting and semantic retrieval clustering scores

TABLE 3
Means and Standard Deviations in Sorting, Retrieval Clustering Scores, and Recall by Age Group, Task, and Measurement Point

	Time point 1				*Time point 2*			
	6–7		*8–9*		*8–9*		*10–11*	
Task	M	SD	M	SD	M	SD	M	SD
Standard task								
Semantic sorting	−.17	.55	.00	.65	.16	.63	.10	.71
Semantic clustering	−.02	.46	.13	.65	.20	.48	.50	.51
Recall	7.5	2.4	7.9	2.5	9.1	1.8	9.7	2.0
Instruction task								
Semantic sorting	.28	.64	.47	.67	.80	.41	.87	.36
Semantic clustering	.20	.58	.50	.51	.60	.48	.73	.36
Recall	7.2	2.2	8.3	1.9	8.9	1.7	9.8	1.9
Transfer task								
Semantic sorting					.69	.53	.88	.34
Semantic clustering					.53	.53	.70	.46
Recall					8.7	2.1	8.8	2.3

Note. The 8–9-year-olds at time point 2 were 6–7 years old at time point 1, and the 10–11-year-olds at time point 2 were 8–9 years old at time point 1.

were included in the analyses. Table 3 represents the mean clustering-ARC scores for sorting during study, clustering at retrieval, and mean recall scores as a function of age, instruction condition, and time.

Longitudinal Development of Semantic Grouping Strategies

To investigate if 6–7-year-olds are nonstrategic before they become prompted strategic at 8–9 years of age, and if 8–9-year-olds first are prompted strategic before they spontaneously use the strategy at 10–11 years old, semantic sorting during study, semantic clustering at retrieval, and recall scores were analyzed by separate 2 Instruction X 2 Time analyses of variance (ANOVAs) per age group (i.e., 6–7- and 8–9-year-olds at time point 1).

Sorting during study

For both 6–7- and 8–9-year-old children, a main effect of instruction was found, $F(1, 25), = 29.1$, $p < .001$, $\eta^2_p = .54$; $F(1, 24), = 33.4$, $p < .001$, $\eta^2_p = .58$, respectively. For both groups, this main effect indicates that ARC sorting scores were significantly higher after than before instruction. The ANOVA analysis also revealed a main time effect in 6–7-year-olds, $F(1, 25) = 14.1$, $p = .001$, $\eta^2_p = .36$, and a marginal-significant main time effect in 8–9-year-olds, $F(1, 24) = 3.4$, $p = .079$, $\eta^2_p = .12$; in both groups semantic sorting was higher at time point 2 than at time point 1. However, the 6–7-year-old children showed semantic sorting scores below chance (ARC <.5) at time point 1 both before and after prompting. After 1.5 years (at 8–9 years old), these children still had below-chance sorting scores before prompting, but after prompting semantic sorting was above-chance. The children that were 8–9 years old at time point 1 had below-chance sorting scores before instruction at both time points, whereas prompting led to almost above-chance (.47) sorting at time point 1 and clearly above-chance sorting (.87) at time point 2.

Clustering at retrieval

Also for clustering at retrieval, both 6–7 and 8–9-year-old children showed a main effect of instruction/prompting; $F(1, 25), = 12.6$, $p = .002$, $\eta^2_p = .34$, and $F(1, 24), = 11.1$, $p = .003$, $\eta^2_p = .32$, respectively. This main effect indicated higher semantic ARC retrieval scores after than before prompting. Also a main effect of time was found for both groups, 6–7 year olds, $F(1, 25) = 9.5$, $p = .005$, $\eta^2_p = .28$; and 8–9-year-olds, $F(1, 24) = 10.6$, $p = .003$, $\eta^2_p = .31$, indicating that semantic clustering during retrieval was higher at time point 2 than at time point 1. The mean retrieval clustering ARC values (see Table 3) indicate that 6–7-year-olds had below-chance clustering retrieval scores at time point 1 in both instruction conditions. At time point 2 when these children were 8–9 years old they still showed below-chance retrieval clustering before prompting, but prompting now led to above-chance semantic clustering. At time point 1, 8–9-old children had below-chance clustering retrieval scores before prompting while these scores were above chance (ARC = .5) after prompting. After 1.5 years when these children were 10–11 years of age, semantic clustering scores before prompting were at the same level (ARC = .5) as after prompting at time point 1, but further increased to .73 after prompting.

Recall

For recall performance, a main effect of time was found in both age groups, indicating that 6–7-and 8–9-year-old children had higher recall scores at time point 2 than at time point 1, $F(1, 25), = 25.2$, $p = .003$, $\eta^2_p = .50$; $F(1, 24) = 34.2$, $p < .001$, $\eta^2_p = .59$, respectively. Further, we established whether there were any intrusions from prior lists that might have acted as a retrieval cue for items of the current list. In total four intrusions from a prior list occurred (that were made by four different children), but only one intrusion was followed by an item from the same semantic category.

Relation Between Strategy Use and Recall

To determine at what age children were nonstrategic, prompted strategic or spontaneously strategic, Pearson's correlations were computed between sorting during study, retrieval clustering and recall separately for the two age groups. These correlations were computed for the two instruction conditions, separately per time point. See Table 4 for all correlations. Significant positive correlations were found between ARC sorting and recall scores and ARC retrieval and recall scores in 8–9-year-olds at time point 1 after prompting. Also after 1.5 years at 10–11 years old these correlations were significant, both now before and after prompting. In the 6–7-year-olds these correlations reached significance only at time point 2 before prompting (i.e., when they

TABLE 4
Pearson's Correlations Between Sorting During Study, Clustering at Retrieval, and Recall by Age Group, Task, and Measurement Point

	Time point 1		*Time point 2*	
Task	*6–7*	*8–9*	*8–9*	*10–11*
Standard task				
Sorting-clustering	.02	.57**	.75**	.59**
Sorting-recall	−.02	.13	.52**	.50*
Clustering-recall	−.47*	.28	.64**	.44*
Instruction task				
Sorting-clustering	.62**	.53**	.25	.45*
Sorting-recall	.36	.44*	.32	.61**
Clustering-recall	.34	.44*	.27	.49*

Note. The 8–9-year-olds at time point 2 were 6–7 years old at time point 1, and the 10–11-year-olds at time point 2 were 8–9 years old at time point 1.
$^*p < .05$. $^{**}p < .01$.

were 8–9-years old), but because of mean group ARC values that are smaller than <.5 these correlations are not further interpreted.

Summarizing the longitudinal data, 6–7-year-olds had below-chance (ARC < .5) semantic sorting and retrieval clustering scores in both instruction conditions at time point 1, indicating that they are nonstrategic (i.e., mediation deficient). After 1.5 years when these children were 8–9 years old, sorting and retrieval clustering scores in the task without prompting were increased and correlated with recall but were still below chance. After prompting they did show above-chance sorting and retrieval organization, and correlations (*r*) of .32 and .27 were found for sorting-recall and retrieval clustering-recall respectively. Although these latter correlations did not reach significance, the fact that significant correlations of both scores with recall were present in the transfer task allows for the conclusion that these children were prompted strategic (i.e., production deficient) at time point 2. The 8–9-year-olds also had below-chance semantic sorting and retrieval clustering scores at time point 1 before prompting, but the prompt at time point 1 led to above-chance sorting and retrieval clustering scores that were related to increased recall levels. This pattern at time point 1 indicates that the 8–9-year-olds were prompted strategic for both strategies. After 1.5 years when they were 10–11 years old, they were still prompted strategic with respect to the sorting strategy, but had become spontaneously strategic with respect to retrieval clustering (ARC ≥ .5); retrieval clustering scores related to recall both without and with prompting at time point 2).

Longitudinal Relation Between Working Memory Capacity and Semantic Grouping Strategy Use

To examine if a child's short-term memory and/or working memory capacity at time point 1 was a precursor of successful strategy use (i.e., sorting or retrieval clustering) at time point 2, the following analyses steps were performed (this was done separately for the task without and with a grouping instruction). First, subgroups of children that did or did not apply the sorting or clustering retrieval strategy at longitudinal measurement point 2 were created on the basis of having an ARC score $\geq .5$ or $< .5$ (according to recommendations by Coyle & Bjorklund, 1997). Second, in strategic and nonstrategic groups correlations were computed between forward and backward digit-span scores measured at time point 1 (as a measure of short-term memory and working memory capacity, respectively) and recall scores at time point 2. Because we had specific expectations about these correlations (i.e., we expected that time point 1 digit-span scores would correlate positively with time point 2 recall scores), one-tailed tests were used in correlation analyses. When significant correlations were found, regression analysis was done to test whether short-term memory or working memory capacity explained significant variance in recall performance.

Standard task (without prompting)

Regarding the sorting strategy, there were 33 nonstrategic and 18 strategic children. Results of the correlation analyses showed that in nonstrategic sorters, no significant correlations were found between short-term memory time point 1 scores and recall time point 2 scores ($r = .17$, $p = .39$) or between working memory time point 1 scores and recall time point 2 scores ($r = .10$, $p = .58$). In strategic sorters, these correlations were also not significant (time point 1 short-term memory and time point 2 recall: $r = .21$, $p = .40$; time point 2 working memory and time point 2 recall: $r = -.08$, $p = .77$). For the clustering retrieval strategy, 28 children were classified as nonstrategic and 23 as strategic. It was found that in nonstrategic clusterers, digit-span backward scores at time point 1 correlated positively with recall scores at time point 2 ($r = .36$, $p = .03$, one tailed) and regression analysis revealed that digit-span backward scores at time point 1 explained 13% of the variance in recall scores at time point 2 ($p = .06$). The correlation between short-term memory time point 1 and recall time point 2 was not significant ($r = .11$, $p = .60$). In strategic clusterers, no significant correlations were found between short-term memory time point 1 scores and recall time point 2 scores ($r = .29$, $p = .19$) or between working memory time point 1 scores and recall time point 2 scores ($r = -.32$, $p = .14$).

Instruction task (prompting)

For the sorting strategy, 6 children were classified as nonstrategic and 45 children were classified as strategic. Given the low number of nonstrategic sorters, no further statistical analysis could be carried out in this group. In strategic sorters, time point 1 digit-span backward scores were positively correlated with time point 2 recall scores ($r = .26$, $p = .043$), but regression analysis demonstrated that digit-span backward did not explain significant variance in recall performance ($p > .1$). No significant correlations were found with short-term memory ($r = .15$, $p = .32$). Regarding the clustering retrieval strategy, there were 14 nonstrategic children and 37 strategic children. In these groups, no significant correlations were found between short-term memory time point 1 scores and recall time point 2 scores ($r = .03$, $p = .84$ and $r = .10$, $p = .57$, respectively) or between working memory time point 1 scores and recall time point 2 scores ($r = .42$, $p = .13$; $r = -.18$, $p = .54$, respectively).

TABLE 5
Mean Age, Gender Distribution, and Mean IQ Scores of All Strategy Subgroups Formed at Time Point 2; Means and Standard Deviations in Sorting, Retrieval Clustering Scores, and Recall by Strategy Subgroup in the Transfer Task; and Pearson's Correlations Between Sorting During Study, Clustering at Retrieval, and Recall by Strategy Subgroup in the Transfer Task

	Strategy group for sorting						*Strategy group for clustering*					
	Prompted strategic			*Spontaneously strategic*			*Prompted strategic*			*Spontaneously strategic*		
Descriptives	%	M	SD	%	M	SD	%	M	SD	%	M	SD
Age		9.5	1.3		9.5	1.4		9.0	1.1		10.1	1.3
Gender (% female)	48.2			50.0			47.4			55.6		
IQ		106.0	11.2		111.0	14.0		109.7	13.5		108.4	11.1
Semantic sorting		.89	.31		.94	.27		.74	.52		.99	.04
Semantic clustering		.71	.38		.79	.30		.51	.47		.88	.24
Recall		9.3	1.4		10.3	1.6		8.3	2.2		10.4	1.7
Correlations												
Sorting-clustering		.69**			.46			.77**			−.23	
Sorting-recall		.35			−.11			.79**			.39	
Clustering-recall		.44*			.33			.71**			.33	

Note. For spontaneously strategic sorting-clustering in the sorting group, $p = .058$ (trend significant). For prompted strategic sorting-recall in the sorting group, $p = .077$ (trend-significant).
$^{*}p < .05$. $^{**}p < .01$.

Strategy Transfer

To answer our research question of whether prompted strategic children and spontaneous strategy users differed in the ability to transfer the semantic grouping strategy to a new task, children were first classified as prompted strategic children or spontaneous strategy users on the basis of their individual ARC scores in the sort-recall task administered at time point 2. A child was classified as prompted strategic when ARC scores were $<.5$ in the task without instruction and $\geq .5$ in the task with instruction. A child was considered as spontaneously strategic when ARC scores were $\geq .5$ both in the task without and with instruction. This division in prompted strategic children and spontaneous strategy users was done separately for the sorting and clustering retrieval strategy. To examine possible differences in (the success of) strategy transfer between these different children groups, independent *t* tests were run for sorting scores, retrieval clustering scores and recall scores in the transfer task administered at time point 2. Also, correlations were computed between the strategy measures (ARC sorting and retrieval scores) and recall performance to evaluate if prompted strategic children were able to successfully apply the semantic grouping strategy in the transfer task. See Table 5 for mean ARC scores for sorting during study, clustering at retrieval, and mean recall scores in the transfer-sort-recall task in prompted strategic children and spontaneously strategic children for the sorting and retrieval clustering strategy.

For the sorting strategy, 27 children were classified as prompted strategic and 18 children were classified as spontaneously strategic (the remaining children could not be placed in either group). In the standard task at time point 2 (on the basis of which children were classified as prompted strategic or spontaneously strategic), recall performance differed significantly between prompted strategic and spontaneous strategy users in the task without instruction, $t(43) = -3.40$, $p = .002$, but not in the task with instruction, $t(43) = -0.04$, $p = .97$, indicating that both groups were equally successful in applying the semantic sorting strategy after having received the prompt. However, the analyses in the transfer task revealed that whereas ARC sorting and ARC retrieval scores did not significantly differ between prompted strategic and spontaneously strategic children (who both had ARC scores $\geq .5$), $t(43) = -0.51$, $p = .61$, and, $t(43) = -0.75$, $p = .46$, respectively, the latter had higher recall scores than the former, $t(43) = -2.10$, $p = .040$.

For the clustering retrieval strategy, 19 children were classified as prompted strategic and 18 as spontaneously strategic (the remaining children could not be placed in either group). Results in the standard task showed the expected pattern of recall performance differing significantly between prompted strategic children and spontaneous strategy users in the task without instruction, $t(35) = -2.50$, $p = .016$, but not in the task with instruction, $t(35) = -0.70$, $p = .49$, indicating that both groups were equally successful in using the semantic retrieval strategy in the task after prompting. In the transfer task however, spontaneously strategic children had significantly higher ARC sorting, $t(35) = -2.1$, $p = .047$; ARC retrieval, $t(35) = -3.00$, $p = .005$; and recall scores, $t(35) = -3.30$, $p = .002$ compared to prompted strategic children. This indicates that although both prompted strategic and spontaneously strategic children had above-chance ($>.5$) ARC sorting and ARC retrieval clustering scores (see Table 5), strategy use in the transfer task was higher and more successful in latter than in the former.

Regarding overlap of the strategy subgroups in sorting and retrieval clustering, 13 children were prompted strategic both for the sorting and clustering retrieval measure, and also 13 children were spontaneously strategic both for the sorting and clustering retrieval measure. There were no significant differences in age, gender, or IQ between prompted strategic and spontaneously strategic children (neither for sorting nor for retrieval clustering), except that the latter were significantly older than the former (see Table 5 for mean age, gender distribution, and mean IQ scores in all strategy subgroups).

See Table 5 for the correlations between sorting during study, clustering at retrieval and recall in all four strategy groups (prompted strategic sorters, spontaneously strategic sorters, prompted strategic clusterers, spontaneously strategic clusterers). Summarizing these correlations, prompted strategic sorters and prompted strategic clusterers had significant correlations between sorting scores, clustering retrieval scores, and recall performance. These correlations indicate that these children successfully used the semantic grouping strategy in the transfer task. Most correlations were nonsignificant in spontaneously strategic children, but this is most likely due to ceiling effects. For example, more than 60% of the consistent strategy users (both for the sorting and clustering retrieval strategy) had a recall score of 11 or 12 (maximum = 12), and maximum ARC sorting and retrieval scores were also obtained by large percentages of these children (ranging from 70% to 90% of the children that had the maximum ARC score of 1).

DISCUSSION

The present longitudinal study had three goals: (a) to investigate if 6–7-year-olds first are nonstrategic before they are prompted strategic at 8–9 years of age, and if 8–9-year-olds first are prompted strategic before they are able to spontaneously use the semantic grouping strategy after age 10 (i.e., at 10–11 years of age); (b) to examine if short-term memory or working memory capacity predicts semantic strategy use longitudinally; and (c) to examine whether prompted strategic demonstrate strategy transfer when not prompted. These three goals will be separately discussed subsequently.

Longitudinal Development of Semantic Grouping Strategies

The present longitudinal study, that included a subset of children from our earlier cross-sectional study (Schleepen & Jonkman, 2012), confirmed the progression of the majority of the children from mediation deficient at 6–7 years old to production deficient at 8–9 years old to consistent strategy use after age 10 years (i.e., 10–11 years old), now using a two-cohort longitudinal instead of cross-sectional design.

The present longitudinal data revealed that children that were nonstrategic at 6–7 years old progressed to being prompted strategic, i.e., being able to use the strategy after prompting, at 8–9 years old. This was the case for both the sorting and clustering retrieval strategy as was shown by ARC sorting and retrieval clustering scores $\geq .5$ only after prompting at time point 2. In 8–9-year-old children, the sorting and clustering retrieval strategy showed a different succession of developmental phases. With respect to the clustering retrieval strategy it was found that 8–9-year-old prompted strategic children (who at this age only had clustering scores above-chance leading to higher recall after prompting) were spontaneously strategic at 10–11 years old. This was shown by above-chance clustering and significant clustering-recall relations in the sort-recall task without and with prompting. Thus, after 1.5 years these children still successfully applied the clustering strategy during the retrieval of pictures, now also without prompting. With respect to the sorting strategy, at 8–9 years old children only applied sorting above-chance after prompting and this was still the case 1.5 years later. The reason that 8–9-year-old children were still prompted strategic at time point 2 (10–11 years old) may be explained by older children sorting more covertly when not receiving any prompts (making it more difficult to measure sorting), which would be supported by earlier reports of older children using more covert strategies during selective learning whereas younger children used more overt strategies like naming or pointing (Schwenck et al., 2009). Prompting might have then only led to more overt sorting without leading to any memory benefits as shown by recall scores being 9.7 and 9.8 items, respectively, before and after prompting.

Importantly, whereas different lists were used for the two sort recall tasks (prompted or not prompted) within one session, the same lists were used 1.5 years later. However, it is unlikely that this influenced our results because we verified that 88.5% of the 8–9-year-olds and 80% of the 10–11-year-olds did not remember any of the pictures they had studied 1.5 years earlier at the beginning of time point 2. Also, excluding the children that did remember any pictures from time point 1 from the statistical analyses did not change the results.

Most former (longitudinal) studies focused primarily on utilization deficiencies and measured on a finer time scale (e.g., six-month time lags), which better allowed for the study of individual differences in strategy use and indeed these studies reported large individual differences in strategy use (Kron-Sperl et al., 2008; Schneider et al., 2004; Schneider et al., 2009). Whereas the study of individual differences is very important, the present study shows that over extended time periods of 1.5 years, from the age of 6 years onward the majority of the children do seem to follow a relatively fixed pattern of first being mediation deficient, followed by production deficient before becoming a spontaneous and succesful strategy user around 10–11 years old. Such knowledge about the developmental stage in which the majority of childrern in a certain age range is very relevant for the design of educational programs often targeting children within certain fixed age groups.

Longitudinal Relation Between Working Memory Capacity and Semantic Grouping Strategy Use

Because our previous cross-sectional study (Schleepen & Jonkman, 2012) showed a relation between working memory capacity and successful use of the semantic grouping strategy, a third goal was to examine this relationship in a longitudinal design. Specifically, we examined whether working memory capacity at time point 1 (inferred from digit-span backward scores) predicted strategy use 1.5 years later at time point 2. This was not the case, as working memory capacity only (trend significantly) predicted recall scores in nonstrategic clusterers 1.5 years later. Several reasons may be given for why we did not find that working memory capacity is predictive for later (successful) strategy use. First, it could be argued that a developmental shift took place in the factors contributing to successful use of this strategy. That is, whereas children might have relied on working memory-resources during strategy implementation at time point 1, during the 1.5-year transition period other factors not measured in this study (e.g., a child's knowledge base; Bjorklund, 1987) might have become to play a more important role in the development of the semantic grouping strategy. Second, while the backward digit-span test has been shown to have adequate internal consistency (Waters & Caplan, 2003) and thus can be considered a suitable measure of working memory in children (St Clair-Thompson, 2010), it may be that using a composite score reflecting performance on several (partly different) working memory tests may have better predictive value than the score on a single working memory test. Future longitudinal research that addresses the working memory capacity–semantic strategy use link should take these points into consideration.

Strategy Transfer

With respect to strategy transfer we found that prompted strategic sorters and prompted strategic clusterers were able to generalize the semantic sorting as well as clustering retrieval strategy to the near transfer sort-recall task, as was shown by ARC scores $\geq .5$. Moreover, significant positive correlations between strategy use (ARC scores) and recall performance in both groups indicate that they successfully used the semantic sorting and clustering retrieval strategy in the transfer task. However, prompted strategic sorters and prompted strategic clusterers differed in the extent to which their semantic strategy use in the transfer task was comparable to that of spontaneous strategy users. That is, prompted strategic sorters showed comparable transfer of both the sorting and clustering retrieval strategy as spontaneous strategy users for sorting (equal ARC scores). In contrast, for retrieval clustering, prompted strategic clusterers generalized these strategies to a lesser extent to the transfer task than spontaneous strategy users (had lower ARC sorting and clustering retrieval scores). This is in accordance with the lower recall scores that were found in prompted strategic children compared to spontaneous strategy users. It has to be noted though that lower recall in prompted strategic children was found irrespective of whether the prompted group was defined based on sorting or cluster scores. Thus although strategy use was comparable between prompted and spontaneous users and both showed transfer of the grouping strategy to a new task, the latter showed more frequent and more efficient strategy transfer, leading to higher recall benefits.

As far as we know, only one other study investigated strategy transfer in prompted strategic children separately for the sorting and clustering retrieval measure (Schwenck et al., 2009). In this study it was reported that neither prompted strategic sorters nor prompted strategic clusterers were able to transfer their semantic grouping skills, even though these children became prompted strategic after having received rather extensive training in using semantic grouping strategies. The absence of strategy transfer in the study of Schwenck et al. (2009) may be explained by the fact that the prompted strategic children in their study were younger (between 4 and 8 years old) than in our study (8–11 years old at time point 2 when strategy transfer was studied). The present findings add to this literature by showing that children 8 years of age and older who became strategic after only having received a general grouping prompt, show successful strategy transfer without prompting, although their performance in the transfer task is still inferior compared to that of spontaneous strategy users. If more extensive training would have been provided (e.g., a training similar to the one provided by Schwenck et al. [2009], and also including training of metacognitive skills), prompted strategic children would possibly have shown equal benefit of application of the strategies as spontaneous strategy users. This hypothesis should be investigated in future studies.

Although not investigated here, different factors could possibly explain the less efficient strategy transfer in prompted strategy users. Two such factors are the development of working memory capacity and metacognitive skills. Because strategy use is likely to be less automatized in prompted than spontaneous strategy users application of the strategy likely required much more mental resources (working memory capacity) from prompted users and it may be that at the studied age of 6–9 years old mental (working memory) capacity fell short of meeting such demands leading to lower recall. Several studies have demonstrated that experts require less cognitive resources than novices for their performance in several memory tasks (Allen, McGeorge, Pearson, & Milne, 2004; Beilock, Wierenga, & Carr, 2002; Rowe & McKenna, 2001). The finding of prompted users showing better transfer of sorting than clustering might also support such a capacity explanation, because in the present task clustering depended much more on working memory resources than sorting as the pictures were visible during study reducing working memory load substantially. Another factor of importance might be immaturity in metacognitive skills causing a failure in children's ability to recognize the performance benefits of strategy use, thereby not leading to subsequent consistent use of it. Findings by Melot (1998) support this, by demonstrating that among 6–9-year-old children, those children that better understood that strategy use had the potential of improving their memory performance showed higher strategy use on a posttest. Such metacognitive immaturity may also explain why transfer was better in prompted strategic sorters than prompted strategic clusterers because during sorting children see their own groupings, which is thought to increase a child's metacognitive awareness of the performance benefits of strategy use (Schwenck et al., 2009). Although both working memory and metacognitive skills seem to play an essential role in successful strategy transfer, future researchers should investigate which of these factors contributes most to this.

AUTHOR NOTES

Tamara M. J. Schleepen is a post-doctoral researcher at Maastricht University (the Netherlands). During her Ph.D. program, her research focused on the development pattern and biological basis of memory (strategies). Currently, she also studies how arithmetic skills develop and which factors are underlying this development. **Lisa M. Jonkman** is currently an associate professor in developmental cognitive neuroscience at Maastricht University (the Netherlands). Her research focuses on the cognitive and neurobiological development of attention, executive control, and memory in childhood and adolescence.

REFERENCES

Achenbach, T. M. (1991). *Manual for the child behavior checklist/4–18*. Burlington, VT: University of Vermont, Department of Psychiatry.

Allen, R., McGeorge, P., Pearson, D., & Milne, A. B. (2004). Attention and expertise in multiple target tracking. *Applied Cognitive Psychology*, *18*, 337–347.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

Bauer, R. H. (1977). Memory processes in children with learning disabilities: Evidence for deficient rehearsal. *Journal of Experimental Child Psychology*, *24*, 415–430.

Beilock, S. L., Wierenga, S. A., & Carr, T. H. (2002). Expertise, attention, and memory in sensorimotor skill execution: Impact of novel task constraints on dual-task performance and episodic memory. *The Quarterly Journal of Experimental Psychology. Human Experimental Psychology*, *55*, 1211–1240.

Bjorklund, D. F. (1987). How age changes in knowledge base contribute to the development of children's memory: An interpretive review *Developmental Review*, *7*, 93–130.

Bjorklund, D. F., & Coyle, T. R. (1995). Utilization deficiences in the development of memory strategies In F. E. Weinert & W. Schneider (Eds.), *Memory performance and competencies: Issues in growth and development* (pp. 161–180). Mahwah, NJ: Erlbaum.

Bjorklund, D. F., & de Marchena, M. R. (1984). Developmental shifts in the basis of organization in memory: The role of associative versus categorical relatedness in children's free recall. *Child Development*, *55*, 952–962.

Bjorklund, D. F., & Jacobs, J. W. (1985). Associative and categorical processes in children's memory: The role of automaticity in the development of organization in free recall. *Journal of Experimental Child Psychology*, *39*, 599–617.

Bjorklund, D. F., Miller, P. H., Coyle, T. R., & Slawinski, J. L. (1997). Instructing children to use memory strategies: Evidence of utilization deficiencies in memory training studies. *Developmental Review*, *17*, 411–441.

Cinan, S. (2003). Executive processing in free recall of categorized lists. *Learning and Motivation*, *34*, 240–261.

Clerc, J., & Miller, P. H. (2013). Utilization deficiencies and transfer of strategies in preschoolers. *Cognitive Development*, *28*, 76–93.

Cox, B. D. (1997). The rediscovery of the active learner in adaptive contexts: A developmental-historical analysis of transfer of training. *Educational Psychologist*, *32*, 41–55.

Coyle, T. R., & Bjorklund, D. F. (1997). Age differences in, and consequences of, multiple-and variable-strategy use on a multitrial sort-recall task. *Developmental Psychology*, *33*, 372–380.

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology*, *65*, 171–237.

Day, S. B., & Goldstone, R. L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist*, *47*, 153–176.

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquiere, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology*, *103*, 186–201.

Dehn, M. J. (2008). *Working memory and academic learning: Assesment and intervention*. Hoboken, NJ: Wiley.

Flavell, J. H. (1970). Developmental studies of mediated memory. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 5, pp. 181–211). New York, NY: Academic Press.

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working-memory: A critical building-block for reading development and vocabulary acquisition. *European Journal of Psychology of Education*, *8*, 259–272.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*, 177–190.

Gershberg, F. B., & Shimamura, A. P. (1995). Impaired use of organizational strategies in free recall following frontal lobe damage. *Neuropsychologia*, *33*, 1305–1333.

Hollingshead, A. B. (1975). *Four factor index of social status*. New Haven, CT: Yale University of Department of Sociology.

Holmes, J., & Adams, J. W. (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology*, *26*, 336–366.

Jeyakumar, S. L. E., Warinner, E. M., Raval, V. V., & Ahmad, S. A. (2004). Balancing the need for reliability and time efficiency: Short forms of the Wechsler Adult Intelligence Scale–III. *Educational and Psychological Measurement*, *64*, 71–87.

Kron-Sperl, V., Schneider, W., & Hasselhorn, M. (2008). The development and effectiveness of memory strategies in kindergarten and elementary school: Findings from the Wurzburg and Gottingen longitudinal memory studies. *Cognitive Development*, *23*, 79–104.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition*, *29*, 10–17.

Melot, A. M. (1998). The relationship between metacognitive knowledge and metacognitive experiences: Acquisition and re-elaboration. *European Journal of Psychology of Education*, *13*, 75–89.

Murphy, M. (1979). Measure of category clustering in free recall. In R. Puff (Ed.), *Memory organization and structure* (pp. 51–81). New York, NY: Academic Press.

Neath, I., Bireta, T. J., & Surprenant, A. M. (2003). The time-based word length effect and stimulus set specificity. *Psychonomic Bulletin & Review*, *10*, 430–434.

Reese, H. W. (1962). Verbal mediation as a function of age level. *Psychological Bulletin*, *59*, 502–509.

Ringel, B. A., & Springer, C. J. (1980). On knowing how well one is remembering: The persistence of strategy use during transfer. *Journal of Experimental Child Psychology*, *29*, 322–333.

Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin*, *76*, 45–48.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology–General*, *126*, 211–227.

Rowe, R. M., & McKenna, F. P. (2001). Skilled anticipation in real-world tasks: Measurement of attentional demands in the domain of tennis. *Journal of Experimental Psychology–Applied*, *7*, 60–67.

Schlagmuller, M., & Schneider, W. (2002). The development of organizational strategies in children: Evidence from a microgenetic longitudinal study. *Journal of Experimental Child Psychology*, *81*, 298–319.

Schleepen, T. M. J., & Jonkman, L. M. (2012). Children's use of semantic organizational strategies is mediated by working memory capacity. *Cognitive Development*, *27*, 255–269.

Schneider, W., Kron, V., Hunnerkopf, M., & Krajewski, K. (2004). The development of young children's memory strategies: First findings from the Wurzburg Longitudinal Memory Study. *Journal of Experimental Child Psychology*, *88*, 193–209.

Schneider, W., Kron-Sperl, V., & Hunnerkopf, M. (2009). The development of young children's memory strategies: Evidence from the Wurzburg Longitudinal Memory Study. *European Journal of Developmental Psychology*, *6*, 70–99.

Schwenck, C., Bjorklund, D. F., & Schneider, W. (2009). Developmental and individual differences in young children's use and maintenance of a selective memory strategy. *Developmental Psychology*, *45*, 1034–1050.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York, NY: Oxford University Press.

St Clair-Thompson, H. L. (2010). Backwards digit recall: A measure of short-term memory or working memory? *European Journal of Cognitive Psychology*, *22*, 286–297.

Waters, G. S., & Caplan, D. (2003). The reliabity and stability of verbal working memory measures. *Behavior Research Methods, Instruments & Computers*, *35*, 550–564.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children–Third Edition*. New York, NY: Psychological Corporation.