ELSEVIER

Learning and Instruction 14 (2004) 89–109

Learning and Instruction

www.elsevier.com/locate/learninstruc

The relation between intellectual and metacognitive skills from a developmental perspective

Marcel V.J. Veenman [a,*], Pascal Wilhelm [b], Jos J. Beishuizen [a]

[a] *Department of Developmental and Educational Psychology, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands*

[b] *Department of Instructional Technology, University of Twente, P.O. 217, 7500 AE Enschede, The Netherlands*

Abstract

The first objective of this study was to establish to what extent metacognitive skill development is associated with intelligence. As a second objective, the generality vs. domain-specificity of maturing metacognitive skills was investigated. Both issues have major implications for the training and transferability of metacognitive skills. Participants from four age groups (fourth-, sixth-, and eighth-graders, and university students) performed four inductive learning tasks, representing different domains. Intelligence, metacognitive skillfulness and learning performances were assessed for each participant. Results show that metacognitive skillfulness is a general, person-related characteristic across age groups, rather than being domain-specific. Moreover, metacognitive skills appear to develop and to contribute to learning performance, partly independent of intelligence. Educational implications are discussed.

---

1. Introduction

Metacognition has been recognized as a most relevant predictor of learning (Brown, 1978; Flavell, 1976, 1979; Glaser, 1990; Veenman & Elshout, 1995; Wang, Haertel, & Walberg, 1990, 1993). This study addresses the nature of metacognitive skill development. More specifically, it is investigated whether those skills develop within or beyond the boundaries of intellectual growth (Alexander, Carr, & Schwanenflugel, 1995). A related issue concerns the generality vs. domain-specificity of

---

[*] Corresponding author. Tel.: +31-71-527-3463; fax: +31-71-527-3619.

E-mail address: veenman@fsw.leidenuniv.nl (M.V.J. Veenman).

0959-4752/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.learninstruc.2003.10.004

metacognitive skills. If those skills are mainly domain-specific, they cannot be entirely part of intelligence. In that case, they are most likely acquired along with expertise in a domain (Glaser, 1990). If they represent a general disposition, on the other hand, they may develop either along with intelligence, or as a separable repertoire of acquired skills.

1.1. Metacognitive skillfulness

Though not beyond discussion, metacognitive skillfulness often is distinguished from metacognitive knowledge (Alexander et al., 1995; Baker, 1994; Kuhn, 1999; Schraw & Moshman, 1995; Veenman & Elshout, 1999). The latter concerns the declarative knowledge one has about the interplay between personal characteristics, task characteristics and the available strategies in a learning situation (Flavell, 1979). Such knowledge does not automatically lead to the appropriate task behavior. For instance, a student may know that planning one's activities is necessary and yet refrain from doing so for various reasons. The task may be uninteresting or too difficult, or the student may lack the necessary knowledge and skills for mastery of the task. Metacognitive skills, on the other hand, concern the procedural knowledge that is required for the actual regulation of, and control over one's learning activities (Brown, 1978; Brown & DeLoache, 1978; Flavell, 1992; Kluwe, 1987). Task analysis, planning, monitoring, checking, and recapitulation are manifestations of such skills. These skills can be acquired and eventually executed implicitly (Baker, 1994; Reder, 1996; Veenman, 1993), though some argue that awareness of their metacognitive nature is a prerequisite (Nelson, 1996; Schnotz, 1992).

Metacognitive skills appear to be highly interdependent. By means of thorough task orientation, a metacognitively skilled student is likely to focus on relevant information given in the task assignment, necessary for building an adequate task representation. Consequently, a detailed action plan can be designed. Such an elaborate action plan, containing goals and directions for activities, entails the possibility of process control during task performance. Working systematically according to that plan may enable the student to keep track of progress being made. Evaluation or monitoring activities, which are necessary for detecting faulty procedures and mistakes, are more fruitful within the framework of such an action plan. Finally, elaboration activities like drawing conclusions, recapitulating, and generating explanations are more helpful if they are based on a clear trace of regulatory activities (Mettes, Pilot, & Roossink, 1981; Veenman, Elshout, & Meijer, 1997).

1.2. Intellectual ability as the repertoire of cognitive skills

There exist many conceptions of intelligence (see e.g. Brody, 1992; Carroll, 1993; Sternberg, 1990). Here we adopt a rather pragmatic point of view. Intelligence may be perceived as the magnitude and quality of the human cognitive toolbox, which contains basic cognitive operations (Elshout, 1983). The content and quality of this toolbox is not only determined by the biological substratum (e.g. hereditary factors or brain damage), but increasingly by the opportunities one seeks and the environment offers for acquiring useful cognitive strategies (e.g. at home or in educational settings). In the same vein, Humphreys, (1968, 1989) and Snow (1989; Snow & Lohman, 1984) regard intelligence as the acquired repertoire of intellectual or cognitive skills that is available to a person at a particular point of time. An intelligence test samples this repertoire. The main question here is whether metacognitive skills are essentially part of this cognitive toolbox or repertoire. Sternberg (1990; Davidson, Deuser, & Sternberg, 1994), for instance, regards metacognitive skills as a core process component in his triarchic theory of intelligence. Metacognitive skills, however, may also develop relatively independent from intellectual skills. Slife, Weiss and Bell (1985) adequately formulated this research issue: "The question is whether metacognition can be reduced to cognition".

1.3. Intellectual ability and metacognitive skillfulness

There are three alternative models for describing the relation between intellectual ability and metacognitive skillfulness as predictors of learning (Veenman, 1993; Veenman & Elshout, 1991; Veenman et al., 1997). The first model regards metacognitive skillfulness as a manifestation of intellectual ability, or as an integral part of the intellectual toolbox. According to this intelligence model, metacognitive skills cannot have a predictive value for learning, independent of intellectual ability. Support for the intelligence model was obtained by Elshout and Veenman (1992) in an experiment with novices working with a computer-simulated environment for learning calorimetrics. Several other researchers (Cheng, 1993; Hannah & Shore, 1995; Shore & Dover, 1987; Span & Overtoom-Corsmit, 1986; Zimmerman & Martinez-Pons, 1990) reported significant differences in the usage of metacognitive strategies between intellectually gifted and average students. These differences were, however, rather small and they were not consistently obtained for all metacognitive strategies reported (cf. Alexander et al., 1995). Furthermore, correlations with learning performance scores were often not available or not presented. Slife et al. (1985) showed that the metacognitive functioning of students with learning disabilities was less adequate relative to regular students, although both groups were matched on intelligence and domain knowledge. Apparently, the last results do not support the intelligence model.

In a second, contrasting model intellectual ability and metacognitive skillfulness are regarded as entirely independent predictors of learning, i.e. as entirely separated toolboxes. Allon, Gutkin and Bruning (1994) reported low correlations between WISC-R intelligence and metacognition obtained retrospectively by questioning participants about their problem solving activities. Swanson (1990) claimed to obtain further support of the independency model for children performing two Piagetian tasks. His experimental design, however, which forced intelligence and metacognition to be orthogonal factors, does not permit the conclusion that both predictors are fully independent (see Veenman & Elshout, 1991). Indeed, follow-up studies (Maqsud, 1997; Swanson, Christie, & Rubadeau, 1993) showed that metacognition was only partially independent of intelligence.

The last model is a mixed one. According to this mixed model, metacognitive skillfulness is related to intellectual ability to a certain extent, but it also has a surplus value on top of intellectual ability for the prediction of learning. This mixed model has been corroborated by several studies, either with computer simulations in the domains of electricity, calorimetrics, statistics and behavioral psychology, with studying texts in the domains of law, geography and earth sciences, or with problem solving in the domains of math and thermodynamics (Elshout & Veenman, 1992; Elshout, Veenman, & van Hell, 1993; Veenman, 1993; Veenman, Beishuizen, & Niewold, 1997; Veenman & Elshout, 1991, 1999; Veenman, Elshout, & Busato, 1994; Veenman et al., 1997). In an overview of this research, Veenman (1999) showed that the variance accounted for in learning could be attributed uniquely to intellectual ability for 13%, could be attributed uniquely to metacognitive skillfulness for 16.3%, while both predictors shared another 17.2% of variance. Berger and Reid (1989) concluded from their study with mentally retarded individuals, high or low intelligent students with learning disabilities, and normal achieving adults that "IQ mediates metacognition, but does not explain it". Stankov (2000) also argued that metacognition is partly independent of fluid intelligence. Minnaert and Janssen (1999), on the other hand, could not decide between the independency and mixed model when predicting freshmen's academic performance. They, however, used a questionnaire (LERQ), which appears to be a less valid method for assessing metacognitive skillfulness (Veenman & van Hout-Wolters, 2001).

Some of the previously cited research reports do not fully permit a selection among the three alternative models. Either reports lacked a complete set of data (Allon et al. 1994; Shore & Dover, 1987; Span & Overtoom-Corsmit, 1986), or intellectual ability and metacognitive skillfulness were assessed as orthogonal factors (Slife et al., 1985; Swanson, 1990), or paradigmatic differences in the metacognition assessment-method made those reports incomparable (e.g. self-reports vs. thinking-aloud; Veenman & Van Hout-Wolters, 2001). Additionally, samples differed substantially between studies with regard to age and type of student. Despite those problems, most of the research reports with complete data sets support the mixed model.

1.4. Generality vs. domain-specificity of metacognitive skill

Another relevant issue is whether metacognitive skills are general and domain-independent, or domain-specific and knowledge related. The metacognitive skills of experts in a domain may well be integrated with domain-specific knowledge (Glaser & Chi, 1988), whereas the metacognitive skills of novices by definition lack such integration (Veenman, 1993; Veenman & Elshout, 1999). Therefore, the aforementioned three models pertain to relative novice learners. Moreover, the grain of analysis is of importance. The analyses of specific metacognitive activities, such as monitoring (general according to Schraw, Dunkle, Bendixen, & Roedel, 1995; Schraw & Nietfeld, 1998) or metacognitive memory accuracy (task specific according to Kelemen, Frost, & Weaver III, 2000), may yield contradictory information about the domain-independency of those activities. In order to elaborate on this grain-of-analysis problem, three closely related studies will be discussed in detail.

Glaser, Schauble, Raghavan, and Zeitz (1992) investigated the generality vs. domain-specificity of metacognitive activity during discovery learning. Twelve university freshmen passed through three different simulation environments representing microeconomics, electricity and the refraction of light. The underlying structures of these environments differed substantially from one another, varying from representing correlational to causal relations. Counting student activities from thinking-aloud protocols and computer traces revealed that the patterns of discovery behavior differed substantially across domains. Glaser et al. (1992) concluded that discovery behavior, which included metacognitive components, was mainly domain-specific.

This conclusion of Glaser et al. could be challenged as far as metacognitive skills are concerned. Although different learning environments may evoke different overt activities, these divergent activities may spring from similar metacognitive grounds. For instance, overt orienting behavior may differ for text studying (e.g. scanning headings and subheadings, getting grasp of the theme and the overall text structure), relative to math problem solving (e.g. comprehension of the problem statement, making a sketch representing the problem, selecting relevant data and goals). Such activities may even vary within individuals over tasks. The inclination of analyzing the task and its requirements prior to action taking, however, may be a person related characteristic; perhaps not so much a trait, but rather acquired behavior that proved to be effective to task performance in general (Veenman, 1993).

In a study of Veenman et al. (1997) highly similar to Glaser et al. (1992), 14 participants passed through three different simulation environments (calorimetry, statistics, and a fictitious Deton-lab for exploring the explosive power of unknown materials). Quality of metacognitive skillfulness, assessed through the analyses of thinking aloud protocols, proved to be a general, person-related characteristic across environments, rather than a task-specific phenomenon. Moreover, results supported the mixed model across environments. Recently, Veenman and Verheij (2001) corroborated these results while comparing technical students' problem-solving behavior on a paper-and-pencil model-construction task with discovery learning in Deton-lab. In other words, results of the latter two studies support the notion of metacognitive skillfulness as a person-related quality, surpassing tasks and domains.

1.5. Developmental perspective

Most of the aforementioned research pertained to university students. Apart from the inherent problem of restriction of range in terms of intellectual ability, such a sample imposes a serious problem for generalizing the mixed model to younger age groups in particular. Perhaps, developmental processes affect the relation between intelligence and metacognition.

Flavell (1992) related his concept of metacognition (see above) to Piaget's developmental stage of formal-operational thinking (Inhelder & Piaget, 1958). At this stage children are capable of hypothetico-deductive reasoning, which requires metacogitive control. Flavell indicated that Piaget would not expect metacognition to show up before the level of formal-operational thinking has been reached as he argued that: "young children's egocentrism prevents them from being able to introspect or treat their own thought processes as an object of thought" (Flavell, 1992:118; Inhelder & Piaget, 1958). Flavell further adhered to Piaget's theory by postulating an early developmental level of 'proto-metacognition', at which level children do acknowledge that different people may see different things, although they cannot handle the various perspectives people may take. Therefore, metacognitive awareness may arise at the age of 4–6 years as an inclination that something is wrong (Demetriou & Efklides, 1990; Istomina, 1975; Kluwe, 1987; Kuhn, 1999). The further development of metacognitive knowledge and skills occurs in the years thereafter (Alexander et al., 1995; Armbruster, Echols, & Brown, 1982; Campione, Brown, & Ferrara, 1982; Cross & Paris, 1988; Kuhn, 1999).

A relevant research issue, then, is whether the development of metacognitive skills is intelligence-related or relatively intelligence-independent according to the mixed model. Alexander et al. (1995) formulated three hypotheses with regard to this developmental process. The ceiling hypothesis asserts that, initially, metacognitive development is associated with intelligence, but that such intelligence effects diminish over time. For instance, Schneider and Pressley (1997) argued that in the course of cognitive development, the influence of constraints of the information processing system, like memory capacity and processing, is gradually reduced. As a consequence, more resources become available for metacognitive processes. Apart from that, knowledge and experience increasingly become relevant sources of influence on the quality of the metacognitive process. Both arguments are in line with the ceiling hypothesis. The acceleration hypothesis, on the other hand, predicts that intelligence effects on metacognitive skill development become larger with age. Borkowski and Peck (1986) argued that the development of intellectual skills and the development of metacognitive skills are mutually enhanced by one another. Finally, the monotonic development hypothesis emphasizes the continuous growth of metacognitive skillfulness with age, alongside intelligence. In their literature review, Alexander et al. found support for monotonic development of metacognitive knowledge. The evidence for metacognitive skill development, however, was inconclusive. Moreover, their review study did not address to what extent precisely metacognitive skill development was associated with intelligence. As a first research issue of the present study, it is hypothesized that the mixed model pertains to all age levels.

The second research issue concerns the generalizability of metacognitive skills across domains during the development of those skills. This issue refers to the transferability of metacognitive skills across domains (Baker, 1994). In the present study, tasks with the same format but representing different domains are administered to participants of several age groups. It is hypothesized that metacognitive skills are relatively domain-independent for all age groups.

Both research issues may have important implications for the malleability of metacognitive-skill development through instruction. Firstly, validation of the mixed model would require educational researchers to disentangle the various effect sources of training programs for raising intelligence. Metacognitive skill training is an essential part of such programs, perhaps the only effective one (Perkins & Grotzer, 1997). Secondly, the generalizability of metacognitive skills opens opportunities for training those skills, initially within but eventually across learning contexts through ‘high road’ transfer (Salomon & Perkins, 1989).

2. Method

2.1. Participants

The four groups of participants in the study consisted of 28 fourth-graders (mean age 9.5 years) and 28 sixth-graders (mean age 11.6 years) from an elementary school in the urban area of Amsterdam, 30 eighth-graders (mean age 14.1) from a secondary school in the urban area of The Hague, and 27 university students from the Leiden Faculty of Social Sciences (mean age 22.5 years). The youngest group was selected because the onset of metacognitive-skill development occurs around that age. Other age groups represented different points in a developmental track (with the exclusion of age groups from 16 to 19 years for practical reasons of school exams taken at different levels of secondary education in the Netherlands). Both the elementary and secondary schools were chosen on teaching a regular curriculum, and pupils with severe learning problems were excluded from the sample. Social-economic status might be estimated as slightly above average, but distribution of sex was balanced. Parents were informed by a letter and consented to the participation of their children. As Dutch universities are obliged to admit all students who passed the highest level of secondary education without entry selection on GPA, the participating university students covered a relatively broad intelligence range (cf. Veenman, 1993). Female students outnumbered males twice, in line with the distribution of sex for social sciences students. All university students participated voluntarily. For one sixth-grader and one university student intelligence scores were missing.

2.2. Intellectual ability

Because it was expected that inductive learning tasks would draw heavily on both general and spatial reasoning, for each factor two tests of intellectual ability were included. For the general reasoning factor a Number Series Test (Elshout, 1976) assessed inductive reasoning (Carroll, 1993), and a Concrete Syllogisms Test tapped deductive reasoning (Carroll, 1993). For the spatial reasoning factor, both the Hidden Figures Test (Flanagan, 1951), and the Spatial Insight subtest of the Differential Aptitude Test (Evers & Lucassen, 1983) measured flexibility of closure (Carroll, 1993). Moreover, numerical ability was covered with a Math Word Problems Test (adapted from Elshout, 1976, parallel to a WISC-R subtest). In order to ensure that these tests could be used across age groups, test instructions were simplified and simpler items were added to some tests. Moreover, the Concrete Syllogisms test was developed because an Abstract Syllogisms task (Conclusions; Elshout, 1976) contains relational symbols (“<” or “>”) younger children are unfamiliar with. Both tests contained items with identical underlying structures. For instance, an Abstract syllogism test item (Given $A < B < C$, what is the relation between A and C? Answer: $A < C$) was replaced with a corresponding item in the Concrete Syllogisms task (Peter is smaller than Mark and Mark is smaller than John, then Peter is … John; Answer: smaller). Although the Concrete Syllogism test was a verbal version of the symbolic Abstract Syllogism test, thus introducing a verbal factor, both tests appeared to correlate fairly high for eighth-graders and university students ($r = 0.74$, $p < 0.01$). An overall score for intellectual ability was obtained by calculating the unweighted mean of the standardized scores on all five subtests (Cronbach’s alpha = 0.83). Intelligence scores were not converted to age-adjusted IQ scores because they were intended to measure intellectual growth parallel to metacognitive development.

2.3. Tasks

All participants performed four computerized inductive learning tasks, two tasks in the domain of biology (the plant-growing and food tasks) and two in the domain of geography (the otter and ageing tasks). The tasks were implemented in a computerized authoring environment (FILE, Wilhelm, Beishuizen, & van Rijn, in press). The content of each task was chosen from the biology or the geography curriculum in primary and secondary education. Further differences in task familiarity among age groups were avoided (cf. Alexander et al., 1995). For instance, the task format was fairly new to all participants. Also, the problems and the variables presented in each task were plausible to participants from all age groups, thus reducing the impact of age-related differences in prior knowledge.[1] In each task, five independent variables with discrete levels (either two or three levels) could be varied and their effect on the dependent variable could be inspected. The model underlying the relations between the independent and the dependent variables was identical in each task; two independent variables interacted with one another, one variable had a non-linear effect, and two variables were irrelevant. Each task model corresponded to possible real-life phenomena. Fig. 1 shows the interface of the plant-growing task as an example. The task was to find out how different inde-

---

[1] In order to control for differences in prior knowledge, a questionnaire assessing prior knowledge of biology and geography was administered along with the intelligence tests. This questionnaire also specifically addressed the task domains of plant growing, food, otters, and ageing. Although age differences in prior knowledge were found and although prior knowledge was correlated to both metacognition and learning performance on the four inductive learning tasks across age groups, these correlations were only moderate ($0.25 \leq r \leq 0.36$). Within age group correlations were even lower ($-0.34 \leq r \leq 0.29$). Correlation of prior knowledge with intellectual ability varied strongly ($0.05 \leq r \leq 0.58$), also within age groups ($-0.35 \leq r \leq 0.46$). Partialing out prior knowledge, however, did not substantially affect the correlations amongst intellectual ability, metacognition, and learning performance.

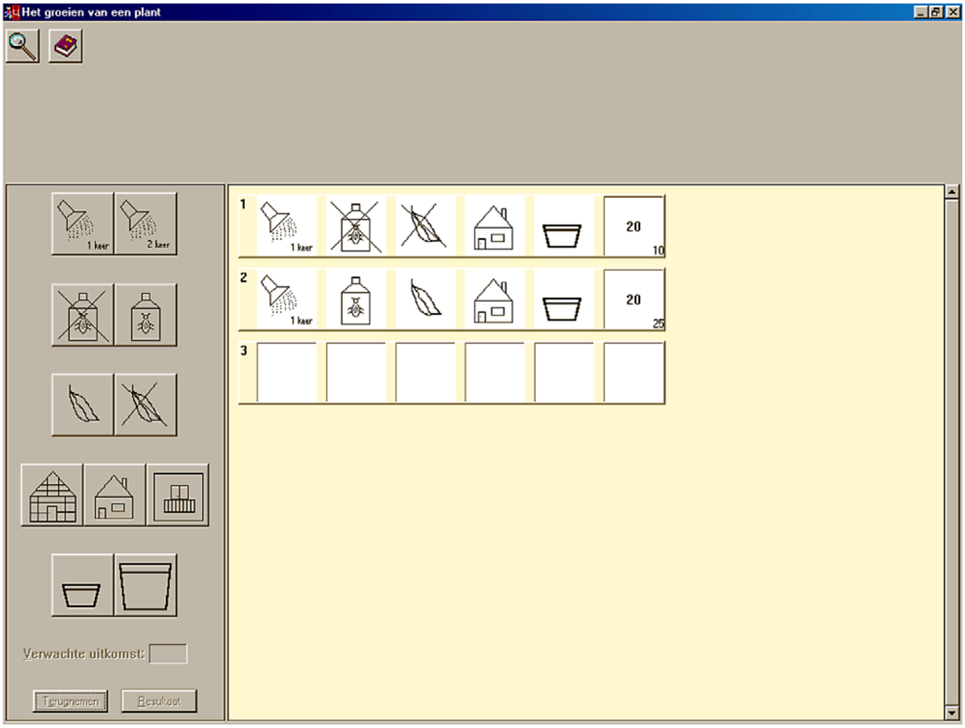

Fig. 1. Interface of the plant-growing task.

pendent variables affected the height of plant growth. Independent variables were: (1) giving water, either once or twice a week; (2) using an insecticide or not; (3) putting dead plant leaves in the flowerpot or not; (4) placing the plant either indoors, on a balcony, or in a greenhouse; and (5) size of the flower pot, either large or small. Distinct levels of plant growth as a dependent variable were 5, 10, 15, 20 and 25 cm high. Variable 4 had a non-linear effect, meaning that growing the plant indoors resulted in five centimeters less growth, relative to a balcony or greenhouse. Variable 2 and 3 did not affect plant growth at all. Variable 1 and 5 interacted, as giving water once or twice a week did not matter for a large pot, but it did matter when a small flowerpot was used. In that case, giving water twice a week would reduce plant growth, while giving water once a week would increase growth, relative to growth in the large flowerpot.

During the food task participants had to find out how eating and drinking habits affected the health status of an imaginary person, called Hans. Independent variables were snacks (i.e. fat consumption), carbohydrates, alcohol, albumen, and supplementary vitamins. In the otter task, the relevance of factors affecting the extinction of otters in the Netherlands had to be found out. Independent variables were extra food provision or not, environmental pollution, natural habitat, media exposure, and enclosing of otter areas to the public or not. In the population-ageing task independent variables that could affect the ageing rate of a population were state of the economy, quality of the educational system, means of living, climate, and general safety. In all cases two variables interacted, one variable had a non-linear effect, and two variables were irrelevant to the dependent variable.

2.4. Learning performance

After completion of each task, the experimenter asked the participants to describe the effects of all independent variables in detail (e.g. "What difference does watering once or twice a week make for the plant growth?"). Learning performance was assessed on the basis of answers to these standardized questions. In fact, nine statements covered all effects for each task. The plant-growing task, for instance, could result in five main effects: Growing the plant indoors made the plant grow smaller relative to growing the plant in a greenhouse (1), growing the plant indoors made the plant grow smaller relative to growing the plant on a balcony (2), and growing the plant on a balcony or in a greenhouse did not make a difference (3). Putting dead leaves in the flowerpot or not (4), or using an insecticide or not (5) did not make a difference. Four correct statements referred to the interaction: watering once or twice a week did not make a difference when a large flowerpot was used (6), watering once or twice a week did make a difference when a small flowerpot was used (7), in a small flowerpot, the plant grew smaller relative to a large pot when water is given twice a week (8), and in a small flowerpot, the plant grew higher relative to a large pot when water is given once a week (9). Two points were awarded when a response to an interview question matched a correct statement. No points were given whenever a statement was incorrect or absent. When a participant stated that an interacting variable had a main effect (e.g. "A big flowerpot is better than a small one"), one point was awarded. Measures of learning performance consisted of the sum of all points on a particular task, with scores ranging from 0 to 18 points. Analyses of learning performance on the same tasks in a parallel sample ($N = 20$) by two independent judges revealed that judgments converged for 85%. Therefore, only one of the judges assessed learning performance from the interview protocols.

2.5. Metacognitive skillfulness

For each task separately, logfiles were automatically scored by the computer on two metacognition measures. The mean number of variables changed per experiment (VOTAT; Chen & Klahr, 1999; Tschirgi, 1980) was obtained for each participant as a negative indicator of metacognitive skillfulness. Varying more than one variable at a time represents poor systematical behavior (Veenman et al., 1997) and a lack of experimental control (Schauble, Raghavan, Glaser, & Reiner, 1991; Shute, Glaser, & Raghavan, 1989). The frequency of scrolling back to earlier experiments, on the other hand, was assessed as a positive indicator of metacognitive skillfulness. Scrolling would indicate a participant's intention to check earlier experimental configurations or to relate the outcomes of experiments. Scores on both measures were standardized and the sign of the negative indicator was inverted.

As was argued before, metacognitive skills appear to be highly interdependent. Good orientation leads to good planning and systematical behavior, which in turn lead to more evaluative control. Therefore, a small set of measurements may adequately represent a broader range of metacognitive skills (Veenman, Elshout, & Groen, 1993). This assumption, however, needed explicit verification in order to validate the logfile-measures of metacognition. As protocol analyses are extremely time-consuming, only a limited number of protocols could be processed. Two other 'blind' judges rated 10% of the plant-growing-task protocols and 5% of the otter-task protocols from tape on the quality of metacognitive skillfulness, for which the criteria were adapted from Veenman (1993; Veenman et al., 1997). Protocols were judged on the quality of (1) orientation (elaborateness of hypotheses generated before each experiment); (2) systematical behavior (planning a sequence of experiments, and avoiding unsystematic behavior such as varying two independent variables between subsequent experiments); (3) evaluation (detection and correction of mistakes); and (4) elaboration (drawing conclusions, relating outcomes of experiments, generating explanations, and recapitulating). It must be emphasized that protocols were judged on the quality of metacognitive activities, not on the quality of information these activities produced. For instance, subjects generating well-considered, though incorrect hypotheses scored high on orientation. Similarly, subjects drawing elaborated, yet incorrect conclusions scored high on elaboration. Moreover, protocol analyses obviously did not include the post-experimental interviews for assessing learning performance. The four categories were rated on a five-point scale (ranging from 0 to 4). The judges performed the analyses together, arguing until agreement was reached. This method of protocol analysis lacks the assessment of an interjudge reliability but enables the judges to scrutinize their judgments mutually, which enhances reliability (Veenman & Elshout, 1995). Sum scores were calculated over the four categories (Cronbach's alphas > 0.95).

2.6. Procedure

The paper-and-pencil intelligence tests were administered during class prior to individual test sessions, which took place in a quiet room at school or at the university. Each participant attended two sessions within approximately 2 weeks. For those who missed a session, care was taken that all of them would catch up with that session later on. The order of domains was counterbalanced over individual sessions. The order of tasks within each session was always the same, i.e. for the biology domain first the plant-growing task and then the food task was presented, whereas for the geography domain the otter task was followed by the ageing task.

The first individual session started with an instruction of the computer interface using another FILE configuration of an everyday-life task (the Peter task; see Wilhelm et al., in press). Participants were instructed how to place levels of each independent variable in an array on the screen by clicking on corresponding icons (see also Fig. 1), how to enter a prediction, and how to obtain the result. This interface instruction included the use of the scroll button for paging through previous experiments. At the onset of the second session participants were asked to show the experimenter how the interface could be used. When they could not demonstrate a particular function, the experimenter reminded them of this function.

Each task started with a short story explaining the purpose of the task. In order to make clear what was expected from them, participants were told they would be asked the following question for each independent variable at the end of the task: "What difference do you think the (independent variable) makes for the (dependent variable)?" During task performance, computer logfiles registered the participant's activities for obtaining metacognition scores. Furthermore, participants were asked to think aloud. Some of these tape-recorded protocols were used for establishing quality ratings of metacognitive skillfulness afterwards.

Participants were allowed to work on each task for 35 min, but they could stop earlier if they wanted to after conducting at least 10 experiments. Pilot studies showed that some participants wanted to stop after they had checked the main effects, requiring at least seven experiments. When this occurred, participants were urged to conduct more experiments through a general instruction ("Could you make some more rows?"). When participants gave notice that they finished the task, they were asked what effect each independent variable had on the outcome. This structured interview was used for determining the learning outcomes on each task.

3. Results

3.1. Intellectual ability

ANOVA on the intelligence scores revealed a significant effect of age ($F(3, 107) = 49.52$, $p < 0.001$). As expected, intelligence increased with age. Mean z-scores were -1.22 ($SD = 0.67$), -0.43 ($SD = 0.63$), 0.34 (SD = 0.96), and 1.36 (SD = 0.93) for ascending age groups.

3.2. Metacognition

ANOVA on the reversed z-scores of VOTAT also showed a significant effect of age ($F(3, 109) = 22.28$, $p < 0.001$), with mean scores of -3.06 (SD = 3.71), -1.00 (SD = 2.50), 1.55 (SD = 2.98), and 2.40 (SD = 1.31) for ascending age groups. Similarly, ANOVA on the z-scores of scrolling back revealed a significant effect of age ($F(3, 113) = 18.49$, $p < 0.001$), with mean scores of -2.02 (SD = 1.55), -0.98 (SD = 1.90), -0.30 (SD = 3.25), and 3.19 (SD = 3.69) for ascending age groups. No significant within task effects were found. Both metacognition measures were positively correlated to learning performance ($0.32 \leq r \leq 0.64$). Therefore, VOTAT and scrolling measures for each task were aggregated.

3.3. Concurrent validity of metacognition scores

Quality of metacognitive skillfulness, judged from thinking-aloud protocols of the plant-growing task, correlated 0.85 ($p < 0.01$, $N = 12$) with the corresponding logfile scores. For the otter task a correlation between both measures of 0.84 ($p < 0.05$, $N = 6$) was obtained.

3.4. Generality of metacognition across domains

A Principal Components Analysis was performed on the metacognition scores for each task. As two separate measures of metacognition were available for each domain (otters and aging for geography; plant-growing and food for biology; $N = 113$), PCA would allow for the extraction of domain-specific components. All metacognition measures appeared to load strongly on the first component extracted (see Table 1). Therefore, this component may be interpreted as representing general metacognition. The second component contrasted metacognition for geography with metacognition for the food task, although the eigenvalue and variance proportion accounted for were low (see Table 1). Separate PCAs for each age group yielded similar patterns of component loadings.

A repeated measures ANOVA on the raw metacognition scores with age groups as between factor and tasks as within factor only revealed a significant main effect of age groups ($F(3, 109) = 38.60$, $p < 0.001$), indicating that metacognition increased with age (see Table 2). Neither the main task effect ($F(3, 327) = 0.06$), nor the interaction ($F(9, 327) = 1.62$) appeared to be significant. Consequently, metacognition scores were aggregated over the four tasks to generate a general metacognition score for further analyses (Cronbach's alpha = 0.92).

Table 1
Unrotated component matrix for metacognitive skillfulness

	Component 1	Component 2
Eigenvalue	3.23	0.32
Variance proportion	0.81	0.08
Metacognition otter (geography)	0.90	−0.28
Metacognition ageing (geography)	0.92	−0.20
Metacognition plant grow (biology)	0.91	0.05
Metacognition food (biology)	0.88	0.44

Table 2
Means (SD) for metacognition scores

	Otter	Ageing	Plant	Food	Total tasks
Grade 4	−1.19 (1.08)	−1.22 (1.29)	−1.29 (1.42)	−1.38 (1.27)	−1.27 (1.05)
Grade 6	−0.75 (0.98)	−0.62 (1.26)	−0.31 (0.92)	−0.30 (0.96)	−0.49 (0.75)
Grade 8	0.35 (1.46)	0.41 (1.26)	0.11 (1.36)	0.38 (1.12)	0.31 (1.16)
Students	1.54 (0.90)	1.47 (1.15)	1.42 (0.96)	1.16 (1.08)	1.40 (0.79)
All	−0.02 (1.53)	0.00 (1.59)	−0.03 (1.52)	−0.04 (1.44)	

Table 3
Means (SD) for performance scores

	Otter	Ageing	Plant	Food	Total tasks
Grade 4	7.50 (2.52)	7.43 (3.27)	7.54 (3.10)	8.50 (3.00)	7.74 (2.07)
Grade 6	9.00 (3.03)	9.86 (2.74)	8.46 (3.47)	10.14 (3.20)	9.37 (2.04)
Grade 8	10.40 (2.50)	11.50 (2.75)	10.67 (3.56)	11.27 (2.27)	10.96 (2.27)
Students	10.67 (2.84)	12.37 (3.31)	11.82 (3.97)	12.63 (3.39)	11.87 (2.33)
All	9.40 (2.97)	10.29 (3.52)	9.62 (3.88)	10.63 (3.30)	

3.5. *Learning performance*

A repeated measures ANOVA on the performance scores with age groups as between factor and tasks as within factor revealed a significant main effect of age groups ($F(3,109) = 19.90$, $p < 0.001$), indicating that performance increased with age (see Table 3). The main task effect ($F(3,327) = 5.73$, $p < 0.001$) showed that performance varied over tasks (see Table 3). The interaction effect, however, was not significant ($F(9,327) = 0.78$), which allowed performance scores to be aggregated over the four tasks for further analyses (Cronbach's alpha = 0.77).

3.6. *Correlational analyses*

Correlations among intellectual ability, overall metacognition, and overall learning performance were calculated for each age group separately, as well as for the entire sample (see Table 4).[2] Next, semipartial correlations (Nunnally, 1967) were calculated by partialing intellectual ability from the correlations between metacognition and learning performance. These semipartial correlations are indications of the unique contribution of metacognition to learning performance, independent of intellectual ability.

For the youngest age group (grade 4), the correlation between intellectual ability and learning performance appeared to be low (see Table 4), although this correlation did not deviate significantly from similar correlations in the other age groups (Fisher-z ratios < 1.25; Guilford, 1965). For the first three age groups (grade 4 up to 8), the semipartial correlations indicated that metacognition significantly contributed to learning performance on top of intellectual ability. In the eldest age group of university students, however, metacognition correlated rather poorly with learning performance, which was also reflected in a low semipartial correlation (see Table 4). The correlation between metacognition and learning performance was significantly lower for students relative to fourth-graders and eighth-graders (Fischer-z ratios > 2.00, $p < 0.05$), but not relative to sixth-graders (Fischer z ratio = 0.77).

---

[2] Correlations were also calculated on factor scores obtained for both metacognition components. Correlations of scores on the first, general metacognition component with intellectual ability and learning performance were virtually identical to those of the raw metacognition scores. Correlations of scores on the second component with intellectual ability and learning performance were close to zero.

Table 4
Correlations among intellectual ability, metacognition, and performance

	Intellectual ability	Metacognition	Semi-part	*N*
Performance grade 4	0.20	$0.74^{\dagger}$	$0.73^{\dagger}$	28
Metacognition grade 4	$0.49^{\dagger}$			
Performance grade 6	$0.51^{\dagger}$	$0.48^{\dagger}$	0.30^{*}	27
Metacognition grade 6	0.40^{*}			
Performance grade 8	0.39^{*}	$0.70^{\dagger}$	$0.58^{\dagger}$	30
Metacognition grade 8	$0.43^{\dagger}$			
Performance students	$0.48^{\dagger}$	0.29	0.11	26
Metacognition students	0.42^{*}			
Performance All	$0.66^{\dagger}$	$0.74^{\dagger}$	$0.38^{\dagger}$	111
Metacognition All	$0.75^{\dagger}$			

$^{*}p < 0.05$; $^{\dagger}p < 0.01$. Semi-part means semi-partial correlation with intellectual ability partialed from the correlation between metacognition and performance.

Overall analyses aggregating all age groups, revealed high correlations amongst intellectual ability, metacognition, and learning performance (see Table 4). The semipartial correlation indicated that metacognition significantly contributed to learning performance on top of intellectual ability. Using regression-analytic techniques (Pedhazur, 1982; Veenman, 1993; Veenman & Elshout, 1999) the unique and shared proportions of variance accounted for in learning performance by intellectual ability and metacognition were estimated. The variance accounted for uniquely by intellectual ability was 2.4%, the variance accounted for uniquely by metacognition was 14.4%, and the variance shared by both predictors was 40.8%. Another 42.4% of variance in learning performance was unaccounted for. These results indicate that, although both predictors have much in common, metacognition contributes to learning performance on top of intellectual ability.

4. Discussion

Results show that metacognitive skillfulness is a general, person-related characteristic across age groups, rather than a domain-specific feature. Although the domains of biology and geography may be related to a certain extent, the content of these tasks varied substantially. The plant-growing task likely elicits everyday associations different from, for instance, the ageing task, which concerned economic and societal effects on age distribution in a population. Therefore, the domain surpassing nature of metacognitive skillfulness, which so far pertained to university students only (Veenman et al., 1997; Veenman & Verheij, 2001), may be generalized to younger age groups. It should be emphasized, however, that the present results only allow for such a generalization over varying task contents, not over task formats or tasks settings, as each task involved inductive learning in a computerized environment. Research in progress now investigates the generality of metacognitive skills over different task formats of text studying vs. problem solving for different age groups (van Hout-Wolters & Veenman, 2001).

Results for separate groups show that metacognitive skillfulness is a primary factor predicting learning performance in grade 4 (accounting for at least 54% of variance), that metacognitive skillfulness is a relevant predictor of performance on top of intellectual ability in grade 8 and 9 (accounting for another 9 to 34% of variance in performance), and that metacognitive skillfulness failed to be a distinct factor in the performance of the eldest age group. Overall analyses including all age groups, however, show that, although both predictors were age-related, metacognitive skillfulness had a predictive value for learning performance on top of intellectual ability (accounting for an additional 14% of variance). This result is even more noteworthy because intelligence measures were purposefully not adjusted to age levels of participants (i.e. transformed to IQ), thus maximizing the potential impact of intellectual ability on the development of metacognitive skillfulness. Although intellectual ability mediates the development of metacognitive skills, the overall results fit the mixed model.

For grade 4 two issues need to be addressed: (1) the low predictive value of intellectual ability; and (2) the relatively high impact of metacognition. According to Elshout (1987) and Raaheim (1988), a curvilinear relationship exists between the impact of intelligence and performance. Intelligence is assumed to play little part in very familiar task situations. In other words, routine kills intelligence. As the familiarity with a task situation declines and task complexity consequently advances, intellectual ability is called upon increasingly. At a certain point of task complexity one can optimally profit from one's intellectual repertoire. This point is called the threshold of problematicity (Elshout, 1987), at which verge one is still capable of managing a relatively unfamiliar problem. If the task complexity, however, moves beyond this threshold, the impact of intellectual ability gradually diminishes. Extreme complex learning tasks may even annihilate the impact of intelligence because subjects cannot see the forest for the trees. The individual position of the problematicity threshold depends on one's initial knowledge or expertise level, one's intellectual ability, and one's metacognitive skillfulness (Veenman & Elshout, 1999). Therefore, the position of the threshold may vary from person to person, and from task to task. For instance, during the acquisition of expertise the individual threshold will gradually shift to a higher level of task complexity.

Most likely, the tasks in this experiment were too complex for fourth-graders in order to have their intellectual repertoire play a relevant role. For instance, from the interview data it appeared that fourth-graders hardly ever discovered any interaction effect. Moreover, inspection of the protocols revealed that these young participants might have suffered from a confirmation bias (Dunbar & Klahr, 1989; Wason, 1977), as they held on to real-life conclusions inconsistent with the empirical data. For instance, some of them would not believe that using a small pot was not always detrimental to plant growth, despite the empirical evidence. The manifold combinations of experimental manipulations and outcomes might have overwhelmed them. Metacognitive, rather than intellectual skills appeared to be helpful in gaining control over this complex task. Indeed, previous research confirmed that whenever learners are confronted with a highly difficult or unfamiliar task, only metacognitive skill contributes to the initial learning process (Veenman, Kok, &

Kuilenburg, 2001; Veenman, Prins, & Elshout, 2002). Metaphorically speaking, if you are stumbling around in the dark, intelligence is of no use but it may help if you go around carefully (e.g. doing things step-by-step).

In the eldest group of university students, on the other hand, metacognition correlated surprisingly low with learning performance. It should be taken into account that the mixed model has been confirmed for students using a variety of tasks different from the present ones. A floor effect in VOTAT was responsible for a restriction of range in the metacognition scores of students in the present study. Correction for this restriction of range would yield a much higher adjusted correlation (Gulliksen, 1961). This floor effect, however, did not result from a relative cognitive simplicity of the tasks for the eldest age group. In that case one would have expected a lower correlation of intelligence with performance (Elshout, 1987; Raaheim, 1988). Furthermore, students overtly expressed in their thinking-aloud protocols that they experienced the tasks as being difficult. These tasks were probably less metacognitively demanding for students, allowing them to be rather homogeneous in systematically varying a limited number of variable levels. Accordingly, a distinction between intellectual and metacognitive complexity could be postulated (cf. Crawford, 1991). Intellectual complexity depends on the number and nature of relations to be found, which affect learning performance according to the inverted U-shaped curve of Elshout and Raaheim. Metacognitive complexity, on the other hand, depends on the number of concurrent strategic alternatives. The learning impact curve for metacognition asymptotically approaches zero, as complexity gets low. As a direction for future research, both forms of complexity could be varied independent of one another. If complexity variation shows differential effects on the learning impact of intelligence and metacognition, this would substantiate the distinction between both forms of complexity. Accordingly, a relevant educational implication would be that metacognitive complexity could be compensated for by metacognitive instruction in an early stage of the learning process, whereas intellectual ability and intellectual complexity could be balanced out by adjusting the number and nature of relations during inductive reasoning (i.e. tuning to the individual threshold of problematicity).

From a developmental perspective, the increase of metacognitive skillfulness with age, from childhood to early adulthood, is significant. Clearly, metacognitive skillfulness represents an acquired repertoire of general skills for managing problem-solving and learning situations. Moreover, correlations between metacognition and intelligence appear to be stable across age groups in line with the monotonic development hypothesis of Alexander et al. (1995). The age-related increase of metacognition, however, is not exclusively determined by intellectual growth. It appears that metacognition develops partly independent of intelligence, albeit to a limited extent. Metacognitive development is associated with, but not intrinsically part of the maturing cognitive toolbox. Agents responsible for this metacognitive development may be found at home (Grolnick & Ryan, 1989; Kontos, 1983) and, more specifically, in the schools (Baker, 1994). It has been established, for instance, that metacognitive training programs may be advantageous to less intelligent pupils or students (Alexander et al., 1995; Brown & Palincsar, 1989; Campione et al., 1982; Cross & Paris, 1988; Veenman et al., 1994).

An interesting research issue, then, would be to what extent concurrent and 'mindful' training of metacognitive skills in multiple domains, directed at 'bridging' and synchronizing the application of those skills across domains, could enhance transfer or generalizability of those skills (Salomon & Perkins, 1989). As an educational implication of the generality of metacognitive skills and of the mixed model, concurrent metacognitive-skill instruction in various domains may be expected to scaffold and strengthen the transferability of metacognitive skills, even in less intelligent pupils or students. Teachers of various disciplines, however, have to come to terms with one another in order to attune metacognitive instructions.

References

Alexander, J. M., Carr, M., & Schwanenflugel, P. J. (1995). Development of metacognition in gifted children: directions for future research. *Developmental Review*, *15*, 1–37.

Allon, M., Gutkin, T. B., & Bruning, R. (1994). The relation between metacognition and intelligence in normal adolescents: Some tentative but surprising findings. *Psychology in the Schools*, *31*, 93–97.

Armbruster, B. B., Echols, C. H., & Brown, A. L. (1982). The role of metacognition in reading to learn: a developmental perspective. *The Volta Review*, *84*, 45–56.

Baker, L. (1994). Fostering metacognitive development. In H. W. Reese (Ed.), *Advances in child development and behavior, vol. 25* (pp. 201–239). San Diego: Academic Press.

Berger, R. S., & Reid, D. K. (1989). Differences that make a difference: Comparisons of metacomponential functioning and knowledge base among groups of high and low IQ learning disabled, mildly mentally retarded, and normal achieving subjects. *Journal of Learning Disabilities*, *22*, 422–429.

Borkowski, J. G., & Peck, V. A. (1986). Causes and consequences of metamemory in gifted children. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 182–200). New York: Cambridge University Press.

Brody, E. B. (1992). *Intelligence, Nature, determinants, and consequences, 2nd ed.* New York: Academic Press.

Brown, A. L. (1978). Knowing when, where, and how to remember: a problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology, vol. 1* (pp. 77–165). Hillsdale, NJ: Erlbaum.

Brown, A. L., & DeLoache, J. S. (1978). Skills, plans, and self-regulation. In R. S. Siegel (Ed.), *Children's thinking: What develops?* (pp. 3–35). Hillsdale, NJ: Erlbaum.

Brown, A. L., & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing, learning, and instruction. Essays in honor of Robert Glaser* (pp. 393–451). Hillsdale NJ: Erlbaum.

Campione, J. C., Brown, A. L., & Ferrara, R. A. (1982). Mental retardation and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 392–490). Cambridge: Cambridge University Press.

Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Chen, Z., & Klahr, D. (1999). All other things being equal: acquisition and transfer of the control of variables strategy. *Child Development*, *70*, 1098–1120.

Cheng, P. (1993). Metacognition and giftedness: the state of the relationship. *Gifted Child Quarterly*, *37*, 105–112.

Crawford, J. D. (1991). Intelligence, task complexity, and the distinction between automatic and effortful mental processing. In H. A. H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 119–144). Hillsdale, NJ: Erlbaum.

Cross, D. R., & Paris, S. G. (1988). Developmental en instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology*, *80*, 131–142.

Davidson, J. E., Deuser, R., & Sternberg, R. J. (1994). The role of metacognition in problem solving. In J. Metcalfe, & A. P. Shimamura (Eds.), *Metacognition* (pp. 207–226). Cambridge: MIT Press.

Demetriou, A., & Efklides, A. (1990). The objective and subjective structure of problem-solving abilities: metacognitive awareness from early adolescence to middle age. In H. Mandl, E. de Corte, S. N. Bennett, & H. F. Friedrich (Eds.), *Learning and instruction in an international context. Volume 2.1. Social and cognitive aspects of learning and instruction* (pp. 161–179). Oxford: Pergamon Press.

Dunbar, K., & Klahr, D. (1989). Developmental differences in scientific discovery processes. In D. Klahr, & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 109–143). Hillsdale, NJ: Erlbaum.

Elshout, J. J. (1976). *Karakteristieke moeilijkheden in het denken [Characteristic difficulties in thinking]*. Amsterdam: University of Amsterdam.

Elshout, J. J. (1983). Is measuring intelligence still useful?. In S. B. Anderson, & J. S. Helmick (Eds.), *On educational testing* (pp. 45–56). San Francisco: Jossey-Bass.

Elshout, J. J. (1987). Problem solving and education. In E. de Corte, H. Lodewijks, R. Parmentier, & P. Span (Eds.), *Learning and instruction* (pp. 259–273). Oxford: Pergamon Books Ltd (Leuven: University Press).

Elshout, J. J., & Veenman, M. V. J. (1992). Relation between intellectual ability and working method as predictors of learning. *Journal of Educational Research*, *85*, 134–143.

Elshout, J. J., Veenman, M. V. J., & Van Hell, J. G. (1993). Using the computer as a help tool during learning by doing. *Computers & Education*, *21*, 115–122.

Evers, A., & Lucassen, W. (1983). *Differentiële Aanleg Testserie*. Lisse: Swets & Zeitlinger.

Flanagan, J. C. (1951–1956). *Flanagan aptitude classification tests*. SRA.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231–235). Hillsdale, NJ: Erlbaum.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *American Psychologist*, *34*, 906–911.

Flavell, J. H. (1992). Perspectives on perspective taking. In H. Beilin, & P. Pufall (Eds.), *Piaget's theory: prospects and possibilities* (pp. 107–141). Hillsdale, NJ: Erlbaum.

Glaser, R. (1990). The reemergence of learning theory within instructional research. *American Psychologist*, *45*, 29–39.

Glaser, R., & Chi, M. T. H. (1988). Overview. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. xv–xxviii). Hillsdale, NJ: Erlbaum.

Glaser, R., Schauble, L., Raghavan, K., & Zeitz, C. (1992). Scientific reasoning across different domains. In E. de Corte, M. C. Linn, H. Mandl, & L. Verschaffel (Eds.), *Computer-based learning environments and problem solving, NATO ASI series F, Vol. 84* (pp. 345–371). Heidelberg: Springer Verlag.

Grolnick, W. S., & Ryan, R. M. (1989). Parent styles associated with children's self-regulation and competence in school. *Journal of Educational Psychology*, *81*, 143–154.

Guilford, J. P. (1965). *Fundamental statistics in psychology and education*. New York: McGraw-Hill.

Gulliksen, H. (1961). *Theory of mental tests*. New York: Wiley.

Hannah, C. L., & Shore, B. M. (1995). Metacognition and high intellectual ability: Insights from the study of learning-disabled gifted students. *Gifted Child Quarterly*, *39*, 95–109.

Humphreys, L. G. (1968). The fleeting nature of the prediction of college academic success. *Journal of Educational Psychology*, *59*, 375–380.

Humphreys, L. G. (1989). Intelligence: Three kinds of instability and their consequences for policy. In R. L. Linn (Ed.), *Intelligence* (pp. 193–216). Urbana: University of Illinois Press.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. London: Routledge & Kegan.

Istomina, Z. M. (1975). The development of voluntary memory in children of preschool age. *Soviet Psychology*, *13*, 5–64.

Kelemen, W. L., Frost, P. J., & Weaver III, C. A. (2000). Individual differences in metacognition: evidence against a general metacognitive ability. *Memory and Cognition*, *28*, 92–107.

Kluwe, R. H. (1987). Executive decisions and regulation of problem solving behavior. In F. E. Weinert, & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 31–64). Hillsdale, NJ: Erlbaum.

Kontos, S. (1983). Adult–child interaction and the origins of metacognition. *Journal of Educational Research*, *77*, 43–54.

Kuhn, D. (1999). Metacognitive development. In L. Balter, & C. S. Tamis-LeMonda (Eds.), *Child psychology. A handbook of contemporary issues* (pp. 259–286). Philadelphia: Psychology Press.

Maqsud, M. (1997). Effects of metacognitive skills and nonverbal ability on academic achievement of high school pupils. *Educational Psychology*, *17*, 387–397.

Mettes, C. T. C. W., Pilot, A., & Roossink, H. J. (1981). Linking factual and procedural knowledge in solving science problems: a case study in a thermodynamics course. *Instructional Science*, *10*, 333–361.

Minnaert, A., & Janssen, P. J. (1999). The additive effect of regulatory activities on top of intelligence in relation to academic performance in higher education. *Learning and Instruction*, *9*, 77–91.

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

Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw-Hill.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research, 2nd ed*. New York: Holt, Rinehart and Winston.

Perkins, D. N., & Grotzer, T. A. (1997). Teaching intelligence. *American Psychologist*, *52*, 1125–1133.

Raaheim, K. (1988). Intelligence and task novelty. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence, vol 4* (pp. 73–97). Hillsdale: Erlbaum.

Reder, L. M. (1996). *Implicit memory and metacognition*. Mahwah: Erlbaum.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist*, *24*, 113–142.

Schauble, L., Glaser, R., Raghavan, K., & Reiner, M. (1991). Causal models and experimentation strategies in scientific reasoning. *Journal of the Learning Sciences*, *1*, 201–238.

Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty, 2nd ed*. Mahwah, NJ: Erlbaum.

Schnotz, W. (1992). Metacognition and self regulation in text processing: Some comments. In M. Carretero, M. L. Pope, R. J. Simons, & J. I. Pozo (Eds.), *Learning and instruction. European research in an international context, vol. 3* (pp. 365–375). Elsmford, NY: Pergamon Press.

Schraw, G., Dunkle, M. E., Bendixen, L. D., & Roedel, T. D. (1995). Does a general monitoring skill exist? *Journal of Educational Psychology*, *87*, 433–444.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, *7*, 351–371.

Schraw, G., & Nietfeld, J. (1998). A further test of the general monitoring skill hypothesis. *Journal of Educational Psychology*, *90*, 236–248.

Shore, B. M., & Dover, A. C. (1987). Metacognition, intelligence and giftedness. *Gifted Child Quarterly*, *31*, 37–39.

Shute, V. J., Glaser, R., & Raghavan, K. (1989). Inference and discovery in an exploratory laboratory. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences* (pp. 279–326). New York: Freeman.

Slife, B. D., Weiss, J., & Bell, T. (1985). Separability of metacognition and cognition: Problem solving in learning disabled and regular students. *Journal of Educational Psychology*, *77*, 437–445.

Snow, R. E. (1989). Aptitude–treatment interaction as a framework for research on individual differences in learning. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences* (pp. 13–59). New York: Freeman.

Snow, R. E., & Lohman, D. F. (1984). Toward a theory of cognitive aptitude for learning from instruction. *Journal of Educational Psychology*, *76*, 347–376.

Span, P., & Overtoom-Corsmit, R. (1986). Information processing by intellectually gifted pupils solving mathematical problems. *Educational Studies in Mathematics*, *17*, 273–295.

Stankov, L. (2000). Complexity, metacognition, and fluid intelligence. *Intelligence*, *28*, 121–143.

Sternberg, R. J. (1990). *Metaphors of the mind: conceptions of the nature of intelligence*. Cambridge: Cambridge University Press.

Swanson, H. L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology*, *82*, 306–314.

Swanson, H. L., Christie, L., & Rubadeau, R. J. (1993). The Relationship between metacognition and analogical reasoning in mentally retarded, learning disabled, average, and gifted children. *Learning Disabilities Research*, *8*, 70–81.

Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development*, *51*, 1–10.

van Hout-Wolters, B. H. A. M., & Veenman, M. V. J. (2001). *Metacognitieve vaardigheden, intelligentie en leerresultaten bij verschillende taaktypen en leeftijdsgroepen [Metacognitive skills, intelligence, and learning results for different tasks and age groups]. Research grant*. The Hague: Dutch National Science Foundation (http://www.ilo.uva.nl/onderzoek/).

Veenman, M. V. J. (1993). *Intellectual ability and metacognitive skill: Determinants of discovery learning in computerized learning environments*. Amsterdam: University of Amsterdam.

Veenman, M. V. J. (1999). It takes two to tango: over het samenspel van intellectuele en metacognitieve vaardigheden. In R. Hamel, M. Elshout-Mohr, & M. Milikowski (Eds.), *Meesterschap; Zestien stukken over intelligentie, leren, denken en probleemoplossen voor Jan J. Elshout [Mastery: Sixteen papers on intelligence, learning, thinking and problemsolving in the honour of Jan J. Elshout]* (pp. 187–196). Amsterdam: Vossiuspers AUP.

Veenman, M. V. J., Beishuizen, J. J., & Niewold, P. (1997). Intellectual ability and metacognitive skills during text studying. In S. Vosniadou, E. Matsagouras, K. Maridaki-Kassotaki, & S. Kotsanis (Eds.), *Abstracts of the 7th European Conference for Research on Learning and Instruction* (pp. 145–146). Athens: Gutenberg UP.

Veenman, M. V. J., & Elshout, J. J. (1991). Intellectual ability and working method as predictors of novice learning. *Learning and Instruction*, *1*, 303–317.

Veenman, M. V. J., & Elshout, J. J. (1995). Differential effects of instructional support on learning in simulation environments. *Instructional Science*, *22*, 363–383.

Veenman, M. V. J., & Elshout, J. J. (1999). Changes in the relation between cognitive and metacognitive skills during the acquisition of expertise. *European Journal of Psychology of Education*, *14*, 509–523.

Veenman, M. V. J., Elshout, J. J., & Busato, V. V. (1994). Metacognitive mediation in learning with computer-based simulations. *Computers in Human Behavior*, *10*, 93–106.

Veenman, M. V. J., Elshout, J. J., & Groen, M. G. M. (1993). Thinking aloud: Does it affect regulatory processes in learning. *Tijdschrift voor Onderwijsresearch*, *18*, 322–330.

Veenman, M. V. J., Elshout, J. J., & Meijer, J. (1997). The generality vs. domain-specificity of metacognitive skills in novice learning across domains. *Learning and Instruction*, *7*, 187–209.

Veenman, M. V. J., Kok, R., & Kuilenburg, J. (2001). Intelligence and metacognitive skillfulness in secondary education. In F. Oser, & U. Baets (Eds.), *9th European Conference on Learning and Instruction, Abstract Volume* (pp. 166). Mainz: Aachen.

Veenman, M. V. J., Prins, F. J., & Elshout, J. J. (2002). Initial learning in a complex computer simulated environment: The role of metacognitive skills and intellectual ability. *Computers in Human Behavior*, *18*, 327–342.

Veenman, M. V. J., & Van Hout-Wolter, B. H. A. M. (2001). The assessment of metacognitive skillfulness. In F. Oser, & U. Baets (Eds.), *9th European Conference on Learning and Instruction, Abstract Volume* (pp. 278). Mainz: Aachen.

Veenman, M. V. J., & Verheij, J. (2001). Technical students' metacognitive skills: relating general vs. specific metacognitive skills to study success. *Learning and Individual Differences*, *13*, 259–272.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *Journal of Educational Research*, *84*, 30–43.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, *63*, 249–294.

Wason, P. C. (1977). On the failure to eliminate hypotheses—a second look. In P. N. Johnson-Laird, & P. C. Wason (Eds.), *Thinking* (pp. 307–314). Cambridge: Cambridge University Press.

Wilhelm, P., Beishuizen, J. J., & van Rijn, H. (in press). Studying inquiry learning with FILE. *Computers in Human Behavior*.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, *82*, 51–59.