Development of metacognitive skills in young adolescents : a bumpy ride to the high road

Stel, M. van der

Chapter 2

Relation between intellectual ability and metacognitive skillfulness as predictors of learning performance of young students performing tasks in different domains

Abstract

The first objective of this study was to establish the relation between intellectual ability and metacognitive skillfulness as predictors of learning performance in young students (aged 12 years). Furthermore, the generality vs. domain specificity of metacognitive skillfulness was investigated. Thirty-two first-year secondary-school students participated in this study. While thinking aloud, they performed two different tasks representing two different domains: A text-studying task for history and a problem-solving task for mathematics. Participants' intellectual ability, metacognitive skillfulness, and learning performance were assessed. Results show that metacognitive skillfulness contributed to learning performance (partly) independent of intellectual ability. Results also show that metacognitive skills predominantly appear to be general. Domain-specific metacognitive skills, however, played a substantial, but minor role as well.

2.1 Introduction

“Things are always difficult at first” is heard frequently in education. Nevertheless, some students are more successful in acquiring expertise in a domain than others. Elshout (1983) and Schoenfeld (1983) simultaneously introduced the concept of the ‘expert-novice’. These students are able to gain expertise rather rapidly compared to others. Does this rapid progression occur because of their intellectual ability, their metacognitive skills, or because of a combination of both?

This study addresses the issue of how intellectual ability and metacognitive skills contribute to the prediction of learning performance. Furthermore, the generality vs. domain specificity of metacognitive skillfulness is investigated. By using two very different tasks (text studying and problem solving) in widely varying domains (history vs. mathematics) the results of Veenman and Spaans (2005) are further elaborated upon. They found that young secondary-school students have rather strong domain-specific metacognitive skills.

Compared to earlier studies on the generality vs. domain specificity of metacognitive skills (Prins, 2002; Veenman, 1993; Veenman & Beishuizen, 2004; Veenman & Spaans, 2005; Veenman & Verheij, 2003), however, participants in the present study are young students (12 years), and the differences in tasks and domains are maximized.

2.1.1 Intellectual ability

Although researchers diverge in their conceptions of intelligence (see e.g., Brody, 1992; Carroll, 1993; Resnick & Glaser, 1976; Sternberg, 1990), they often relate intelligence to learning. We adopted the rather pragmatic point of view on intelligence from Elshout (1983): Intelligence may be perceived as the magnitude and quality of the human cognitive toolbox, which contains basic cognitive operations. The content of this toolbox is determined by the biological substratum (e.g., hereditary factors or brain damage), but also by the opportunities for acquiring useful cognitive strategies at school or at home (Veenman & Spaans, 2005). 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.

2.1.2 Metacognitive skillfulness

Since Flavell (1979) introduced the concept ‘metacognition’, many studies have addressed the issue of the influence of metacognition on learning performance (for an overview, see Wang, Haertel, and Walberg, 1990). Metacognitive *knowledge* refers to the declarative knowledge one has about the interplay between personal characteristics, task characteristics, and available strategies in a learning situation (Flavell, 1979), while metacognitive *skills* concern the procedural knowledge that is required for the actual regulation of, and control over one's learning activities (Brown & DeLoache, 1978). Task analysis, planning, monitoring, checking and recapitulation are manifestations of such skills. An interesting question is whether metacognitive skills are part of the intellectual toolbox or repertoire; or as Slife, Weiss, and Bell (1985) formulated …"whether metacognition can be reduced to cognition".

2.1.3 Relation between intellectual ability, metacognitive skillfulness, and learning performance

Veenman (1993; Veenman & Beishuizen, 2004; Veenman, Elshout, & Meijer, 1997; Veenman & Verheij, 2003;) described three, mutually exclusive models concerning the relation between intellectual ability and metacognitive skillfulness as predictors of learning performance. The *intelligence* model regards metacognitive skillfulness as an integral part of intellectual ability. In this model metacognitive skillfulness does not contribute to learning performance on top of intellectual ability. According to this model, metacognitive skills cannot have a predictive value for learning performance independent of intellectual ability. The second, contrasting model is the *independency* model, in which intellectual ability and metacognitive skillfulness are regarded as entirely independent predictors of learning performance. Finally, in the *mixed* model intellectual ability and metacognitive skillfulness are correlated, but they also have their own, unique contribution to the prediction of learning performance.

Over the last decades, support has been found for each of these models (for an overview, see Veenman & Spaans, 2005; Veenman, Wilhelm & Beishuizen, 2004). Many studies, however, are difficult to compare, due to dissimilarities in assessing metacognitive skillfulness (thinking aloud, observation, questionnaires), in participants (age, level of intellectual ability), and in tasks and domains. Moreover, the focus of some studies is restricted to the relation between intellectual ability and metacognitive skillfulness, thereby excluding the relation of both predictors with learning performance. The evidence found seems to be highly in favor of the mixed model, but many of these studies concerned the metacognitive skillfulness of older secondary-school or university students. It remains to be ascertained more thoroughly whether the mixed model can be generalized to younger students with less developed metacognitive skills, performing different tasks in different domains (see Veenman & Spaans, 2005; Veenman et al., 2004).

Therefore, the first objective of the present study is to establish the relation between intellectual ability and metacognitive skillfulness as predictors of learning performance in young students. We hypothesize that the mixed model can be generalized to first-year secondary-school students, regardless of tasks or domains.

2.1.4 Generality vs. domain specificity of metacognitive skillfulness

A second objective of this study is to establish whether metacognitive skillfulness is general (i.e., domain independent), or domain specific. Earlier studies concerning this issue have yielded contradictory results. On the one hand, researchers have found evidence for general, domain-independent metacognitive skills (Schraw, Dunkle, Bendixen, & Roedel, 1995; Schraw & Nietfeld, 1998; Veenman & Beishuizen 2004; Veenman & Verheij 2003; Veenman et al., 1997, 2004). Despite this considerable number of studies with comparable results, it should be kept in mind that participants, tasks, and domains varied substantially. On the other hand, other reseachers (De Jong, 1992; Glaser, Schauble, Raghavan, & Zeitz, 1992; Kelemen, Frost, & Weaver, 2000) found evidence against a general metacognitive ability.

Veenman et al. (2004) found support for the generality of metacognitive skills among young students (9- to 22-year olds) performing discovery-learning tasks in different domains. Recently, this support for young students' general metacognitive skills could not be corroborated by a study of Veenman and Spaans (2005). In this study 12- and 15-year olds performed two tasks in different domains (solving math word problems and an inductive-learning task for biology). Metacognitive skills of the younger students appeared to be rather domain-specific, whereas those of the older ones turned to be general by nature.

Based on earlier studies, the onset of metacognitive *skill* development for academic tasks is not to be expected before the age of 8 – 10 years (Alexander, Carr, & Schwanenflugel, 1995; Berk, 2006; Kuhn, 1999; Siegler, 1998; Veenman, et al., 2004). The academic metacognitive skills of 12-year olds are still premature and developing, which may explain why these skills may diverge on notably different tasks.

Participants in the present study are rather young and inexperienced in applying academic metacognitive skills. Therefore, we expect their metacognitive skills to be in a transitory phase of development. We hypothesize that children, slightly over the age of 12 years, will use general as well as domain-specific metacognitive skills. In order to allow domain-specific metacognitive activities, the differences in tasks and domains were maximized (see section 2.2.3).

2.2 Method

2.2.1 Participants

Thirty-two first-year secondary-school students (12 boys, 20 girls; average age 12 years and 8 months) participated in this study. They were recruited from 85 students of three different tracks (pre-university education, higher general education and pre-vocational education)[1] of an urban school in the Netherlands.

A history pretest was administered to all 85 students. This test consisted of 16 multiple-choice questions about American history (Cronbach's alpha = .71). Students were instructed not to guess the right answer. The pretest was administered for reasons of selection and played no further role in data analysis. One boy was excluded from further participation due to prior knowledge of the topic, that is, he correctly answered more than 75% of the 16 MC questions. Next, 40 participants were selected out of the remaining 84 students on their intellectual ability (see section 2.2.2). Eight low intelligent students, however, were excluded from the study due to learning or conduct disorders (e.g., dyslexia or ADHD). The mean score on the history pretest for the remaining 32 participants was 4.0 (*sd* = 2.4), i.e., 25% of the maximum score. For mathematics no pretest was administered. Here students need a certain amount of lower-level mathematics knowledge in order to solve more complex problems.

Consent was requested from and given by the participants' parents. After completing the tasks, participants received a small financial reward for their participation.

2.2.2 Intellectual ability

The intellectual ability of 84 students was assessed by a series of ability tests. Three subtests from the Groninger Intelligence test for Secondary Education (GIVO, standardized Dutch intelligence test; Van Dijk & Tellegen, 1994) were selected: Number Series, Verbal Analogies and Unfolding Figures. With these subtests a number of the primary intelligence factors (Carroll, 1993) is assessed: Inductive and deductive reasoning abilities, both verbal and quantitative, and visuospatial ability.

The GIVO, however, lacks a test for assessing memory abilities, another primary factor in Carroll's (1993) model highly relevant to text studying. Therefore, a fourth test (Names & Professions; see Veenman & Beishuizen, 2004) was added. An overall Intellectual Ability score was obtained by transforming the scores on all tests into z-scores and then calculating the mean z-score (Cronbach's alpha = .76).

[1] In Dutch: VWO, HAVO and VMBO-T respectively

To guarantee sufficient variance in intellectual ability the median of the Intellectual Ability scores was calculated and participants were denominated as high (1st quartile), average (2nd and 3rd quartile) or low (4th quartile) in intelligence. Finally, 13 high, 13 average and 6 low intelligent students were selected as participants. In the 4th quartile there were relatively many students with learning or conduct problems. That is the reason why only six students could be selected as participants in the low range of intelligence.

2.2.3 Tasks

Students performed two different tasks representing two different domains. One task was a text-studying task in the domain of history and the other one was a problem-solving task in the domain of mathematics.

2.2.3.1 History task

In an individual session of 50 minutes, participants were asked to study a history text in the same way as they usually do when preparing for a test. They had to read aloud and think aloud while reading the text and answering the questions or assignments embedded in the text. Participants were allowed to study the text for 30 minutes and in the remaining 20 minutes the post-test was administered (see section 2.2.5).

The history text was composed of parts from two of the most frequently used Dutch schoolbooks for history: “MeMo” (Van Boxtel & Schrover, 1998) and “Sprekend verleden” (Buskop, Dalhuisen, & Geest, 1998). To avoid prior knowledge of the subject on the one hand, and to appeal to the zone of tolerable difficulty on the other hand, the text was based on a subject of the second-year curriculum, instead of the first-year curriculum.

The text about slavery and the civil war in the United States of America contained 76 concepts and 1479 words. In the text, three questions or assignments were embedded. These questions were not meant for testing the students’ knowledge, but to elicit (more) metacognitive activities (e.g., ‘There are several reasons why the north and the south were at war with each other. Describe in your own words at least two of these reasons’). From a pilot-study we learned that if a text does not contain such questions, many participants just tend to read linearly.

2.2.3.2 Mathematics task

In another individual session of 50 minutes, participants had to solve five mathematical word problems in 20 minutes. Several categories of problems were presented (e.g., distance, fraction, surface area, percentage problems). For instance, a fraction problem was 'My brother received two bags of marbles. Each bag contains 48 marbles in four different colors: 1/6 is yellow; 3/8 is blue; 1/3 is green and the rest is red. Of which color does my brother have most marbles?' Together with the assignments, participants received a sheet containing the answers and a brief stepwise explanation of how to solve the problems. Participants were free to consult this sheet whenever and as much as they liked. The first 20 minutes were considered as a learning-by-doing phase. Next, the participants handed in all materials and received another series of five parallel problems, which had to be solved without any help in the remaining 30 minutes. This second part is considered as a post-test assessment of learning performance (see section 2.2.5). All problems had to be solved while thinking aloud.

2.2.4 Metacognitive skillfulness

The transcribed thinking-aloud protocols for both tasks were analyzed on spontaneous use of metacognitive skills according to the procedure of Veenman (1993; Veenman & Beishuizen, 2004; Veenman, Kerseboom, & Imthorn, 2000). Metacognitive skillfulness was divided into four subscales: Orientation (O), Planning and Systematic orderliness (P), Evaluation (Ev), and Elaboration (El). Some metacognitive activities are general across both tasks and domains, whereas other metacognitive activities seem to be more domain-specific (see Table 2.1).

To obtain a reliable score for metacognitive skillfulness, scoring criteria were formulated for both tasks. For each subscale it was described which criteria should be met in order to receive a certain rate for the quality[2] of metacognitive skillfulness. This resulted in a method, which allowed assessing general as well as domain-specific metacognitive activities. The scoring method consisted of two steps for each protocol. First, each utterance was coded in the margin as belonging to one of the four subscales (O, P, Ev or El). Secondly, each subscale received a qualitative score according to the formulated criteria. This score was a total score per subscale per protocol. A five-point scale (ranging from 0 to 4)[3] was used for each subscale. It was the quality, not merely the quantity of metacognitive activities that determined the scores. For example, a participant received a higher score for 'deeper' Elaboration (e.g., drawing a conclusion

[2] The data in the tables relate to the quality of metacognitive skillfulness. The same analysis was performed with the quantitative scores of metacognitive skillfulness. Quantitative scores were obtained by counting the frequency of metacognitive activities of each subscale (e.g., if a student evaluated five times, his quantitative Evaluation score was five).The results are comparable with the results of the qualitative data. They follow the same pattern, but are somewhat less pronounced.

[3] Results of a CatPCA (formerly PRINCALS) show that it is permitted to treat the scores for metacognitive skills as interval variables.

in one's own words) than for a superficial one (e.g., summarizing a paragraph almost literally).

Table 2.1 Examples of domain-specific and general metacognitive activities

	History-specific	Math-specific	General
Orientation			
a.			Activating prior knowledge
b.			Goal setting
c.	Predicting the content of the text	Estimating the answer	
d.	Reading titles of paragraphs prior to reading the entire text	Making a sketch of the problem in order to represent the problem	
Planning & systematic orderliness			
a.			Subgoaling
b.			Time management
c.	Designing a reading plan and deciding upon which text parts to pay attention to	Designing a step-by-step action plan, instead of working by trial-and-error	
d.	Note taking (self-instruction for doing so)	Writing down calculations step by step	
Evaluation			
a.			Expressing non-understanding
b.			Comment on own activities
c.	Monitoring text comprehension during reading	Monitoring action plan	
d.	Self-correction after rereading (parts of) the text	Checking an answer by recalculating	
Elaboration			
a.			Recapitulating and drawing conclusions
b.			Relating the answer to the question or problem
c.	Paraphrasing (parts of) the text	Paraphrasing the problem	
d.	Summarizing (self-instruction for doing so); Making inferences during reading	Drawing conclusions while referring to the problem statement	

It is important to emphasize that the judges intentionally avoided the confounding of metacognition scores by the correctness or incorrectness of the content matter. So, an incorrect, but highly elaborated conclusion could equally generate a high score for 'Elaboration' as long as it was in line with the participant's own reasoning.

For each task Cronbach's alpha was calculated over the four subscales: (History .72; Mathematics .83). Sum scores on the subscales were calculated, representing the quality of metacognitive skillfulness for each task. Furthermore, six protocols for each task were simultaneously rated by two judges. This resulted in an alpha interrater reliability of .97 for history and .89 for math.

2.2.5 Learning performance

After the learning phase of both tasks, the learning performance of participants was assessed by a post-test, as was explained to them in advance. For history the post-test consisted of five multiple-choice questions and six essay questions. The MC questions assessed reproductive knowledge (facts and dates, e.g., 'when did the Americans officially abolish slavery?') and the essay questions were meant to assess overall text comprehension (e.g., 'Explain in your own words why Lincoln changed his opinion about slavery several times'). Participants were not allowed to consult the text or their notes while answering the questions.

According to a rating system, points were given for the correctness of answers to each question. MC questions could render one point, while essay questions could render a maximum of four points. A total score was calculated and used as a measure of learning performance in history (Cronbach's alpha = .58).

For math a post-test with five math word problems was administered. For each of the five math post-test problems an equal amount of points could be earned: two points if both the procedure and the answer was correct; one point if either one of them was correct; and zero points if neither of them was correct. A total score was calculated and used as a measure of learning performance in mathematics (Cronbach's alpha = .71).

2.2.6 Procedure

The intellectual-ability test and the history pretest were administered during a group session of 100 minutes. The individual sessions took place during school time. The experiment had a counterbalanced design with respect to task order, meaning that 16 students started with history and 16 students with mathematics. Participants could make use of a pen, pencil, text highlighter, ruler, calculator, and blank sheets of paper for making notes.

All participants were instructed to think aloud while working on both tasks. The experimenter was not allowed to help the students in any way. To encourage the student to keep on thinking aloud the experimenter used standard prompts (e.g., 'please, keep on thinking aloud') whenever the student fell silent. All the utterances of the participants were audio-taped, transcribed, and analyzed in relation to metacognitive skillfulness.

2.3 Results

All participants (*N* = 32) performed the tasks, so there are no missing values in the data.In Table 2.2 the means and standard deviations of the Learning measure and Metacognitive skillfulness for both tasks are depicted, as well as the maximum obtainable score.

Table 2.2 Means and standard deviations

	Mean	**Std. deviation**	**Maximum score**
MetatotHis*	5.03	3.40	16.0
Learning measureHis	8.47	3.28	20.0
MetatotMath*	6.09	3.31	16.0
Learning measureMath	4.44	2.54	10.0

**Note*: MetatotHis means the total score on the quality of metacognitive activities during history task; MetatotMath means the total score on the quality of metacognitive activities during math task.

2.3.1 Relation between intellectual ability, metacognitive skillfulness and learning performance

Correlations between intellectual ability, metacognitive skillfulness, and learning performance on both tasks were calculated (see Table 2.3). As far as the math task is concerned, Intellectual ability correlated significantly with both Metacognitive skillfulness and the Learning measure. The same applies for the correlation between Metacognitive skillfulness and the Learning measure for math. The results on the history task differ partly from the results on the math task. A significant correlation was found between Intellectual ability and the Learning measure, but the correlation between Intellectual ability and Metacognitive skillfulness was not significant. Metacognitive skillfulness, however, correlated significantly with the Learning measure for history.

Table 2.3 Correlations between Intellectual ability, Metacognitive skills and Learning performance for both history and mathematics

	Intellectual ability	**MetatotHis**	**Learning measureHis**	**MetatotMath**
MetatotHis	.27			
Learning measureHis	.40*	.65**		
MetatotMath	.62**	.35*	.38*	
Learning measureMath	.75**	.27	.49**	.72**

* $p < 0.05$ ** $p < 0.01$

To determine whether correlations for both tasks deviate significantly from each other, Fisher-*z* ratios were calculated (Guilford, 1965). There is no difference in correlations between both tasks as long as Metacognitive skillfulness is involved (Fisher-*z* ratio = 1.71, n.s., for the correlations between Intellectual ability and Metacognitive skillfulness, and 0.50, n.s., for the correlations between Metacognitive skillfulness and Learning measures). Intellectual ability, however, plays a more important role in predicting the learning performance for mathematics relative to history (Fisher-*z* ratio = 2.09, $p < 0.05$).

Next, semi-partial correlations (Nunnally, 1967) were calculated by partialing Metacognitive skillfulness from the correlation between Intellectual ability and the Learning measure (i.e., semi-partIntel) and partialing Intellectual ability from the correlation between Metacognitive skillfulness and the Learning measure (i.e., semi-partMeta). These semi-partial correlations (see Table 2.4) are needed to calculate the unique, independent contribution of Metacognitive skillfulness and Intellectual ability to the Learning measures.

Table 2.4 Semi-partial correlations

	Semi-partIntel	**Semi-partMeta**
Learning measureHis	.23	.53**
Learning measureMath	.39**	.33**

**$p < 0.01$

Using regression-analytic techniques (Pedhazur, 1982; Veenman, 1993; Veenman & Verheij, 2003; Veenman & Spaans, 2005) the unique and shared proportions of variance in the Learning measures were distributed to Metacognitive skillfulness and Intellectual ability (see Table 2.5). The math data could be taken as an example. The squared multiple correlation of Intellectual ability and Metacognitive skillfulness for predicting Learning measure in math was calculated from the correlations presented in Table 2.3 and 2.4 (R^2 = the squared correlation between Intellectual ability and Learning measure + the squared semi-partial correlation between Metacognitive skillfulness and Learning measure with Intellectual ability partialled out = $.75^2 + .33^2 = .671$). The unique contribution of Intellectual ability to Learning measure was determined by calculating the squared semi-partial correlation between Intellectual ability and Learning measure with Metacognitive skillfulness partialled from Intellectual ability ($r^2 = .152$). Consequently, it was estimated that Intellectual ability uniquely accounted for 15.2% of the variance in Learning measure for math, Metacognitive skillfulness uniquely accounted for 10.9% of the variance, while both predictors had another 41.0% of variance in common. This procedure was applied to both tasks (see Table 2.5).

Table 2.5 Percentage of variance accounted for in Learning measures

	Intel unique	Meta unique	Shared	Total
History	5.2	28.0	10.8	44.0
Math	15.2	10.9	41.0	67.1

Note: Intel unique means the unique contribution of Intellectual ability to Learning measure; Meta unique means the unique contribution of Metacognitive skillfulness to Learning measure; Shared means the shared contribution of Intellectual ability and Metacognitive skillfulness to Learning measure.

The history results in Table 2.5 show that, despite the variance shared with Intellectual ability, Metacognitive skillfulness substantially added to the prediction of Learning measure in history on top of Intellectual ability. In fact, the unique contribution of Metacognitive skillfulness outweighed the unique contribution of Intellectual ability to Learning measure in history.

2.3.2 Metacognition across domains

A principal component analysis (PCA) with a two-factor solution was performed on the four subscales of metacognitive skillfulness for both tasks (see Table 2.6).

Table 2.6 Unrotated component matrix for metacognitive skillfulness

	Component 1	Component 2
Eigenvalue	3.35	1.71
Variance proportion	.42	.21
metaorientationHis	.41	.21
metaplanningHis	.63	.64
metaevaluationHis	.47	.44
metaelaborationHis	.68	.51
metaorientationMath	.65	-.48
metaplanningMath	.76	-.32
metaevaluationMath	.58	-.67
metaelaborationMath	.87	-.13

The results of the unrotated PCA show that all measures of metacognitive skillfulness substantially load on the first component. This component has an eigenvalue of 3.35 and a variance proportion of .42. Although component loadings are not extremely high, ranging from .41 to .87, this first component may be interpreted as representing general metacognitive skills across the two tasks/domains. On the other hand, we find contrasted loadings in the second component, that is, positive loadings for history and negative ones for math. This points in the direction of a component representing domain-specific metacognitive skills. With an eigenvalue of 1.71 and a variance proportion of .21 this second component cannot be ignored.

2.4 Discussion

This study investigated the relation between intellectual ability and metacognitive skillfulness as predictors of learning performance in young learners. As was expected, results corroborate the mixed model. First, intellectual ability and metacognitive skillfulness are moderately correlated. Moreover, both intellectual ability and metacognitive skillfulness have their own, unique contribution to learning performances on both tasks. These results are similar to results for older age groups obtained in other studies (Veenman, 1993; Veenman & Beishuizen, 2004; Veenman & Verheij, 2003; Veenman et al., 2004). Therefore, the mixed model can be generalized to twelveyear-olds, although the small number of participants may be considered as a possible limitation of this study.

The second research question concerned the generality vs. domain specificity of metacognitive skillfulness. The PCA on the metacognitive data shows a two-component solution: A first component, which can be interpreted as general metacognitive skills, and a second component, which can be interpreted as domain-specific metacognitive skills. Results support our expectation that metacognitive skills of 12 year old students represent a general as well as a domain-specific component. This may indicate that these students are in a transitory phase of metacognitive-skill development. Veenman and Spaans (2005) assumed that metacognitive skills initially develop on separate islands of tasks and domains. Beyond the age of 12, these skills will gradually merge into a more general repertoire that is applicable and transferable across tasks and domains. Among 12-year olds a phase of transition is characterized by applying recently acquired general metacognitive skills, along with a remainder of domain-specific metacognitive skills.

From a developmental perspective, it is interesting to know more about the development of metacognitive skillfulness. On the one hand, the development of metacognitive skills in relation to intellectual ability as predictors of learning performance. On the other hand, the development of the nature of metacognitive skills (general vs. domain-specific). A longitudinal study will offer the opportunity to investigate these aspects of metacognitive development. The present study is the first part of a longitudinal project, where the same students will be followed for three consecutive years.