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

Stel, M. van der

# Chapter 4

## Metacognitive skills and intellectual ability of young adolescents: A longitudinal study from a developmental perspective

## Abstract

This study shows the final results of a longitudinal project where the same participants were followed for three consecutive years as they enter secondary school (aged 12 to 15 years). The first objective of this study was to investigate the development of both quantity and quality of metacognitive skills. The second objective was to establish whether the development of metacognitive skills is intelligence-related or relatively intelligence-independent. Finally, the generality vs. domain-specificity of developing metacognitive skills was investigated. In the first year 32 first-year secondary-school students participated in this study. In the second and third year, respectively 28 and 25 students participated. While thinking aloud, the participants performed two different tasks representing two different domains: A text-studying task for history and a problem-solving task for mathematics. Each year participants were given new tasks, suitable for their age. Participants' intellectual ability was assessed, as well as their metacognitive skills and learning performance for both domains separately. Results of the first two years show a significant growth of both the quantity and the quality of metacognitive skills. In the third year this growth did not continue. Furthermore, results show that metacognitve skills contributed to learning performance partly independent of intellectual ability. Results also show that metacognitive skills appear to be predominantly general by nature over the years. A smaller domain-specific component was found as well in the first two years, while this component disintegrated in the third year. In conclusion, the age around 15 years appears to be a relevant point in time during the developmental trajectory of metacognitive skills. At this age the growth of metacognitive skills is put on hold, while the nature of these skills becomes fully general.

## 4.1 Introduction

Since Flavell (1979) introduced the concept 'metacognition', many studies have addressed the issue of the influence of metacognition on learning performance. Based on a meta-review of studies, Wang, Haertel, and Walberg (1990) concluded that metacognition is the most important predictor of learning performance. The present study focuses on the development of metacognitive skills in relation to intellectual ability. Another important issue concerns the generality vs. the domain specificity of metacognitive skills.

### *4.1.1 Metacognitive skills*

A generally accepted distinction is the one between metacognitive knowledge and metacognitive skills. 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). Having metacognitive knowledge at one's disposal, however, appears to be no guarantee for using this knowledge whenever it is needed. Alexander, Carr and Schwanenflugel (1995) found a discrepancy between childrens' knowledge about monitoring and applying monitoring skills during task performance. In the same vein, Winne (1996) stated that knowledge has no effect on behavior until it is actually being used. Students who are aware of effective strategies still do not always use them (Barnett, 2000; Pressley, Yokoi, Van Meter, Van Etten, & Freebern, 1997). Metacognitive skills concern the *procedural* knowledge that is required for the actual regulation of, and control over one's learning activities (Brown & DeLoache, 1978; Veenman, Elshout, & Meijer, 1997; Veenman, 2011). Task analysis, goal setting, planning, monitoring, checking, and recapitulation are manifestations of such skills. Metacognitive skills can be inferred from overt behavior or utterances by the student, that is, from concrete metacognitive activities (Veenman, Van Hout-Wolters, & Afflerbach, 2006). These activities can be divided in behavior that occurs at the onset of task performance (orientation), during task performance (planning, monitoring, evaluation) and after task performance (reflection and elaboration). Examples of metacognitive activities are given in Table 4.1. Note that some of the behavior in Table 4.1 may be considered as cognitive, but the purposeful application of such cognitive behavior at the appropriate moment results from metacognitive skillfulness. It reflects the intention to attain control over the cognitive task. Several researchers (Bowen, Shore, & Cartwright, 1992; Brown, 1980; Christoph, 2006; Markman, 1977, 1979; Mevarech & Fridkin, 2006; Pressley & Afflerbach, 1995; Schoenfeld, 1992; Shore & Lazar, 1996) have investigated the use of metacognitive skills while performing different tasks (e.g., reading comprehension or problem solving) often focusing on a separate component of metacognitive skills (e.g., planning or monitoring activities). The present study, however, includes a broad range of metacognitive skills referring to orientation, planning, monitoring, as well as elaboration skills in both problem-solving and text-studying tasks. While assessing metacognitive skills over a developmental trajectory two different perspectives can be taken: The quantity and the quality of metacognitive skills. The quantity concerns the frequency of these skills being applied, whereas the quality concerns the depth or the extent to which they are applied. An example of the latter is that making a sketch of the problem in order to represent the problem is considered as a deeper orientation than just reading (a part of) the problem statement. Using metacognitive skills more frequently does not automatically imply that these metacognitive skills have a higher level of quality. Therefore, this study focuses on the development of both the quantity and the quality of metacognitive skillfulness.

### *4.1.2 Metacognitive skills from a developmental perspective*

The first objective of this study was to investigate the development of metacognitive skills. Flavell (1992) related the concept of metacognition to Piaget's developmental stage of formal-operational thinking (Inhelder & Piaget, 1958). At this stage children are capable of hypothetico-deductive reasoning, which requires metacognitive control. Flavell indicated that Piaget would not expect metacognition to show up before the stage of formal-operational thinking has been reached. More recent studies, however, showed that, alongside with the "theory of mind" (ToM), that is, the understanding of one's own and other people's state of mind, (Wellman, 1990), young preschoolers already start to develop some metacognitive awareness (Blöte, Van Otterloo, Stevenson, & Veenman, 2004; Demitriou & Efklides, 1990; Kuhn, 1999). Larkin (2006) found a relation between ToM, metacognitive knowledge and strategy use in two 5- to 6-year-olds. A further metacognitive development in later childhood concerns not only the metacognitive knowledge, but also the onset of the development of metacognitive skills. Although Whitebread et al. (2009) found some planning and monitoring activities in playful situations with youngsters as young as 5 years old, it is generally assumed that the development of metacognitive skills in educational contexts commences around the age of 8-10 years (Berk, 2006; Kuhn, 1999; Siegler, 1998; Veenman et al., 2006; Veenman, Wilhelm, & Beishuizen, 2004). In a cross-sectional study concerning the frequency of metacogntive skills, Veenman et al. (2004) found a linear growth in metacognitive skills between the age of 9 and 22 yrs. In another study (Veenman & Spaans, 2005) a significant growth of the quality of students' metacognitive skills (12

– 15 years) was found. A growth in both frequency and quality of metacognitive skills was found in two studies with participants aged between 13 and 15 yrs. and 12 and 14 yrs. respectively (Van der Stel, Veenman, Deelen, & Haenen, 2010; Van der Stel & Veenman, 2010). In line with these results, metacognitive skills are expected to increase in quantity (frequency) as well as in quality (depth) over the years (hypothesis 1).

### *4.1.3 Relation between intellectual ability, metacognitive skills, and learning performance from a developmental perspective*

A related research issue is whether the development of metacognitive skills is intelligence-related or relatively intelligence-independent. Several researchers (Alexander et al., 1995; Borkowski & Peck, 1986; Schneider & Pressley, 1997; Van der Stel & Veenman, 2008; Van der Stel, et al., 2010; Veenman et al., 2004; Veenman & Spaans 2005) investigated metacognitive ability in relation to intellectual ability.

An interesting question is whether metacognitive skills are part of intelligence, that is, "whether metacognition can be reduced to cognition" (Slife, Weiss, & Bell, 1985). Veenman (1993; Veenman & Beishuizen, 2004; Veenman & Verheij, 2003; Veenman et al., 1997) described three, mutually exclusive models concerning the relation between intellectual ability and metacognitive skills as predictors of learning performance. The *intelligence* model regards metacognitive skills as an integral part of intellectual ability. According to this model, metacognitive skills do not have a predictive value for learning performance independent of intellectual ability. Sternberg (1990), for instance, advocates such an inclusive position of 'metacomponents' in his triarchic theory of intelligence. The second, contrasting model is the *independency* model, in which intellectual ability and metacognitive skills are regarded as entirely independent predictors of learning performance. Finally, in the *mixed* model intellectual ability and metacognitive skills are correlated, but metacognitive skills has a surplus value on top of intellectual ability for 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 et al., 2004). Many studies, however, are difficult to compare, due to dissimilarities in assessing metacognitive skills (thinking aloud, observation, questionnaires), in participants (age, educational background), and in tasks and domains. Moreover, the focus of some studies is restricted to the relation between intellectual ability and metacognitive skills, thereby excluding the relation of both predictors with learning performance. The evidence found so far seems to be in favor of the mixed model, albeit many of those studies concerned the metacognitive skills of older secondary-school or university students in cross-sectional designs. From the perspective of the development of metacognitive skills, it remains to be ascertained more thoroughly in a longitudinal design whether the mixed model can be generalized to younger students with initially developing metacognitive skills, performing different tasks in different domains (see Veenman & Spaans, 2005; Veenman et al., 2004). More specifically, the role of intellectual ability in the *development* of metacognitive skills will be addressed in this study. Alexander et al. (1995) formulated three developmental hypotheses with regard to the relation between intelligence and the development of metacognition. According to the *ceiling* hypothesis, initial effects of intelligence on the development of metacognition diminish over time. The *acceleration* hypothesis, on the other hand, assumes that the impact of intelligence on the development of metacognition increases with age. The *monotonic development* hypothesis, finally, assumes that both intelligence and metacognition show a monotonic growth over age independent of each other. In their literature overview, Alexander et al. (1995) found support for the monotonic development of metacognitive knowledge. Gifted children showed a general superiority in their declarative metacognitive knowledge, relative to non-gifted children at all ages. Giftedness effects were persistent throughout childhood with gifted children from early elementary school to junior high school showing a similar superiority in this knowledge. With regard to metacognitive skills, however, results were inconclusive. Young gifted children were not more spontaneously strategic than non-gifted children. However, a giftedness advantage showed up in the middle school and high school years for both spontaneous and complex strategy use. In a cross-sectional study, Veenman and Spaans (2005) obtained evidence in favor of a monotonic development of metacognitive skills. They obtained support for a monotonic maturation effect of both intellectual ability and metacognitive skills in students of 13 and 15 years performing a problem-solving task and an inductive-learning task.

In the present study, it is hypothesized that metacognitive skills develop alongside, but not fully dependent on intellectual ability, regardless of tasks and domains. A monotonic development of both metacognitive skillfulness and intellectual ability as predictors of learning performance is expected, in line with the mixed model and the monotonic development hypothesis (hypothesis 2).

### *4.1.4 Metacognitive skills across domains*

Another objective of this study was to establish whether metacognitive skills are general or domain specific. From a developmental as well as from an instructional perspective it is relevant to know not only whether metacognitive skills develop, but also how they develop: Whether they develop from being general into becoming domain specific or the other way around? Earlier studies concerning this issue yielded contradictory results. Despite differences in assessment, age groups, tasks and domains, some researchers found evidence for general metacognitive skills (Schraw, Dunkle, Bendixen, & Roedel, 1995; Schraw & Nietfeld, 1998; Veenman & Beishuizen 2004; Veenman & Verheij 2003; Veenman et al., 1997, 2004). On the other hand, De Jong (1992) found that the quality and quantity of metacognitive activities of secondary-school students varied substantially across tasks. Glaser, Schauble, Raghavan, and Zeitz (1992) showed that metacognitive activities of university students varied across different discovery-learning tasks, although improvement between subsequent tasks did not rule out the existence of general strategies. Kelemen, Frost, and Weaver (2000) concluded that individual differences in meta-memory accuracy were not stable across consecutive sessions and tasks, which they interpreted as evidence against a general metacognitive ability.

Veenman et al. (2004) found support for the generality of metacognitive skills among young novices (aged 9 to 22 yrs.) performing discovery-learning tasks in different domains. This support for young students' general metacognitive skills could not be corroborated by a study of Veenman and Spaans (2005). In their study 13- and 15-year olds performed two tasks in different domains (math word-problem solving 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 out 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 et al., 1995; Berk, 2006; Kuhn, 1999; Siegler, 1998). 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. Veenman and Spaans (2005) assumed that metacognitive skills might be initially acquired within separate tasks and domains, and then progressively become a generalized repertoire across tasks and domains.

Participants in the present study are rather young adolescents and inexperienced in applying academic metacognitive skills. Spear (2000) characterized adolescence as a transitional developmental period. "Adolescence is the gradual period from childhood to adulthood……adolescence is a period of transitions rather than a moment of attainment" (p.417). Therefore, metacognitive skills of young adolescents are expected to be in a transitory phase of development, which implies that both general and domain-specific metacognitive skills will be used. It is hypothesized that the participants will initially use general as well as domain-specific metacognitive skills. It is also hypothesized that the initially acquired domain-specific metacognitive skills tend to generalize during further development. Therefore, older students are expected to resort less to domain-specific metacognitive skills than younger students do (hypothesis 3).

## 4.2 Method

### *4.2.1 Participants*

In the first year of this three-year-longitudinal project, 32 first-year secondary-school students participated. They were recruited from 85 students of three different tracks (pre-university education, higher general education and pre-vocational education) of an urban school in the Netherlands. This school is known because of its large diversity of children, thus representing a broad educational level of the students, a broad range of social economic status of parents, and various ethnic backgrounds. It was chosen not to work with more than one school in order to avoid confounding variables, such as differences in teachers, pedagogical/didactic philosophy, schoolbooks, etc. Students with learning or conduct disorders (e.g., dyslexia or ADHD) were excluded from the study. Participants were distributed equally over the three tracks. Participants were selected on their intellectual ability (see section 4.2.3). Consent was requested from and given by the participants' parents. In the second year, four students withdrew as participants due to changing residence or school. In the third year, another three students withdrew for the same reason. In the third year, 25 students (8 boys, 17 girls; average age 14 years and 7 months) participated in the third part of this study[7]. The data in the present study refer to the 25 students that participated in all three years. After completing the tasks, participants received a small financial reward.

### *4.2.2 Metacognitive skills*

All transcribed thinking-aloud protocols were analyzed on metacognitive skills according to the procedure of Veenman (Prins, Veenman, & Elshout, 2006; Van der Stel & Veenman, 2008, 2010; 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). In Table 4.1 general metacognitive activities across both tasks and domains as well as more specific metacognitive activities for text-studying and problem-solving tasks are

[7] Effects of selective loss of participants over the years were checked for and not found. The same applies for effect of gender.

described for each subscale of metacognitive skillfulness (Meijer, Veenman, & Van Hout-Wolters, 2006; Pressley, 2000; Pressley & Afflerbach, 1995).

**Table 4.1** 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 | |

The scoring method consisted of two steps for each protocol. First, an utterance was coded in the margin if belonging to one of the four subscales (O, P, Ev or El). This resulted in a quantitative score obtained by counting the frequency of metacognitive activities on each subscale (e.g., if a student evaluated five times, the quantitative Evaluation score was five). Secondly, a score for the quality of metacognitive skills was judged from the protocols. To obtain a reliable score for the quality of metacognitive skills, scoring criteria were formulated for each subscale. This resulted in a method, which allowed for assessing general as well as domain-specific metacognitive activities. A five-point scale (ranging from 0 to 4) was used for each subscale. For example, a participant received a higher score for 'deeper' Elaboration (e.g., drawing a conclusion in one's own words) than for a superficial one (e.g., summarizing a paragraph almost literally). It is important to emphasize that the judges intentionally avoided the confounding of metacognition scores with 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.

Each year, 6 protocols of each task were rated by two judges separately. The interrater reliability was computed on the summed scores over the four subscales of metacognition. Since the interrater reliability was high, the remaining protocols were analyzed and rated by one judge. Cronbach's alpha interrater reliability ranged from .77 to .93 for the quantitative scores and from .89 to .97 for the qualitative scores.

### *4.2.3 Intellectual ability*

Each year, students' intellectual ability 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 three subtests a number of the primary intelligence factors (Carroll, 1993) is assessed: Both verbal and numerical inductive and deductive reasoning abilities and visuospatial ability. The GIVO, however, lacks a test for assessing memory abilities, another primary factor in Carroll's (1993) model highly relevant to the prediction of school performance (Crone et al., 2006). Therefore, a fourth test (Names & Professions, requiring the memorization of word pairs; see Veenman & Beishuizen, 2004) was added. In order to determine the growth in intellectual ability, the raw scores of the aforementioned four subtests were compared. Furthermore, raw scores were transformed into z-scores and for each participant a mean z-score was calculated over the four subtests for each year. This resulted in a total

score of the participant's intellectual ability for each year. These scores were used in the correlational analyses.
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. Participants were selected at random from each quartile, so as to attain an equal distribution over the quartiles.

### *4.2.4 Tasks*

To ensure the novelty of tasks, each year participants were given *new* tasks with task demands adapted to their age. All tasks were piloted beforehand. In order to answer the question whether metacognitive skills are general or domain specific, students were asked to perform 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. In order to allow domain-specific metacognitive activities, the differences in tasks and domains were maximized.

#### *4.2.4.1 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 were also asked to read aloud and to 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 a post-test was administered (see section 4.2.5).

The history texts were composed of texts 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 topic on the one hand, and to appeal to the zone of tolerable difficulty (Vygotsky, 1978) on the other hand, the text was based on a subject of the curriculum that was one year ahead (e.g., the first-year text was based on a subject of the second-year curriculum, instead of the first-year curriculum). In all texts, three activating questions or assignments were embedded. These were not meant for testing the students' knowledge, but to elicit text-studying 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 (Meijer, Veenman, & Van Hout-Wolters, 2006) it appeared that if a text did not contain such activating questions and assignments, many participants just tended to read linearly. In the first year, a text about slavery and the civil war in the United States of America was used. In the second year, a text about the First World War was used. Finally, in the third year, a text about politics and economics in the United States of America in the Thirties of the last century was presented.

#### *4.2.4.2 Mathematics task*

In another individual session of 50 minutes, participants practiced to solve mathematical word problems for 20 minutes. Five problems were presented in the first year, six in the second year, and five in the third year. Several categories of problems were presented. In the first year, the categories of problems were distance, fraction, percentage, and surface area of rectangles. In the next year, the categories of problems were content, surface area of a triangle, fraction, percentage, and algebra. In the last year, the categories were calculation of probability, quadratic equation, Pythagoras' theorem, statistics, and formula with a square root. The tasks for each year were composed of adaptations of math problems from one of the most frequently used Dutch schoolbooks for math ("Getal en Ruimte"; Vuijk et al., 2003).

Together with the assignments, participants received a sheet containing the answers and a brief stepwise explanation of problem solutions. 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 parallel problems, which had to be solved in the remaining 30 minutes, without the option to consult an answer sheet. This second part is considered as a post-test assessment of learning performance (see section 4.2.5). All problems had to be solved while thinking aloud.

### *4.2.5 Learning performance*

After the learning phase of both tasks, learning performance was assessed by a post-test, as was explained to the participants in advance. Each year the post-test for history consisted of five multiple-choice questions and six essay questions. The multiple-choice questions assessed reproductive knowledge (facts and dates, e.g., 'What was the name of the Austrian-Hungarian Crown prince?'). The essay questions were meant to assess overall text comprehension (e.g., 'Describe why things went wrong in agriculture and explain what Roosevelt did to restore agriculture economically').

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. Multiple-choice 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 was .58, 1st year; .51, 2nd year; .80 3rd year). For each test the maximum obtainable score was 20 points.

After the learning-by-doing phase for math, learning performance was assessed by the post-test. In each post-test, items were parallel to the items in the learning phase, that is, the surface structure of the post-test items differed from the one of the learning-task items, but the deep structure was the same. Post-tests with five (1st year and 3rd year) and six (2nd year) math word problems were administered. In the first year two points per item could be earned 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 (Cronbach's alpha = .58). Due to an increase of the complexity of the problems, another scoring system was chosen in the second and third year: For the first five math problems in the second year, an equal amount of 10 points could be earned. Problem 6 consisted of three sub-problems that were independent of each other and, therefore, was valued with a maximum of 30 points. So, the maximum obtainable score in the second year was 80 points. A total score was calculated and used as a measure of learning performance in mathematics (Cronbach's alpha = .69). In the third year, five points per (sub)problem could be earned with a maximum obtainable score of 45 points (Cronbach's alpha = .77). Because of the differences in the number of obtainable scores per item over the years, the mean proportion of right answers (*p*-value) was calculated for all questions in each year as well as the mean *p*-value per year[8]. The *p*-values were very similar over the years.

### *4.2.6 Procedure*

Each year, the intellectual-ability tests were administered during a group session. The individual, thinking-aloud sessions took place during school time. The experiment had a counterbalanced design with respect to task order, meaning that half of the participants started with history and the other half with mathematics. Participants could make use of a pen, pencil, text highlighter, ruler, calculator, and blank sheets of paper for making notes. In order to compare the results of the three years, only the data of the participants that performed all tasks over the years ($N$ = 25) were used in the statistical analyses.

All participants were instructed to think aloud while working on the individual tasks. The experimenter refrained from helping students in any way. Whenever a student

[8] Mean p-values for History over the years were .51, .54 and .56 respectively. Mean p-values for Math over the years were .44, .69 and .66 respectively. ANOVA and post-hoc test (Bonferroni) were performed on the math data. No significant difference was found.

fell silent, the experimenter used standard prompts (e.g., 'please, keep on thinking aloud') in order to encourage students to think aloud. All utterances of participants were audio-taped, transcribed, and analyzed on metacognitive skills.

## 4.3 Results

### *4.3.1 Development of metacognitive skills and intellectual ability*

In order to analyze growth in intellectual ability and metacognitive skills, results of the three consecutive years were compared. First, ANOVA was performed on the raw subscale scores of intellectual ability with age as within-subjects factor. A significant effect of age was found [$F(7,18) = 7.40$, $p <.001$, $\eta^2 = .78$]. Pairwise comparisons (comparing the first to the second year, and the second to the third year) showed an incremental change in intellectual ability over the years (see Table 4.2).

**Table 4.2** Means (and standard deviations) of intellectual ability

| **Intellectual ability *N* = 25** | **13 years** | **14 years** | **15 years** |
|---|---|---|---|
| Number series | 18.32 (4.03) | 19.20 (4.23) | 21.04 (3.61) |
| Verbal analogies | 13.88 (3.27) | 16.40 (3.61) | 17.64 (3.14) |
| Unfolding figures | 10.48 (4.67) | 12.72 (4.95) | 14.48 (4.62) |
| Memory | 7.76 (2.22) | 9.22 (2.59) | 9.30 (3.11) |
| Total | 12.61 (2.80) | 14.38 (2.83) | 15.68 (2.73) |

Furthermore, separate ANOVAs with repeated measures were performed on the quantitative and the qualitative subscale scores of metacognition of both tasks with age as within-subjects factor. A significant age effect was found for the frequency of metacognitive skills in math [$F(8,17) = 4.32$, $p < .01$, $\eta^2 = .67$], whereas for history no significant age effect was found [$F(8,17) = 2.14$, $p > .05$, $\eta^2 = .50$]. ANOVAs on the quality of metacognitive skills revealed a significant age effect for both tasks [$F(8,17) = 2.90$, $p < .05$, $\eta^2 = .58$, math] and [$F(8,17) = 3.28$, $p < .02$, $\eta^2 = .61$, history]. Means and standard deviations are shown below in Tables 4.3 (quality) and 4.4 (frequency).

**Table 4.3** Means (and standard deviations) of quality of metacognition

| **Quality of metacognition** ***N*** **= 25** | **13 years** | **14 years** | **15 years** |
|---|---|---|---|
| Orientation Math | 1.52 (.65) | 1.80 (.82) | 1.60 (.71) |
| Planning Math | 1.84 (1.54) | 2.36 (1.10) | 2.48 (.92) |
| Evaluation Math | 1.12 (1.10) | 1.40 (.76) | 1.36 (.99) |
| Elaboration Math | 1.36 (.99) | 1.48 (.92) | 1.08 (.81) |
| Total Math | 5.96 (3.33) | 7.04 (2.61) | 6.52 (2.50) |
| Orientation History | .56 (.58) | .92 (.76) | 1.04 (.79) |
| Planning History | 2.00 (1.50) | 1.64 (.91) | 1.96 (1.27) |
| Evaluation History | .80 (1.00) | 1.44 (1.35) | .72 (.98) |
| Elaboration History | 1.84 (1.55) | 1.96 (1.02) | 2.00 (1.22) |
| Total History | 5.24 (3.77) | 5.88 (3.20) | 5.72 (3.30) |

**Table 4.4** Means (and standard deviations) of quantity (frequency) of metacognition

| **Quantity of metacognition** | **13 years** | **14 years** | **15 years** |
|---|---|---|---|
| Orientation Math | 6.84 (1.95) | 9.44 (3.13) | 6.52 (1.80) |
| Planning Math | 8.40 (4.11) | 12.16 (4.30) | 10.88 (2.74) |
| Evaluation Math | 2.72 (2.50) | 5.04 (4.83) | 2.84 (2.35) |
| Elaboration Math | 3.24 (2.63) | 5.88 (4.00) | 3.68 (2.54) |
| Total Math | 21.20 (8.44) | 32.52 (10.90) | 23.88 (6.58) |
| Orientation History | .24 (.66) | .56 (.87) | .80 (1.04) |
| Planning History | 7.36 (6.73) | 9.00 (7.70) | 8.60 (6.77) |
| Evaluation History | 1.44 (1.98) | 3.56 (3.93) | 1.60 (2.23) |
| Elaboration History | 9.32 (9.44) | 7.40 (4.55) | 8.92 (5.89) |
| Total History | 18.36 (13.83) | 20.52 (12.83) | 19.52 (13.38) |

Pairwise comparisons were performed in order to look closer into changes on subscale level over the years. These tests revealed different developmental patterns at subscale level (see Figures 4.1 and 4.2).

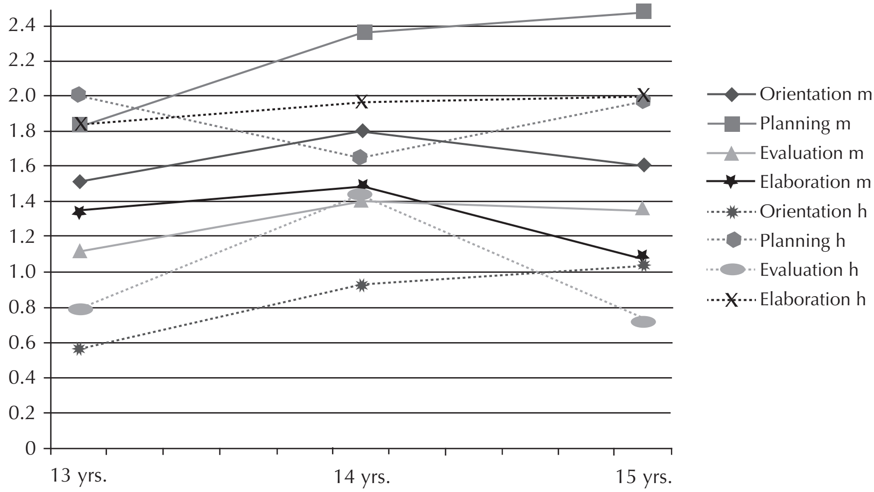


**Figure 4.1** Qualitative metacognition scores across age (m = math; h = history)

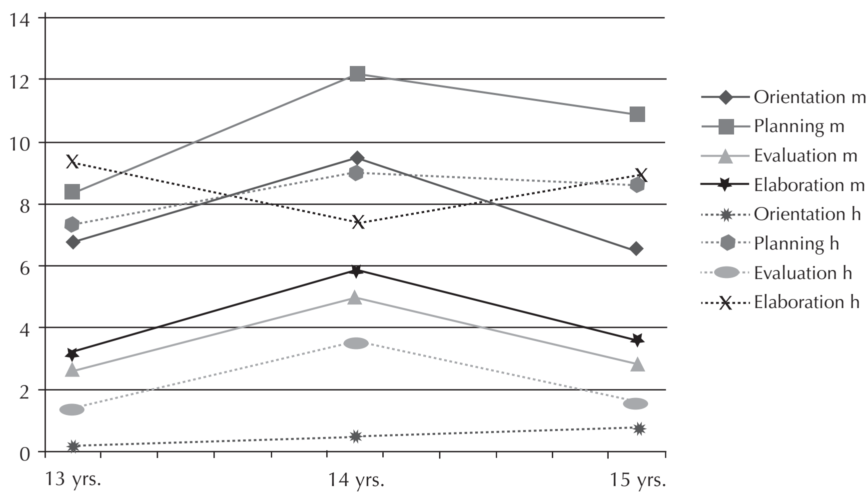


**Figure 4.2** Quantitative metacognition scores across age (m = math; h = history)

The majority of the metacognitive scores did not increase continuously over the years. Two main patterns can be observed: (1) Growth between the first and the second year, followed by stabilization in scores between the second and the third year, and (2) Growth between the first and the second year followed by regression between the second and third year.

### *4.3.2 Testing the mixed model*

To determine whether developmental processes affect the relation between intellectual ability and metacognition as predictors of learning performance, the correlations between these three variables over the three consecutive years were compared (see Table 4.5).

**Table 4.5** Correlations and Semi-partial correlations

| | IA | Meta | Semi-PartIA | Semi-PartMeta |
|---|---|---|---|---|
| **Quality** | | | | |
| LP Math13 | .77** | .70** | .43** | .30* |
| Meta Math13 | .61** | | | |
| LP Math14 | .54** | .74** | .20 | .55** |
| Meta Math14 | .49* | | | |
| LP Math15 | .79** | .59** | .52** | .13 |
| Meta Math15 | .66** | | | |
| LP History13 | .46* | .68** | .28* | .58** |
| Meta Hist13 | .28 | | | |
| LP History14 | .24 | .35 | .16 | .30* |
| Meta Hist14 | .26 | | | |
| LP History15 | .25 | .60** | .00 | .55** |
| Meta Hist15 | .41* | | | |
| **Quantity** | | | | |
| LP Math 13 | | .73** | .40** | .30* |
| Meta Math13 | .64** | | | |
| LP Math14 | | .40 | .46** | .27 |
| Meta Math14 | .24 | | | |
| LP Math15 | | .34 | .70** | .08 |
| Meta Math15 | .42* | | | |
| LP History13 | | .61** | .32* | .51** |
| Meta Hist13 | .26 | | | |
| LP History14 | | .23 | .21 | .20 |
| Meta Hist14 | .15 | | | |
| LP History15 | | .43* | .15 | .38* |
| Meta Hist15 | .23 | | | |

*Note:* LP means Learning performance * $p < .05$, ** $p < .01$

As far as the math task was concerned, intellectual ability correlated significantly with both quality of metacognitive skills and learning performance in the three consecutive years. The same applies for the correlation between quality of metacognitive skills and learning performance. The correlation between frequency of metacognitive skills for math and learning performance was significant only in the first year. The correlation between intellectual ability and frequency of metacognitive skills was significant, except for the second year.

Results on the history task differ partly from results on the math task. Only in the first year a significant correlation was found between intellectual ability and learning performance. The correlation between intellectual ability and quality of metacognitive skills was significant in the third year only. The correlation between quality of metacognitive skills and learning performance was significant in the first and the third year. The same applies for frequency of metacognitive skills and learning performance. No significant correlations were found between intellectual ability and frequency of metacognitive skills for history.

To test the mixed model, semi-partial correlations for each age group (Nunnally, 1967) were calculated by partialling metacognitive skill from the correlation between intellectual ability and learning performance (i.e., Semi-PartIA) and partialling intellectual ability from the correlation between metacognitive skills and learning performance (i.e., Semi-PartMeta). These semi-partial correlations (see Table 4.5) are needed to calculate the unique, independent contribution of metacognitive skills and intellectual ability to learning performance. Using regression-analytic techniques (Pedhazur, 1982; Van der Stel & Veenman, 2008, 2010; Veenman & Spaans, 2005) the unique and shared proportions of variance in learning performance were distributed to metacognitive skills and intellectual ability (see Table 4.6).

History results in Table 4.6 show that, despite the variance shared with intellectual ability, both frequency (QN) and quality (QL) of metacognitive skills, substantially added to the prediction of learning performance on top of intellectual ability. Between 13 and 14 years the unique contribution of metacognition decreased in order to increase again between 14 and 15 years. The unique contribution of intellectual ability to learning performance in history faded out over the years. Math results show an increasing contribution of metacognitive skillfulness to the prediction of learning performance on top of intellectual ability between 13 and 14 years. With 15 years, however, this unique contribution practically disappeared. The unique contribution of intellectual ability to the learning performance on top of the quality of metacognitive skills decreased substantially between 13 and 14 years, however, to reappear with 15 years. In order to check whether the contribution of metacognitive skills differed significantly over the years, Fisher-*z* ratios were calculated for pairs of correlations (Guilford, 1965). All Fisher-*z* ratios were smaller than 1.46, meaning that none of the correlations differed significantly.

**Table 4.6** Percentage of variance accounted for in Learning performance

| | Intel unique | | Meta unique | | Shared | | Total | |
|---|---|---|---|---|---|---|---|---|
| | QL | QN | QL | QN | QL | QN | QL | QN |
| History 13 | 7.9 | 9.9 | 33.0 | 25.9 | 13.3 | 11.1 | 54.2 | 46.9 |
| History 14 | 2.5 | 4.5 | 9.0 | 3.8 | 3.5 | 1.4 | 15.0 | 9.7 |
| History 15 | 0.0 | 2.3 | 30.1 | 14.3 | 6.1 | 3.8 | 36.2 | 20.4 |
| Math 13 | 18.8 | 15.8 | 8.8 | 9.2 | 40.6 | 43.7 | 68.2 | 68.7 |
| Math 14 | 4.2 | 21.0 | 29.7 | 7.5 | 25.0 | 8.3 | 59.0 | 36.8 |
| Math 15 | 26.6 | 51.8 | 1.6 | 0.0 | 38.0 | 12.8 | 66.2 | 64.6 |

*Note:* Intel unique means the unique contribution of Intellectual ability to Learning performance; Meta unique means the unique contribution of Metacognitive skills to Learning performance; Shared means the shared contribution of Intellectual ability and Metacognitive skills to Learning performance. Total means the total contribution of Intellectual ability and Metacognitive skills to Learning performance. QL = qualitative metacognition scores; QN = quantitative metacognition scores.

### *4.3.3 Metacognitive skills across domains*

The generality vs. domain specificity of metacognitive skills was investigated by performing a principal component analysis (PCA) on the metacognition scores. For each year separately, a principal component analysis with a two-factor solution was performed on the four subscales of metacognitive skills for both tasks (see Table 4.7).

The unrotated solutions of the PCAs show that all measures of quality of metacognitive skills substantially load on the first component (see Table 4.4). This component has eigenvalues of 3.53 (13 yrs.), 3.28 (14 yrs.), and 3.87 (15 yrs.), with variance proportions of .44; .41 and .48 respectively. Moreover, in the first two years a second component contrasting the two domains was extracted with eigenvalues of 1.78 (13 yrs.), and 1.40 (14 yrs.), and with variance proportions of .22, .17 respectively. Loadings on a second component of the third year with an eigenvalue of 1.11 (15 yrs.), and a variance proportion of .14 did not contrast the two domains (see Table 4.7). The same analysis was performed on quantitative scores of metacognitive skills. Results were in line with those of the qualitative data.

**Table 4.7** Unrotated component matrix for the quality of metacognitive skillfulness

| | Component1 13 yrs | Component2 13 yrs | Component1 14 yrs | Component2 14 yrs | Component1 15 yrs | Component2 15 yrs |
|---|---|---|---|---|---|---|
| Eigenvalue | 3.53 | 1.78 | 3.28 | 1.40 | 3.87 | 1.11 |
| Variance proportion | .44 | .22 | .41 | .17 | .48 | .14 |
| Metaorientation Hist | .50 | .49 | .51 | -.45 | .21 | .86 |
| Metaplanning Hist | .72 | .48 | .49 | -.46 | .86 | .19 |
| Metaevaluation Hist | .48 | .52 | .77 | -.42 | .71 | -.44 |
| Metaelaboration Hist | .72 | .37 | .89 | -.14 | .87 | .04 |
| Metaorientation Math | .61 | -.57 | .40 | .23 | .46 | -.27 |
| Metaplanning Math | .78 | -.31 | .58 | .61 | .73 | .04 |
| Metaevaluation Math | .55 | -.67 | .69 | .13 | .68 | -.17 |
| Metaelaboration Math | .86 | -.16 | .63 | .59 | .79 | .17 |

### *4.3.4 Exploratory analysis*

After the individual math sessions of the third year were completed, the experimenter had the impression that the frequency of using the step-wise explanation sheet during the math task had changed over the years. Together with the assignments, participants received this sheet containing the answers and a brief stepwise explanation of problem solutions. Participants were free to consult this sheet whenever and as much as they liked. The impression was that older participants made more use of the sheet to check their own solutions, that is, older participants would consult the sheet more *after* solving a math problem, relative to younger participants. Change in the frequency of using the sheets *prior* to attempts to solve a problem was not suspected. A change in using the sheet could have caused the unexpected low unique contribution of metacognition to learning performance for math in the 3$^{rd}$ year. In order to test this assumption, the sheet use prior and after problem solving was coded in all math protocols. Because the number of math assignments was not the same each year, the sheet-use scores were recoded in a relative frequency of sheet usage. ANOVA with repeated measures on consulting the explanation sheets *after* problem solving with age as within-subjects factor revealed a significant effect of age [$F(2,23) = 6.82$, $p<.005$, $\eta^2 = .37$]. Fourteen-year-olds consulted the sheets more than 13-year-olds, and 15-year-olds more than 14-year-olds. The mean percentage of assignments for which 13-year-olds consulted the sheet after problem solving was 10 ($SD = 23.9$); 14-year-olds ($M = 24$, $SD = 28.7$);

15-year-olds ($M$ = 41, $SD$ = 33.2). ANOVA on the scores *prior* to problem solving did not reveal a significant age effect [$F(2,23) = .04$, $p > .05$, $\eta^2 = .00$]. Each year, the mean percentage of assignments for which the sheet was consulted prior to problem solving was about 30.

## 4.4 Discussion

This longitudinal study investigated the development of both the quantity (frequency) and the quality (depth) of metacognitive skills in young adolescents. According to the first hypothesis, metacognitive skills were expected to increase in frequency as well as in depth over the three consecutive years. Results show an overall growth of quantitative and qualitative scores of metacognitive skills over the first two years (between 13 and 14 yrs.). Between the second year and the third year (between 14 and 15 yrs.), this growth did not continue. So, the first hypothesis is partly corroborated: Metacognitive skills do grow in frequency and in depth, but growth was not continuous over the three consecutive years. In prior research (Veenman & Spaans, 2005; Veenman et al., 2004) a continuous growth of metacognitive skillfulness was found. It should be mentioned, however, that these studies were concerned with larger intervals between measurements, and did not focus on the development between 14 and 15 years in particular. Moreover, these studies had a cross-sectional design and the same tasks were used over the years. Therefore, it is presumed that growth is arrested only temporarily between 14 and 15 yrs. According to dynamic-systems theories (Siegler, DeLoache, & Eisenberg, 2010), a class of theories that focus on how change occurs over time in complex systems, individual children acquire skills at different ages and in different ways. Their development entails regressions as well as progress. Development of metacogitive skills seems to be in line with the notion of dynamic-systems theories: During development both progress and regression occur, and not all components of metacognitive skillfulness develop at the same pace.

The second hypothesis concerned the relation between metacognitive skills and intellectual ability as predictors of learning performance over age groups. Results were expected to be in line with the mixed model and the monotonic development hypothesis. In the present study, a unique contribution of metacognitive skills to learning performance and a shared contribution of metacognitive skills with intellectual ability to learning performance was found in all three consecutive years, with the exception of the frequency of metacognitive skills in math in the 3rd year. The unique contribution of the quality of metacognitive skills in math in the 3rd year was rather small (1.6%). In a cross-sectional study with the same tasks and the same age groups, however, a much higher unique contribution of metacognitive skills (42.8%) in 15-year-olds was found (Van der Stel, et al., 2010). The small unique contribution in the present study could be the result of the more frequent use of the explanation sheets in the 3rd year. Consulting the sheet after problem solving, that is, comparing the solution given on the sheet with one's own solution, could be acquired behavior promoted by the teachers of this particular group. It goes without doubt that this is useful learning behavior, but this acquired behavior probably suppressed the unique contribution of spontaneous metacognitive skills during problem solving. In conclusion, the mixed model was found over the years for both history and math, albeit less convincing than expected for math in the 3rd year.

Another part of hypothesis 2 was the expected monotonic development of metacognitive skills and intellectual ability as predictors of learning performance in line with Alexander's monotonic development hypothesis. Results of the first two years of the present study point in the direction of a monotonic development of metacognitive skills: A continuous growth of metacognitive skills with age was found, alongside intellectual growth. Results of the 3rd year, however, show a continued growth in intellectual ability, but no further growth in metacognitive skills. Despite the (temporary) stabilization in metacognitive growth, results predominantly agree with the monotonic development hypothesis. Here it is hypothesized that intellectual development does not direct metacognitive development, as is the case in both the ceiling and the acceleration hypotheses (see section 4.1.3). Moreover, the mixed model was found each year, confirming an independent contribution of metacognition to learning performance (except for the quantity of metacognition in 3rd year math). Thus, intellectual development does not direct metacognitive development. Therefore, the current results are considered to agree most with the monotonic development hypothesis, which is relevant to the training of metacognitive skills in education. Metacognitive skills can be trained successfully at different ages and in various tasks and domains (cf. Campione, Brown & Ferrara, 1982; Chinnappan & Lawson, 1996; Kramarski & Mevarech, 2003; Masui & De Corte, 1999; Pressley & Gaskins, 2006).

The third and last research question concerned the generality vs. domain specificity of metacognitive skills. In the first two years, the solutions of the principal component analysis (PCA) on the metacognitive data show a highly similar two-component solution: A first component with rather high component loadings, which may be interpreted as representing general metacognitive skills across domains, and a second, weaker component with contrasted component loadings, which may be interpreted as representing domain-specific metacognitive skills. In the third year, however, the solution of the PCA changed: The first component still can be interpreted as representing general metacognitive skills, but the structure of the second component has become much more scattered. It no longer can be interpreted as a domain-specific component. Results support our expectation that metacognitive skills of rather young and inexperienced adolescents represent a general as well as a domain-specific component. Veenman and Spaans (2005) assumed that metacognitive skills initially develop on separate islands of tasks and domains. They also assumed that beyond the age of 12 years, these skills merge into a more general repertoire that is applicable and transferable across tasks and domains. Among young adolescents, a phase of transition could be characterized by applying recently acquired general metacognitive skills, along with a remainder of domain-specific metacognitive skills. In line with hypothesis 3, it was expected that the initially acquired domain-specific metacognitive skills would tend to generalize during development. Although the present results corroborate hypothesis 3, the generalization process was less gradual than expected. Next to drawing on a repertoire of general metacognitive skills, students continue to apply domain-specific metacognitive skills between the age of 12 and 14 years. This may indicate that these students are still in a transitory phase of metacognitive-skill development. The use of both general and domain-specific metacognitive skills for a longer period of time could be explained by the overlapping-waves model (Siegler, 1998). According to this model, children initially use multiple strategies, and with age and experience they will selectively rely on more advanced strategies.

From an instructional perspective, it would be advisable to extend the training of domain-specific metacognitive skills in a particular learning context to more general, domain-surpassing ones (Veenman et al., 2004). Students will profit from an explicit training in metacognitive skills more effectively if that training surpasses a particular learning context. Metacognitive skills, acquired in separate domains, may gradually be generalized across domains (Schneider & Pressley, 1997). This process can be considered as high road transfer (Salomon & Perkins, 1989). Teachers should initially encourage students to develop their domain-specific metacognitive skills. As a next step, teachers should pay attention to the generalized applicability of the students' repertoire of metacognitive skills across domains. If teachers from various disciplines attune their instructions regarding metacognitive skills, transfer of metacognitive skills could be facilitated, thus providing students with tools for performing new tasks in new domains. It should be acknowledged that there are some limitations to the present study. A first limitation concerns the generalizability due to the small sample size. The timeconsuming method of protocol analysis of individual sessions did not allow for larger samples. A second limitation could be the dissimilarity in tasks over the years. Repeatedly measuring learning performance in a longitudinal design, however, makes it inevitable to use *new* tasks each year. By piloting the tasks and consulting teachers, efforts were made to balance the relative difficulty level of the tasks for each age group.

During the last decade, neurocognitive developmental research showed that changes in the adolescent brain are non-linear, non-synchronous and with large individual differences (Casey, Getz & Galvan, 2008; Steinberg, 2005; Toga, Thompson & Sowell, 2006). The prefrontal cortex matures until late adolescence (Toga et al., 2006). Veenman et al. (2004) found that a continued growth also applies to metacognitive skills, they continue to develop till at least the age of 22 yrs. The current study show that different components of metacognitive skillfulness develop neither at the same pace, nor continuously. Knowledge about the developmental trajectory of the various components of metacognitive skillfulness will enable teachers to teach the right things at the right time. For future research it would be a challenge to look for opportunities to combine research that describes developmental changes based on behavioral experiments with research based on new methods like functional imaging. Such studies would make it possible to focus on *processes* of change rather than focusing on steady states at different ages.

## 4.5 Conclusions

Results show that the age of 15 years is a significant point in time during the developmental trajectory of metacognitive skills. It seems that the development of metacognitive skills of students between 12 and 14 years is dominated by growth, resulting in an increase in both frequency and quality of metacognitive skills. Once this growth of metacognitive skills has developed up to a certain level, development of metacognitive skills is not longer dominated by growth, but by the generalized application of these skills. Students around the age of 15 yrs. will increasingly be able to transfer the metacognitive skills that were acquired in certain tasks and domains to new tasks and domains. The present study shows that the age of 15 yrs. is a significant landmark in the development of metacognitive skillfulness on the way to adulthood.