Psychological Topics 23 (2014), 1, 115-141

Original scientific paper – UDC – 159.953.5.072-053.5
37.015.3

Metacognitive Knowledge in Relation to Inquiry Skills and Knowledge Acquisition Within a Computer-Supported Inquiry Learning Environment

Zrinka Ristić Dedić

Centre for Educational Research and Development,
Institute for Social Research in Zagreb, Croatia

Abstract

The study examines two components of metacognitive knowledge in the context of inquiry learning: metatask and metastrategic. Existing work on the topic has shown that adolescents often lacked metacognitive understanding necessary for optimal inquiry learning (Keselman & Kuhn, 2002; Kuhn, 2002a; Kuhn, Black, Keselman, & Kaplan, 2000), but demonstrated that engagement with inquiry tasks may improve it (Keselman, 2003; Kuhn & Pearsall, 1998).

The aim of the study is to investigate the gains in metacognitive knowledge that occur as a result of repeated engagement with an inquiry learning task, and to examine the relationship between metacognitive knowledge and performance on the task.

The participants were 34 eighth grade pupils, who participated in a self-directed experimentation task using the FILE programme (Hulshof, Wilhelm, Beishuizen, & van Rijn, 2005). The task required pupils to design and conduct experiments and to make inferences regarding the causal structure of a multivariable system. Pupils participated in four learning sessions over the course of one month. Metacognitive knowledge was assessed by the questionnaire before and after working in FILE.

The results indicate that pupils improved in metacognitive knowledge following engagement with the task. However, many pupils showed insufficient metacognitive knowledge in the post-test and failed to apply newly achieved knowledge to the transfer task. Pupils who attained a higher level of metacognitive knowledge were more successful on the task than pupils who did not improve on metacognitive knowledge. A particular level of metacognitive understanding is a necessary, but not sufficient condition for successful performance on the task.

Keywords: metacognitive knowledge, inquiry learning, inquiry skills, metacognitive skills, knowledge acquisition

✉ Zrinka Ristić Dedić, Institute for Social Research in Zagreb, Centre for Educational Research and Development, Amruševa 11, 10000 Zagreb, Croatia. E-mail: *zrinka@idi.hr*

Introduction

The present study examined the relationship between metacognitive knowledge and strategic performance within the context of an inquiry learning domain. Inquiry learning, also called inductive learning by some authors (Beishuizen, Wilhelm, & Schimmel, 2004; De Jong & Van Joolingen, 1998; Veenman & Spaans, 2005; Veenman, Wilhelm, & Beishuizen, 2004; Wilhelm & Beishuizen, 2003), is defined as an educational activity in which learners investigate real or virtual phenomena and draw conclusions based on what they have learned regarding the causal status of the features in a multivariable system (Kuhn, Black, Keselman, & Kaplan, 2000). In a typical inquiry learning task, learners participate in the whole scientific inquiry cycle, from formulating investigative questions and stating hypotheses, through experimentation and validation of evidence, to making causal inferences and re-building theories. Learners typically direct their own investigatory activities and conduct experiments in the absence of instruction or explicit feedback regarding the accuracy of their conclusions or the validity of their approaches to the task. However, learners may be prompted by a researcher to formulate investigative intent, express their expectations, or draw and justify conclusions (Kuhn, Garcia-Milà, Zohar, & Andersen, 1995). Through engagement with inquiry learning tasks that resemble real scientific research in its simplest, generic form, learners construct and expand their understanding of the scientific content, while at the same time developing inquiry skills and an understanding of the nature of scientific knowledge and scientific thinking (Ben-David & Zohar, 2009; Kuhn, 2001, 2002a, 2002b).

Due to the complexity of the inquiry learning task, and its qualities that require learners to take a systematic and reflexive approach in order to draw accurate and valid conclusions, inquiry learning is a domain in which metacognition is especially important. In most conceptualisations of metacognition, a distinction is usually made between metacognitive knowledge and the regulation of cognition (Brown, 1978; Schraw & Dennison, 1994).

Metacognitive knowledge refers to memory-retrieved declarative knowledge about the interplay between person, task and strategy characteristics (Flavell, 1979). It can be retrieved from long-term memory as a result of a purposeful and deliberate search, but may also be activated unintentionally based on cues present in the learning situation (Efklides, 2006).

Unlike declarative metacognitive knowledge, the regulation of cognition is the procedural component of metacognition, often referred to as metacognitive skills (Veenman, 2011). These skills pertain to the deliberate use of a person's procedural knowledge for monitoring, guiding and regulating one's learning activity and problem-solving. These skills have a built-in feedback mechanism (Veenman, Van Hout-Wolters, & Afflerbach, 2006) and are amenable to practice and automatization (Veenman & Elshout, 1999).

In the context of inquiry learning, two components of metacognitive knowledge are especially relevant: knowledge of task objectives (metatask) and knowledge of strategies (metastrategic) (Kuhn, 1999, 2000, 2001, 2002b; Kuhn & Pearsall, 1998).

The metatask component relates to an understanding of the nature and requirements of the task, i.e. knowledge of task features and demands that call for the use of a particular strategy. In an inquiry learning task, metatask understanding refers to representation of the task objective as one of identifying the causal structures of the investigated phenomena and examining the effects of individual variables.

The metastrategic component refers to an understanding of the strategies available in one's repertoire of strategies that are potentially applicable to the task and that are likely to succeed in achieving the task objective. As such, it involves both declarative knowledge of potential strategies and the conditions for their use. In the context of inquiry learning, where a control-of-variables strategy (CVS) has a central role, metastrategic knowledge includes recognizing the necessity for a control-of-variables strategy if the resulting causal inferences are to be valid, as well as recognizing the inadequacy of other invalid experimentation strategies.

In Kuhn's (1999, 2001, 2002b) conception, meta-level knowledge is procedural and operates in the real selection and regulation of the use of inquiry strategies. Indeed, in order to emphasize the procedural nature of the meta-level, Kuhn actually used the term 'metastrategic knowing' instead of 'knowledge' in most of her work (Kuhn, 1999, 2001, 2002b). Although this meta-level is distinct from the performance level at which the actual exercise of strategies occurs, the changes at the performance level are determined and mediated by the meta-level of thinking (Kuhn et al., 1995, 2000; Kuhn & Pearsall, 1998). While meta-level knowledge is governing the selection and application of strategies, feedback from the performance level also informs the meta-level and leads to enhanced meta-level awareness and understanding of the task purpose and the strategies. This increased understanding at the meta-level guides subsequent improved strategy selection and gradually leads to changes in the distribution of the strategies observed at the performance level in a continuous cycle in which the meta-level both guides and is modified by the performance level (Kuhn, 2001, 2002a, 2002b, 2005; Kuhn & Pearsall, 1998).

Zohar and Ben David (2008) and Zohar (2012) analysed Kuhn's conception of metastrategic knowledge in relation to other relevant conceptions in the field. They claimed that this concept addresses "the when, why and how" of using a thinking strategy, which corresponds to what Flavell termed knowledge about tasks and knowledge about strategies (Flavell, Miller, & Miller, 2002). Kuhn's conception is similarly related to Schaw's categories of procedural and conditional metacognitive knowledge (Schaw, 1998; Schaw & Dennison, 1994). However, Veenman (2011) opposed such a conception and claimed that metacognitive knowledge encompasses only declarative knowledge and (declarative) conditional knowledge about the utility of strategies, but not the actual use of those strategies. The same author argued that Kuhn's notion of metastrategic knowledge actually obscures the boundary between metacognitive knowledge and skills.

Although a prerequisite for successful performance on the task, activation of metacognitive knowledge does not necessarily lead to appropriate strategic response to the task and is not sufficient for mastery in strategic performance, partly because metacognitive knowledge may be incorrect or incomplete (Efklides, 2009; Veenman, 2011). Furthermore, even having adequate metatask and metastrategic understanding does not necessarily ensure the adequate execution of an appropriate strategy (Veenman, 2011, 2012). Metacognitive knowledge, being declarative in nature, cannot be readily applied on task processing (Efklides, 2009). Instead, procedural knowledge, i.e. the presence of metacognitive skills, is necessary for the adequate execution of appropriate strategies (Veenman et al., 2004). In an inquiry learning task, it is possible that the pupil knows that it is necessary to use a control-of-variables strategy, but fails to apply it if he/she lacks the motivation or procedural knowledge about how to execute these strategies and how to monitor and regulate their application. As a result, Veenman (2011, 2012) concluded that metacognitive knowledge often poorly predicts learning outcome, and has no effect on behaviour until it is actually being used.

Nevertheless, the importance of metacognitive knowledge is reflected in the fact that it directs the interpretation of a learning situation and task demands and provides a database from which the learner can retrieve information about which strategies to apply for the control and regulation of learning in a particular task (Efklides, 2006, 2009). This suggests that metacognitive knowledge can control strategic performance indirectly through expectations that influence the interpretation of situational and task demands (Efklides, 2009).

Previous research in the field of inquiry learning consistently demonstrates that pupils experience problems in every phase of the scientific inquiry cycle and suggests that their difficulties are not limited to performance level, but extend into the area of meta-level understanding and the explicit management of one's strategic performance (Keselman, 2003; Keselman & Kuhn, 2002; Kuhn et al., 2000; Zohar & Peled, 2008). Although learners lack the metacognitive understanding necessary for optimal inquiry learning at the beginning of the learning process, microgenetic studies have demonstrated that pupils show improvement in inquiry skills and metacognitive functioning when they have an opportunity to participate repeatedly in situations that demand the application of these skills, even without direct instruction and with no feedback beyond that provided by the pupils' own activity (Kuhn, 1995, 2001, 2002a; Schauble, 1990, 1996).

The present study is focused on metacognitive knowledge and the extent to which it is influential in inquiry learning task performance. The definition of metacognitive knowledge used in the study addresses both metatask and metastrategic components in the context of a control-of-variables strategy and accepts the notion of its declarative nature up to the point at which it is used for the management and regulation of one's learning process.

The study examined the manner in which metacognitive knowledge is acquired and used for guiding the learning process in an educational context that otherwise provided only scarce opportunities for pupils' participation in inquiry tasks, and in self-directed experimentation tasks in particular.

The relevance of the examination of metacognitive knowledge lies also in the observation that, in the domain of inquiry learning, metacognitive knowledge is of importance in its own right. Indeed, it is a key educational objective in science education related to the development of an understanding of scientific methods and processes of scientific knowledge acquisition.

One of the goals in the current work was to investigate the gains in metacognitive knowledge that occur as a result of pupils' repeated engagement with an inquiry learning task. It was expected that initial levels of metacognitive understanding would be low due to pupils' lack of experience in dealing with inquiry learning tasks, but that they would improve significantly. Another goal was to examine the role of metacognitive knowledge in the process of knowledge acquisition and in the development of inquiry and metacognitive skills. To do so, we examined whether pupils who acquired higher metacognitive knowledge during inquiry learning process were more successful: a) at the metacognitive level, in planning and organising the activities on the task, b) at the performance level, in the use of the control-of-variables strategy, c) in making valid inferences regarding the (non)causal status of independent variables in the system, and d) in the acquisition of knowledge about the system, reflected in correct conclusions about its causal structure.

Method

Participants

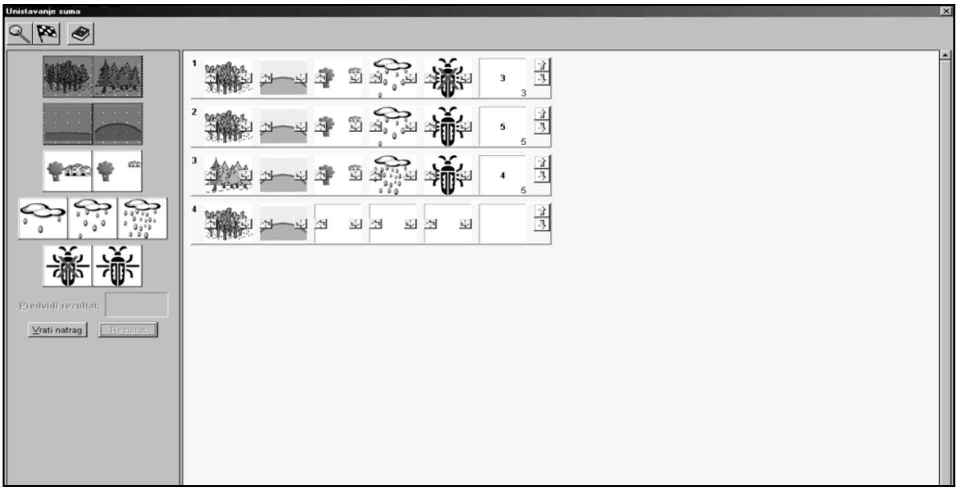
Participants were 34 eighth grade pupils from two primary schools in Zagreb that were chosen purposefully. The selection of participants was based on purposive (maximum variation) sampling (Patton, 1990), with the aim of selecting a heterogeneous group of pupils by purposefully picking a wide range of variation on dimensions of interest for the topic of inquiry learning. The participants were selected from a whole cohort of 114 pupils based on: cognitive abilities (inductive reasoning measured by TN-10, Pogačnik, 1997), self-reported Interest for science measure[1] (Ristić Dedić, 2010) and gender. The participating pupils covered a relatively broad range on intelligence and interest for science measures. The distribution of gender was also balanced.

The Task

The task used in the study was a self-directed experimentation task in the domain of biology. The content of the task was specifically designed for the purposes of the study. The task was loosely related to the sixth grade curriculum for the subject *Nature*, in that it was based on its topic "Benefits of forests, forest devastation and protection" (MZOS, 2006). It was not assumed, however, that the variables used in the task were covered in detail during regular teaching hours or that classroom teaching was based on the principles of inquiry learning. By choosing task content that was thematically connected to curriculum content, it was ensured that participants were at least partially acquainted to the theme and that they had some prior theories about the examined variables and their effects.

The task was presented in a computer-supported environment using the FILE software (*Flexible Inquiry Learning Environment*, Hulshof, Wilhelm, Beishuizen, & van Rijn, 2005; Wilhelm, Beishuizen, & Van Rijn, 2005) that was translated and adapted for the purposes of the study. The interface of the "Forest Devastation" task in FILE is presented in Figure 1.

Figure 1. *The Interface of the FILE Task "Forest Devastation"*



[1] The Interest for science scale assessed pupils' interest in science in general and included items such as *I would like o be a scientist, I like reading about science topics, Carrying out experiments is fun* etc. The split on the scale was based on the median score. On the intelligence test the split was based on a division of results into thirds.

The task required participants to design and conduct experiments and to make inferences regarding the causal structure of a multivariable causal system. The task presented five independent variables with discrete levels or categories and one dependent variable. The participants conducted experiments by choosing and varying the levels of independent variables and observing the effects on the dependent variable. By doing so, the participants were in a position to make inferences about the relationships between independent and dependent variables.

The task of the participants in the study was to determine if and how the independent variables affected forest devastation. The dependant variable was measured on a scale from 1-5 (1 – *very low level of forest devastation* to 5 – *very high level*). Independent variables were: 1) type of forest (deciduous or evergreen), 2) configuration of the land (lowland or hills), 3) distance from a settlement (close or far), 4) frequency of acid rain (rarely, often, and very often) and 5) presence of blight (yes or no).

The model underlining the relationships between independent variables and the dependent variable was such that:

a) Two variables ('configuration of the land' and 'distance from settlement') had no effect on the outcome),
b) Two variables ('presence of blight' and 'type of forest') were causal, but they interacted with one another: at one level of the first variable, both categories of the second interacting variable yield an identical result, while at the second level of the variable, there is a difference between the two categories of the second variable.
c) One variable had a curvilinear effect: two levels of the variable 'frequency of acid rain' result in the same outcome value, while the third level yields a different result.

Tasks of the same structure (i.e., same number of variables and types of relationships between variables) have been used in several studies examining inquiry learning amongst pupils in primary schools (e.g. Keselman, 2003; Kuhn & Katz, 2009; Kuhn & Pearsall, 1998; Kuhn et al., 1995, 2000; Kuhn, Pease, & Wirkala, 2009; Veenman & Spaans, 2005; Veenman et al., 2004; Wilhelm & Beishuizen, 2003).

The choice of variables for the task model was based on a pretest examination with another group of eighth grade pupils in another school. It was important to ensure that the proposed variables and relationships between variables were plausible to participants, and that the task causal model corresponded to possible real-life phenomena. In the pretest, prior theories of pupils were tested in order to construct the task in a way that the task causal model supported and challenged prior causal and non-causal theories of the majority of pupils.

Procedure

An introductory group meeting was organised with participants to demonstrate the FILE program. Participants were told that the study examined how primary school pupils solved one computer-supported task that might help them develop thinking skills. Participants were introduced to the FILE program through a task about a boy named Peter who is late for school. This task, which had the same interface and structure as the "Forest Devastation" task, has been previously applied in several studies (e.g. Hulshof, 2001; Wilhelm & Beishuizen, 2003). The process of conducting the experiments in FILE was demonstrated to participants by conducting two random experiments (uncontrolled comparison). Special attention was paid not to point to or suggest correct or valid strategies for working on the task. Participants were told that learners have to find out about relationships between presented features and Peter's arrival time at school.

Following this demonstration, a metacognitive knowledge questionnaire was applied.

After this introductory group meeting, all other sessions were individual. All participants performed the FILE task during regular school time. In order to track changes in the processes of inquiry skill development and knowledge acquisition, the "Forest Devastation" task was applied in four learning sessions over the course of one month.

Full description of the task interface and the procedure applied for conducting the experiments is provided in Ristić Dedić (2010). The task required participants to actively engage in various stages of inquiry processes: from investigation intent, through hypothesis generation, experiment design and prediction of outcome, to validating evidence and making inferences.

The "Forest Devastation" task was presented to participants in the context of a story about a group of researchers who are working on a forest protection project. The participants were given the task of helping researchers by finding out the level of forest devastation in relation to features identified as having possible effects. The participants were told that they had to find which features do or do not have an influence on forest devastation.

During the learning sessions, participants autonomously made decisions about the features examined in each experiment, the order and the total number of experiments per session.

Participants were additionally instructed to think aloud during the process of inquiry. Web-camera, digital recorder as well as computer log-files were used to record participants' activities on the task. The researcher also asked questions before, during and after each conducted experiment. The questions were the following: a) before each experiment: *What do you want to find out now? How will you do it?*; b) after choosing the features: *What do you expect to find out?*; c) after inspecting the outcome: *What did you find out? How did it go? / What can you conclude about the effect of the features? / How do you know?*

At the end of each learning session, participants were asked to express their theories regarding the effects of each feature. After the final learning session, the metacognitive knowledge questionnaire was applied again.

Instruments/Measures

1. Metacognitive Knowledge Questionnaire

The metacognitive knowledge questionnaire was applied twice, following introductory meeting and immediately after the final learning session. This paper-and-pencil measure of metatask and metastrategic knowledge was developed specifically for the purposes of the study, but was based on measures and procedures used by Keselman (2003) and Kuhn and Pearsall (1998).

The questionnaire had two sections:

a) Understanding of task goals and strategies

In the first section, metacognitive knowledge was assessed through two direct measures (open-ended questions) that were framed as an imaginary conversation between the pupil and another class colleague. In this scenario, the participant was asked to explain their conceptions about the task goal and strategies needed to meet the goal for the task introduced to participants at the beginning of the study:

> *Imagine the following situation: "You talk to your colleague in another class about what we did during our meeting. You explain that you are going to participate in the study by working on a computer task. Your colleague is asking you…".*
>
> *What do you need to do on the task? What is the goal here, what do you need to achieve?* (Question examining participants' understanding of task goal).
>
> *How do you decide which features to choose for each line (experiment) to solve the task successfully?* (Question examining the inquiry approach participants deemed appropriate for meeting the task goal).

In the second application of the questionnaire, the introductory scenario was changed slightly, but the questions regarding task goal and strategies remained the same:

> *Throughout the research, you developed some research skills and learned something about how to run experiments. Now, you are likely in a position to share your knowledge with other class colleagues. Please recall what you did on the task and imagine a situation in which you have to show the task to your colleagues and teach them how to solve it.*

Coding of the Responses

The coding scheme developed by Kuhn and Pearsall (1998) was used for the coding of participants' responses to these questions examining metatask and metastrategic knowledge. While these two components were coded separately, levels of attainment for each component were conceptually equated across components (e.g. *Level 1* in both components denoted an orientation towards positive outcomes). Table 1 presents these coding schemes.

Table 1. *Metatask and Metastrategic Levels (Adapted from Kuhn and Pearsall, 1998)*

Level	**Metatask**		**Metastrategic**	
	Description of the level	Typical response	Description of the level	Typical response
0	No insight into the goal of the task, procedural description	*Select the little images and click on "result" button*	No awareness of the need for a strategic approach (Procedural description)	*Select whatever you like, You click on the little images*
1	Orientation towards attainment of positive outcome	*Help Peter to arrive at school on time*	Choosing instances believed to yield a positive outcome	*Select icons according to how these things could help Peter to be fast*
2	Analysis at instance level (Find out the result for different combinations of features)	*Examine how late Peter would be in different cases*	Choosing different instances to observe outcomes	*Try different combinations and check the results*
3	Analysis at feature level, but without its isolation (Find out which features are making a difference in outcome)	*Find out what influences how much Peter would be late/ uncover the reasons for Peter's delayed arrival to school*	-	-
4	Analysis at feature level with reference to multiple features (Find out whether features 1,2 and 3 are making a difference)	*Find out what difference is made when Peter makes choices (type of shoes, if he is having breakfast, how he goes to school, etc.)*	Comparing instances in an uncontrolled fashion (multiple features are varied)	*Compare how late Peter is depending on what he chooses e.g. what kind of breakfast he had, how he went to school, etc.*
5	Analysis at feature level with focus on single feature at a time (Find out whether feature 1 is making a difference)	*Find out how each feature influences the time when Peter arrives at school*	Comparing instances in which a single feature is varied and other features are not mentioned	*Compare when Peter gets to school using a running bike or city bike*

Table 1. - *Continued*

Level	Metatask		Metastrategic	
	Description of the level	Typical response	Description of the level	Typical response
6	-	-	Comparing instances in which a single feature is varied and other features are held constant	*Change just one element e.g., types of shoes, to be able to conclude if it makes a difference*

The participants' responses were double coded to check inter-rater reliability. Cohen Kappa coefficients ranged between .79 and .94, and indicated acceptably high values in comparison to benchmarks set by Landis and Koch (1977). Percentage agreements between the two coders were between 85.3% and 97.1%.

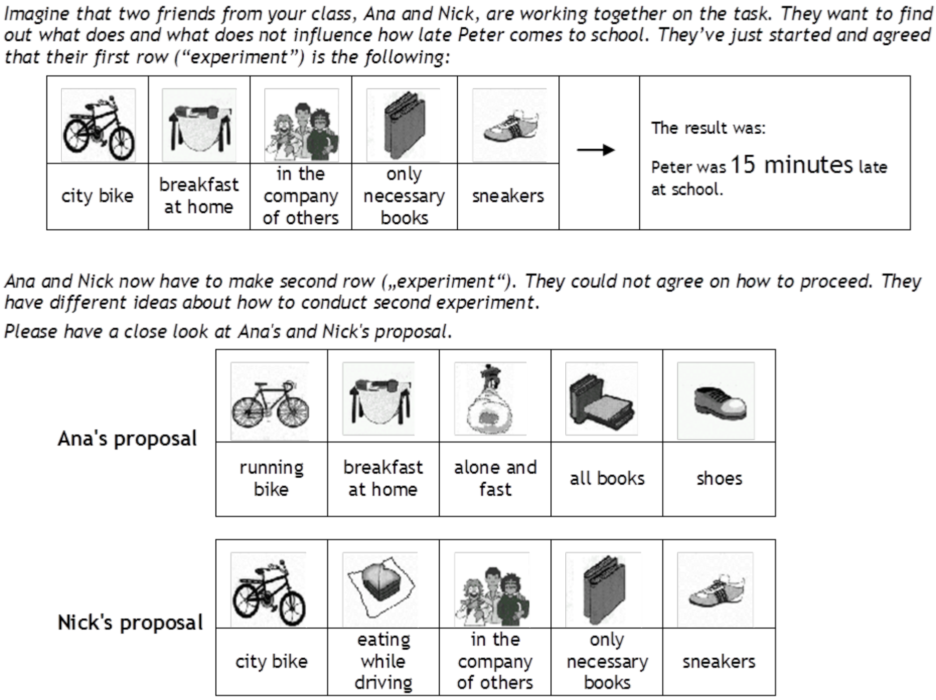
b) Application of metacognitive knowledge

The second section of the questionnaire tested the application of metacognitive knowledge on the "Peter" task. As this task had the same structure as the main "Forest Devastation" task but was different in content, it acted as a transfer measure of metacognitive knowledge.

Participants had to evaluate the strategies applied to the task by two imaginary pupils who are working together on the task. In this scenario, the pupils disagree about how to design the second experiment and which features to select for the combination of independent variables. The first pupil suggests using a combination that enables a valid comparison with the first experiment (i.e., corresponds to the control-of-variables strategy), while the second pupil suggests using a combination that corresponds to the invalid experimentation strategy 'Hold one thing constant at a time'.

The following format was used:

Figure 2. *The Transfer Task*



The participants had to decide which proposal is better and why (in the form of a response to an open-ended question). They also had to determine what could be revealed about the causal status of features if they followed Ana's or Nick's proposal. These questions examined participants' understanding of the value of control-of-variables strategies for making inferences about relationships between variables. They also indirectly assessed participants' understanding of the goal of the task as an inquiry task.

The coding scheme applied to the responses in this section was developed in an exploratory (inductive) manner, based on the collected participants' responses. For responses regarding the participants' preferred proposal, the following scheme was used:

0 – No explanation

1 – Non-analytic response/ theoretic explanation (*Peter should not hurry to school, it might be dangerous for him*)

2 – Dominance of participants' theories over evidence: (Participants preferred proposal that has more favourable combination of features, according to participants' theories / Participants did not show preference because in their opionion both proposals bring similar result or some other combination of features was favoured)

3 – Understanding of the value of CVS.

The coding scheme applied to responses to the question *What could you find out if you followed Ana's/Nick's proposal?* was the following:

1 – Correct (*Whether feature X influences the result or not* – when CVS was used; *Nothing* – when CVS was not used)

0 – Incorrect.

2. Task Performance

Task performance was measured through a set of related, but separate measures of inquiry skills, metacognitive skills applied to the task and acquired knowledge about the causal model.

Inquiry Skills

Two measures of participants' inquiry skills for performing the task were used:

1. Experimentation Skills: Usage of a control-of-variables strategy. As the key strategy for valid experimentation, usage of CVS represents a crucial measure of inquiry skills. When pupils use this strategy, they vary only one variable per experiment, which enables them to make a valid inference about the effect of this variable. The measure has also been used by some authors (Veenman et al., 2004) as a measure of metacognitive skilfulness, where varying more than one variable at a time represents poor systematic behaviour and a lack of experimental control.

The use of CVS was determined using the average number of independent variables changed per experiment in each learning session. This measure was calculated automatically through analysis of the computer logfiles of each learning session.

2. Validity of Inferences. This measure was based on an analysis of verbal protocols regarding participants' responses to the question *"What have you found out?"*, posed by the researcher after each experiment, along with the inspection of experimental data in computer logfiles that documented participants' choice of independent variables in each experiment. Statements regarding the relationship between each independent variable and forest devastation were coded as valid when a controlled comparison of two experiments was followed by a correct conclusion regarding the causal status of the only uncontrolled independent variable and the justification of the inference made reference to the evidence. If there was no evidence of a controlled comparison in the produced data set, or if participants interpreted the evidence incorrectly, the inferences were coded as invalid.

For each learning session, the percentage of valid inferences with respect to the total number of inferences made during the session was calculated.

Twenty percent of participants' responses were double coded in a test of inter-rater reliability. Cohen Kappas were between .90 and .98, and the percent agreement was between 96.9% and 99.2% (depending on independent variable).

Metacognitive Skills: Plan and Organisation of Inquiry Activity

For each learning session, participants' verbal protocols were reviewed to determine the level of plan and organisation of inquiry activity during task performance. The dominant level of plan and organisation achieved at each learning session was coded as follows (based on Schauble, 1996):

1. *General plan,* when participants' experimentation reflected the existence of a plan that took into account the whole structure of the problem. These global plans guided participants in generating evidence throughout the experimentation process and resulted in systematically organised data sets.

2. *Partially organised activity,* when there was evidence for a guiding plan that helped participants run several experiments in a row, but these plans were abandoned or forgotten before their completion.

3. *Local chaining,* when participants showed the tendency to compare pairs of experiments in isolation, without being aware of the wider structure of the problem.

4. *No plan,* when participants ran experiments in an order that did not seem logical and when they did not express any sign of planning activity, even at the level of individual experiment.

Acquired Knowledge About the Causal Structure of the Task Model

Acquired knowledge was assessed based on the "theory interview" held at the end of each learning session, which probed participants' theories on the relationships between variables. For each independent variable, the following questions were posed to participants: *Does feature X have an effect on the level of forest devastation?* and *How does feature X affect the outcome?*

Participants' answers were compared to correct statements about the effects of independent variables. Two points were awarded to a completely correct response, and zero points were awarded to false or non-existant answers. One point was awarded to a partially correct answer, which was possible in the case of three independent variables (where frequency of acid rain had a non-linear relation to the dependent variable, and the type of forest and presence of blight had a main and interacting effect on the dependant variable). Based on the scores received for each independent variable, a composite measure of acquired knowledge was constructed as the sum of points gained for each variable, with scores ranging from 0 to 10.

Results

Changes in Metacognitive Knowledge Related to Repeated Engagement with an Inquiry Learning Task

As previously mentioned, the present study examined change in metacognitive knowledge through both metatask and metastrategic components. Pupils' understanding of the task objective and of the strategies needed to accomplish the task was tested twice: before the first learning session and after the fourth session.

Table 2 presents levels of attainment for the metatask and metastrategic components at these two measurement points (before - after). Both components represented pupils' responses to open-ended direct questions about the task goal and strategies.

Due to the small number of participants and a violation of the assumption of a normal distribution of measures, differences between time points were analysed using non-parametric Wilcoxon sign tests.

Table 2. *Attained Levels of Metatask and Metastrategic Knowledge Before and After Engagement with the "Forest Devastation" Task: Number of Pupils at Each Level*

Metatask knowledge	**Pre** ***N***	**Post** ***N***
0. No insight into the goal of the task	4	1
1. Orientation towards attainment of positive outcome	18	1
2. Analysis at instance level	3	1
3. Analysis at feature level, but without its isolation	9	13
4. Analysis at feature level with reference to multiple features	0	14
5. Analysis at feature level with focus on a single feature at a time	0	4
Metastrategic knowledge		
0. No awareness of the need for a strategic approach	20	16
1. Choosing instances believed to yield a positive outcome	14	0
2. Choosing different instances to observe outcomes	0	6
4. Comparing instances in an uncontrolled fashion	0	0
5. Comparing instances in which a single feature is varied and other features are not mentioned	0	6
6. Comparing instances in which a single feature is varied and other features are held constant	0	6

Metatask Knowledge

Prior to engaging with the "Forest Devastation" task, pupils demonstrated low metatask knowledge. For the majority of participants (18, or 53%), the stated goal of the task was to achieve a positive outcome (*level 1*). However, the responses of a substantial portion of participants (9, or 26%) could be classified as *level 3* (the highest achieved level at this point of measurement), a level characterized by the intent to investigate the impact of various combination of features to Peter's tardiness. At this level, pupils recognized that the aim of the task was to discover the impact of features, but talked about them in a general way without the intent to single out the effects of individual features.

After the fourth learning session, only three participants achieved the lowest levels of metatask knowledge (*0* to *2*), while the majority of pupils (27, or 79%) achieved levels *3* and *4*. At these levels, participants recognized that the task goal was to analyse which features lead to which outcome, but conceptually were not focused on the need to analyze the effect of each individual feature. The highest level (*level 5*), characterized by a focus on analysis at the level of an individual feature, was achieved by only four participants (11%).

The observed advancement in metatask knowledge following engagement with the task was statistically significant (T=16.89, z=-4.39, p<.001, r=.53)[2].

Metastrategic Knowledge

Prior to engaging with the task, metastrategic knowledge was even lower than metatask knowledge, with no pupils demonstrating metastrategic knowledge above *level 1*. Most participants (20, or 59%) exhibited no insight into the need for some form of strategic approach. Typical responses at this *0 level* were: *"You can select an icon randomly"* or *"You just click on the icon"*. Responses at *level 1*, evident in 14, or 41% of pupils, were related to the search for a positive outcome.

However, pupils demonstrated significant progress in metastrategic knowledge following engagement with the "Forest Devastation" task (T=4.05, z=-3.54, p<.001, r=.43).

In the second application of the questionnaire, a significant proportion of students (12, or 35%) achieved the highest level of metastrategic knowledge, which is characterized by an understanding of the need for controlled comparison by varying one feature at a time (in order to determine the effect of this feature). Still, a large number of participants (16, or nearly 50%) remained at the lowest levels.

Metacognitive Knowledge on a Transfer Task

Table 3 presents pupils' responses on the metacognitive knowledge transfer task ("Peter task") before and after engaging with the "Forest Devastation" task. In

---

[2] Because non-parametric tests generally have less statistical power than their parametric counterparts, the indicator of effect size (Pearson correlation coefficients) was also used. These coefficients were calculated according to the formulas in Field (2005). For the interpretation of effect size, Cohen's (1992) recommendations were used.

this task, pupils were exposed to two different proposals - experimentation plans. They were then asked to identify which proposal is better and why (metastrategic knowledge), and to explain what one can find out by following these alternative plans (metatask knowledge).

Table 3. *Number of Pupils Who Selected Each Proposal for Experimentation in the Peter Task - Before and After Engagement with the "Forest Devastation" Task*

Transfer task	**Pre** *N*	**Post** *N*
Chose valid experimentation proposal	11	11
(Adequate explanation)	(2)	(3)
Chose invalid experimentation proposal	9	7
Claimed that both proposals are equally appropriate	2	10
Claimed that none of the proposals are appropriate	10	5
Don't know/No answer	2	1

Prior to engaging with the "Forest Devastation" task, a proposal that included a valid experimentation strategy was selected by 11 pupils (32%), while an alternative proposal was selected by slightly fewer pupils (9, or 26%). However, pupils' explanations of their responses revealed that only two pupils chose a valid proposal with an understanding of the value of controlled comparison, where they demonstrated highly developed metastrategic knowledge. Furthermore, these pupils demonstrated their metatask understanding by making appropriate conclusions about what could be achieved by following each proposal. They correctly reported that, by using uncontrolled comparison, no valid conclusion could be made, while controlled comparison might lead to a conclusion about the effect of the varied feature on Peter's arrival time.

Other pupils explained their choice based on the claim that their preferred proposal presented a more favorable combination of features or the combination that would enable Peter to arrive to school on time. The same argumentation was used by pupils who chose the opposite proposal (which used an invalid experimentation strategy) and by pupils who selected the answer 'both/none of the proposals are good'. These responses demonstrated that these pupils were not able to expand their focus beyond a consideration of experiment outcomes (although the actual outcomes were not even provided). This focus is further demonstrated by answers given to the questions regarding what pupils could discover if they followed each experimentation plan: *We would find out that Peter will be more/less late* and *He would have arrived faster if he had chosen some other features*.

After four learning sessions, only 11 pupils (32%) selected a valid experimentation proposal. Of these, only three participants made a valid argumentation of their choice based on the usage of CVS in their selected proposal.

The remaining pupils who considered the valid experimentation proposal to be better (8 of 11) were not aware of the fact that this was the only procedure that allowed for a valid causal inference and backed up their selection by the claim that this proposal presents a "better" combination of features (according to their expectations).

The selection of opposing proposals, made by seven pupils (21%), was explained using the same claims. A conceptually identical explanation was also given by pupils who claimed that both proposals were equally good and by pupils who claimed that no proposal was good. These pupils stated that both combinations of features led to similar results, or believed that some other combination of features would give a more favorable result.

Changes in pupils' responses on the metacognitive transfer task before and after engaging with the "Forest Devastation" task was measured using a marginal homogeneity test, which indicated no statistically significant values (MH_M=25.5, *Std. MH*=.367, p=.71).

Metacognitive Knowledge and Performance on the "Forest Devastation" Task

In order to examine the relationship between metacognitive knowledge and performance on the task, a derived measure of metacognitive knowledge was used instead of separate measures of metatask and metastrategic components, as proposed by Kuhn and Pearsall (1998)[3]. This derived measure used the highest value exhibited by a pupil in either metatask or metastrategic component. As such, the range for this new measure was between *level 0* (for pupils that achieved the lowest level in both components of metacognitive knowledge) and *6* (for pupils who achieved this level within the metastrategic component).

The relationships between metacognitive knowledge and performance were tested in two situations: metacognitive knowledge as exhibited in the first application of the questionnaire was related to task performance in the first learning session, while metacognitive knowledge in the second questionnaire application was related to task performance in the fourth learning session.

Metacognitive Knowledge and Task Performance in the First Learning Session

Prior to engaging with the "Forest Devastation" task, the achieved range of the new derived measure of metacognitive knowledge was between *0* and *3*. To test differences in task performance between groups of pupils who achieved different

[3] These authors argued that a joint measure of metacognitive knowledge represents a more appropriate measure than individual measures of each component, due to the fact that the assessment method is verbal and might underestimate the actual level of metacognitive understanding held by pupils.

levels of metacognitive knowledge, pupils were divided into two groups: lower (*levels 0* and *1*) and higher (*levels 2* and *3*) metacognitive knowledge.

Mann-Whitney tests were used to test the differences between these two groups in their use of CVS, the validity of their inferences and the accuracy of their theories, while a chi square test was used to test differences in the planning/organization of inquiry activity.

Table 4 presents the results of these tests as well as central values for measures of task performance during the first learning session for groups of pupils with lower and higher levels of metacognitive knowledge.

Table 4. *Median Values (C) of Measures of Strategic Performance in the First Learning Session, Grouped According to Level of Metacognitive Knowledge (MK) Attained Before Engagement with the Task*

Strategic performance: 1st session	Lower MK (*Levels 0-1*) *N*=22	Higher MK (*Levels 2-3*) *N*=12	*Sig.*
Average number of features changed per experiment	2.39	2.26	*U*=127.50 *p*=.87 *r*=.03
Percentage of valid inference	44.44	43.33	*U*=124.50 *p*=.79 *r*=.05
Accuracy of theories on the relation between IVs and DV	5.00	6.00	*U*=97.50 *p*=.21 *r*=.22

Table 5 presents the results examining group differences in the application of metacognitive skills of planning and organising the inquiry activity.

Table 5. *Number of Pupils Using Each Type of Plan in the First Learning Session, Grouped According to Level of Metacognitive Knowledge (MK) Attained Before Engagement With the Task*

Metacognitive skills	Lower MK (*Levels 0-1*) *N*=22	Higher MK (*Levels 2-3*) *N*=12	*Sig.*
General plan	2	0	χ^2=6.71
Partially organised activity	5	5	*df*=3
Local chaining	8	7	*p*=.08
No plan	7	0	φ=.44[4]

[4] Here the effect size measure is Cramer's V of Phi.

In the first learning session, no statistically significant differences were observed between the lower and higher metacognitive knowledge groups in strategic performance. The difference observed in planning and organisation of inquiry activity was also not statistically significant. However, a Phi value of .44 suggests a medium to higher effect size in the direction of a more highly planned and organised experimentation process for the group of pupils that achieved higher levels of metacognitive knowledge.

Metacognitive Knowledge and Task Performance in the Last Learning Session

Following engagement with the task, pupils were allocated to levels from *2* to *6* on the joint measure of metacognitive knowledge. Pupils were divided into three groups with respect to the attained level of metacognitive knowledge: low (*2* or *3*), medium (*4*) and high (*5* or *6*). The values on measures of task performance for these groups during the final learning session are shown in Table 6 and Table 7.

Here, only the differences between the extreme groups (low and high) in the use of CVS, validity of inferences, accuracy of pupils' theories, and the level of plan and organisation of inquiry activity were tested. This decision was justified by the small number of participants and the corresponding low statistical power of tests, and by an intent to alleviate the effects of multiple comparisons between groups.

Table 6. *Median Values (C) of Measures of Strategic Performance in the Last Learning Session, Grouped According to Levels of Metacognitive Knowledge (MK) Attained After Engagement With the Task*

Strategic performance: 4th session	Low MK (*Levels 2-3*) N=13	Medium MK (*Level 4*) N=7	High MK (*Levels 5-6*) N=14	*Sig.* (extreme group comparison)
Average number of features changed per experiment	2.11	1.85	1.81	U=40.00 p=.01 r=.48
Percentage of valid inference	54.17	66.67	80.56	U=48.00 p=.04 r=.40
Accuracy of theories on relation between IVs and DV	8.00	10.00	9.07	U=54.00 p=.07 r=.35

The results of Mann-Whitney tests revealed statistically significant differences in the use of CVS and the validity of inferences, while the differences in the accuracy of pupils' theories about the relationships between variables were not statistically significant, but suggested a medium size effect. The group of pupils who attained high levels of metacognitive knowledge after engagement with the task experimented more systematically, made more valid inferences and acquired more knowledge about the causal structure of the task than the group that remained at a low level of metacognitive knowledge.

Table 7. *Number of Pupils Using Each Type of Plan in the Last Learning Session, Grouped According to the Level of Metacognitive Knowledge (MK) Attained After Engagement With the Task*

Metacognitive skills	Low MK (*Levels 2-3*) *N*=13	Medium MK (*Level 4*) *N*=7	High MK (*Levels 5-6*) *N*=14	*Sig.* (extreme groups comparison)
General plan	4	4	6	χ^2=1.12
Partially organised activity	5	2	6	*df*=2
Local chaining	4	0	2	*p*=.57
No plan	0	1	0	φ=.20

A Chi square test for the difference between the low and the high metacognitive knowledge groups in the planning and organisation of inquiry activity revealed no statistically significant difference, but a small effect size.

Discussion

The first question raised in this study focused on changes in metacognitive knowledge following repeated engagement with an inquiry learning task.

The data presented here suggest that pupils demonstrated improvement in metatask and metastrategic knowledge after four learning sessions within the FILE system. This finding is consistent with other microgenetic studies of inquiry learning (Kuhn et al., 1992, 1995; Schauble, 1990, 1996), which also demonstrated that the participation of pupils in repeated learning sessions that required self-directed experimentation is sufficient (even without specific feedback) for inducing changes at the strategic and meta-levels.

Prior to engaging in the "Forest Devastation" task, pupils' understanding of the nature and requirements of the task, as well as their understanding of the strategies that are applicable to the task, proved to be quite insufficient. This suggests that an understanding of the task requirements could be expected and facilitated only when pupils started working on the task, where they gained direct experience in experimentation and making causal inferences.

With direct and repeated experience working on the inquiry learning task, pupils' understanding of the task changed from one oriented towards achieving a positive outcome to one aimed at exploring the effects of individual features. Pupils also demonstrated improvement in their understanding of the strategies applicable to the task. A significant number of pupils acquired an understanding of the appropriateness of the control-of-variables strategy for responding to the task goals. However, nearly half of the pupils continued to demonstrate insufficient metastrategic understanding at post-test and remained at the lowest level, indicating an inability to formulate any strategic approach or to recognize the need for such an approach. Overall, greater progress in metatask knowledge (when compared to metastrategic knowledge) might be the result of the possibility that questions posed by the researcher during pupils' work and interviews provided more cues for gaining an understanding of the task goal than for gaining an understanding of the strategies applicable to the task. Alternatively, this difference might also be the result of the possibility that metastrategic knowledge is existent, but not completely explicit. For pupils, it might be easier to reach insights into and explicate what is necessary to accomplish in an inquiry learning task than to find out how to achieve that goal.

The data further suggest that pupils mostly failed to apply newly achieved metacognitive knowledge to the transfer "Peter task". In this task, where pupils were asked to apply metacognitive knowledge to a concrete example, no advancement in metatask and metastrategic knowledge was observed. Instead, where pupils were required to demonstrate their understanding of the task goals and to explain why a control-of-variables strategy is the optimal strategy for achieving the goal of identifying the effects of individual features, pupils demonstrated a tendency to regress to a level that represents the engineering approach to experimentation (Schauble, Klopfer, & Raghavan, 1991), i.e., an approach that aims at determining the optimal combination of features that lead to the desirable outcome. Obviously, pupils did not recognize the applicability of the previously acquired valid inquiry strategies, despite the fact that the transfer task used an identical task model (with different content) and required the same type of activities that had already been used on the main task. According to the measures applied, only three pupils had sufficiently developed metatask and metastrategic knowledge that enabled the application of this knowledge to new learning situations. These findings are in contrast with the results related to improvement in metatask and metastrategic knowledge observed through direct questioning. The fact that metacognitive knowledge gained through engagement with the "Forest Devastation" task could not be applied in the context of a parallel transfer task suggests that metacognitive gains were fragile and restricted only to the immediate learning situation. These findings also imply that, while mere repeated engagement in experimentation activity can improve meta-level competency, "stronger" interventions are probably needed for transfer to a variety of contexts.

Arguably, transfer was limited by the fact that the content of the task was from the daily life domain, which is the domain (together with the social domain) that is particularly problematic for the establishment of valid experimentation and reasoning approaches, mainly because of difficulties in coordinating the experimental evidence with the extremely rich and elaborate theories of pupils (Kuhn & Pearsall, 1998; Kuhn et al., 1995).

Results of other studies also point to the limited transfer of metacognitive understanding, even when some interventions for building metacognitive knowledge were applied (Keselman, 2003; Keselman & Kuhn, 2002; Kuhn et al., 2000; Zohar & Ben David, 2008). These studies suggest that, in order to ensure higher and universal metacognitive gains and to secure successful transfer to new learning situations, more sustained practice and the direct instruction of (meta-level) strategies in a variety of contexts is needed.

Arguably, pupils did not demonstrate improved metacognitive understanding on the transfer task in part because of the way in which this task was structured. In this task, two fictional pupils disagreed over how to carry out a second experiment. In the "Forest Devastation" task, pupils often began their experimentation by setting the combination that would result in a minimal outcome (even when they understood that CVS was necessary) because they would like this "minimum combination" to serve as a reference point for comparison with other experiments. On the transfer task, it seems plausible that pupils did not recognize the need to create a second experiment in a manner that would allow for a valid comparison with the first experiment, but instead believed that it was first necessary to establish the optimal combination of features. Bearing this in mind, it is possible to assume that, although pupils might have had general understanding of the need to examine the effect of individual features, the presentation of only the first two experiments did not encourage pupils to think in a valid way. If this assumption is correct, it is likely that pupils, regardless of their initial unfavorable result on the transfer task, would likely advance faster (in relation to the "Forest Devastation" task) if they were given the opportunity to experiment with a new content task. While this study did not allow for an examination of this assumption, other studies have suggested such a possibility (Kuhn et al., 1992, 1995).

The second question raised in the present study focused on the relationship between metacognitive knowledge and task performance. It was assumed that successful strategic performance would be impossible without an understanding of the task goal as one of analyzing the effects of individual features and without an awareness of the need for varying features in the experiments in a way that would allow for valid inference making (Kuhn, 1999, 2002a, 2002b).

The data presented in this study suggest that groups of pupils who achieved lower and higher levels of metacognitive knowledge prior to actual engagement with the "Forest Devastation" task did not differ with respect to task performance during the first learning session. In this session, both groups changed a similar number of features per experiment, made an equal percentage of valid conclusions and acquired the same level of knowledge about the causal model. This finding, indicative of an absence of a relationship between metacognitive knowledge and task performance might be attributed to three issues. Firstly, initial low levels of metacognitive knowledge among all participants (the highest level achieved by any pupil was *3*) might represent equally inappropriate mental frameworks for initiating inquiry activities, irrespective of the metacognitive level achieved. Secondly, achieved metacognitive knowledge prior to actual engagement with the task might be viewed as a reflection of pupils' declarative knowledge that was based on a superficial familiarisation with the task content and procedures, but not on their own experiences and reflections of that experience. It might be argued that true metacognitive understanding can only be facilitated by actual experimentation and the application of cognitive and metacognitive skills and that metacognitive knowledge measured prior to such activities was not an accurate reflection of actual understanding. Finally, due to verbal mediation and a question format that gives pupils freedom in deciding how to respond, it is possible that the paper-and-pencil measure of metacognitive knowledge underestimated pupils' true level of understanding.

However, the data does indicate that pupils who achieved higher levels of metacognitive knowledge prior to actual engagement with the task were somewhat more organised and systematic when it came to actual execution of the inquiry activity. This observed relationship between metacognitive knowledge and applied metacognitive skills suggests that higher metacognitive knowledge is likely to lead to more appropriate (meta)strategic responses by providing input for making decisions about which strategic approach to apply to the task.

A slightly different relationship between metacognitive knowledge and strategic performance was observed for the last learning session. Here, the group of pupils with high metacognitive knowledge (measured after engagement with the task) performed better on the task than the group of pupils who did not improve on the metacognitive knowledge measure. However, these groups were quite similar in terms of the experimentation plan and organisation of the learning activity in the last session.

The data further suggest that even achievement of the highest levels of metacognitive knowledge (*levels 5* and *6*) did not guarantee strategic success, as some of the pupils who reached a high metacognitive level did not use valid strategies consistently. This finding confirms Kuhn and Pearsall's observation (1998) that a particular level of metacognitive knowledge is a necessary, but not sufficient condition for successful performance on the task. Pupils might know that the task goal is to determine the effects of individual features and might understand the need for systematic and strategic action, but still might conduct experiments in an unorganised and unsystematic way. It might be argued that the reasons for not applying metacognitive knowledge might be related to motivation (e.g., the task might be boring or irrelevant to pupils, while it requires an investment of effort), strategic deficits (pupils might lack the inquiry skills necessary for adequate execution of activities), or metacognitive flaws (pupils' skills in planning, monitoring, evaluation and control of the learning process might be underdeveloped). Future research efforts should further examine the interacting relationships between these elements and attempt to determine causality in these relationships.

References

Beishuizen, J., Wilhelm, P., & Schimmel, M. (2004). Computer-supported inquiry learning: Effects of training and practice. *Computers & Education, 42*, 389-402.

Ben-David, A., & Zohar, A. (2009). Contribution of meta-strategic knowledge to scientific inquiry learning. *International Journal of Science Education, 31*(12), 1657-1682.

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.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.

De Jong, T., & Van Joolingen, W.R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*, 179-201.

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

Efklides, A. (2009). The role of metacognitive experiences in the learning process. *Psicothema, 21*(1), 76-82.

Field, A. (2005). *Discovering statistics using SPSS* (Second edition). London, UK: Sage Publications Ltd.

Flavell, J.H. (1979). Metacognition and cognitive monitoring. *American Psychologist, 34*, 906-911.

Flavell, J.H., Miller, P.H., & Miller, S.A. (2002). *Cognitive development* (4th ed.). Upper Saddle River: Prentice Hall.

Hulshof, C.D. (2001). *Discovery of ideas and ideas about discovery: The influence of prior knowledge on scientific discovery learning in computer-based simulations.* (Unpublished doctoral dissertation). University of Twente, Twente.

Hulshof, C.D., Wilhelm, P., Beishuizen, J.J., & van Rijn, H. (2005). FILE: A tool for the study of inquiry learning. *Computers in Human Behavior, 21*, 945-956.

Keselman, A. (2003). Supporting inquiry learning by promoting normative understanding of multivariable causality. *Journal of Research in Science Teaching, 40*(9), 89-921.

Keselman, A., & Kuhn, D. (2002). *Facilitating self-directed experimentation in the computer environment*. Retrieved from http://citeseer.ist.psu.edu/509879.html.

Kuhn, D. (1995). Microgenetic study of change: What has it told us? *Psychological Science, 6*(3), 133-139.

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

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science, 9*(5), 178-181.

Kuhn, D. (2001). Why development does (and does not) occur: Evidence from the domain of inductive reasoning. In J.L. McClelland & R.S. Siegler (Eds.), *Mechanisms of cognitive development: Behavioral and neural perspectives* (pp. 221-249). Mahwah, NJ: Lawrence Erlbaum Associates.

Kuhn, D. (2002a). A multi-component system that constructs knowledge: Insights from microgenetic study. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 109-130). Cambridge: Cambridge University Press.

Kuhn, D. (2002b). What is scientific thinking and how does it develop? In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 371-393). Malden, MA: Blackwell Publishing Ltd.

Kuhn, D. (2005). *Education for thinking.* Cambridge, MA: Harvard University Press.

Kuhn, D., Black J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction, 18(4)*, 495-523.

Kuhn, D., Garcia-Milà, M., Zohar, A., & Andersen, C. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development, 60*, 1-128.

Kuhn, D., & Katz, J. (2009). Are self-explanations always beneficial? *Journal of Experimental Child Psychology, 103*(3), 386-394.

Kuhn, D., & Pearsall, S. (1998). Relations between metastrategic knowledge and strategic performance. *Cognitive Development, 13*, 227-247.

Kuhn, D., Pease, M., & Wirkala, C. (2009). Coordinating the effects of multiple variables: A skill fundamental to scientific thinking. *Journal of Experimental Child Psychology, 103*, 268-284.

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Ministarstvo znanosti, obrazovanja i športa (MZOS) (2006). *Nastavni plan i program za osnovnu školu.* Zagreb: Ministarstvo znanosti, obrazovanja i športa.

Patton, M.Q. (1990). *Qualitative evaluation and research methods.* Beverly Hills: Sage.

Pogačnik, V. (1997). *Priručnik za Test nizova.* Jastrebarsko: Naklada Slap.

Ristić Dedić, Z. (2010). *Ispitivanje motivacijskih i metakognitivnih čimbenika procesa istraživačkog učenja u računalno podržanom okruženju* [An examination of motivational and metacognitive processes in inquiry learning within a computer-supported learning environment]. (Unpublished doctoral dissertation). Filozofski fakultet Sveučilišta u Zagrebu, Zagreb.

Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology, 49,* 31-57.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology, 32*(1), 102-119.

Schauble, L., Klopfer, L.E., & Raghavan, K. (1991). Students' transition from an engineering model to a science model of experimentation. *Journal of Research in Science Teaching, 28*, 859-882.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, *26,* 113-125.

Schraw, G., & Dennison, R.S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19,* 460-475.

Veenman, M.V.J. (2011). Learning to self-monitor and self-regulate. In R.E. Mayer & P.A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 197-218). New York: Routledge.

Veenman, M.V.J. (2012). Metacognition in science education: Definitions, constituents, and their intricate relation to cognition. In A. Zohar & Y.J. Dori (Eds.), *Metacognition in science education: Trends in current research* (Vol. 40, pp. 21-36). Dordrecht: Springer.

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., & Spaans, M.A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences, 15*, 159-176.

Veenman, M.V.J., van Hout-Wolters, B.H.A.M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning, 1*, 3-14.

Veenman, M.V.J., Wilhelm, P., & Beishuizen, J.J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction, 14,* 89-109.

Wilhelm, P., & Beishuizen, J.J. (2003). Content effects in self-directed inductive learning. *Learning and Instruction*, *13*, 381-402.

Wilhelm, P., Beishuizen, J.J., & van Rijn, H. (2005). Studying inquiry learning with FILE. *Computers in Human Behavior*, *21*, 933-943.

Zohar, A. (2012). Explicit teaching of metastrategic knowledge: Definitions, students' learning, and teachers' professional development. In A. Zohar & Y.J. Dori (Eds.), *Metacognition in science education: Trends in current research* (Vol. 40, pp. 197-223). Dordrecht: Springer.

Zohar, A., & Ben David, A. (2008). Explicit teaching of meta-strategic knowledge in authentic classroom situations. *Metacognition and Learning*, *3*, 59-82.

Zohar, A., & Peled, B. (2008). The effects of explicit teaching of metastrategic knowledge on low- and high-achieving students. *Learning and Instruction*, *18*, 337-353.

Received: November 20, 2013

PRAZNA PARNA STRANICA