EFFECTS OF OBSERVATIONAL LEARNING ON STUDENTS' USE OF AND ATTITUDE TOWARDS READING AND LEARNING STRATEGIES

JISKA KNIEP* & TANJA JANSSEN**

** University of Applied Sciences, Amsterdam, the Netherlands** University of Amsterdam, the Netherlands*

Abstract

Previous research has shown that observation can be effective for learning in various domains, such as writing, reading and creative art work. By observing models at work, students can develop strategic knowledge and they might also change their conception of what the modeled skill involves. The question in this study is which instructional approach is more effective for students' processes of studying text: learning by observation or learning through practice? We designed an intervention that enables students to learn from observation of video models. The models in these videos are peers, who read and learn a history text while thinking aloud. In a pre-test/post-test control group design we assigned 52 Dutch students (15-18 years old) to one of two conditions; one observational learning condition and a control condition (learning through practice). In the observational learning condition, students were asked to observe and evaluate the thinking processes of two peers on video, and decide which was the weaker and which was the stronger better performer of a study task. In the control condition, students received direct instruction and practice in reading and learning strategies. We measured students' attitude and self-reported use of learning strategies with a questionnaire (at pre- and post-test) and a learner report (post-test only). Students' use of strategies at pre- and post-test was measured with a think aloud task with eight participants of each condition. This showed that students in the experimental condition checked their own understanding of the text while studying for a history test more often than students in the control condition. Moreover, students in the experimental condition reported more learning experiences in their learner reports than those in the control group. In these reports, this group also reported almost 10% more metacognitive learner experiences.

Keywords: observational learning, (meta)cognitive learning strategies, studying reading

Kniep, J. & Janssen, T. (2014). Effects of Observational Learning on Students' Use of and Attitude towards Reading and Learning Strategies. L1-Educational Studies in Language and Literature, 14, p. 1-21. http://dx.doi.org/10.17239/L1ESLL-2014.01.02

Corresponding author: Jiska Kniep, Hogeschool van Amsterdam, Wibautstraat 2-4, 1091 GM Amsterdam, email: j.kniep@hva.nl

1. INTRODUCTION

Secondary school students have to read many study texts in order to succeed in exams and tests in various school subjects. When it comes to students understanding what they read, Schellings, Aarnoutse and Van Leeuwe (2006) distinguish two levels: the identification of words in a text and the construction and integration of meaningful content within their prior knowledge and experiences. The identification of words is essential in the reading process, but it is not sufficient to understand texts. In order to do this, readers need to construct meaning when they read texts and they need to integrate new information into existing knowledge.

Study reading is even more complex than reading for understanding. Cognitive, metacognitive and affective processes influence this task. For example, the student's motivation and prior knowledge of the content, the student's relationship with the teacher, learning aims, and the difficulty of the content itself (Janssen, Ten Dam & Van Hout-Wolters, 2002).

Reading for study includes more than just comprehending what a text is about. There is a strong focus on memorizing; activities that take place before, during and after reading are all directed to achieving that goal (Brand-Gruwel, Schellings & Teulings, 1998). Orientation and setting goals are activities that take place beforehand. During the process, learners determine the main idea of the text and connect the information with prior knowledge. Pintrich and De Groot (1990) refer to this as elaboration. Controlling, rereading and producing test questions are features of the last phase.

There are many strategic activities which are essential for successful reading; to name but a few: activation and use of prior knowledge, determining the reading purpose, establishing relations between words, sentences and paragraphs including the prediction of information and the creation of images; determining the nature and type of a text; identifying the theme and main idea along with a summary of the text; posing and answering one's own questions; planning, steering, monitoring and correcting of one's own reading behavior; evaluating the purpose of texts; and reflecting on the reading activities which have been executed, and their results (Schellings et. al., 2006, p. 550-551).

In the reading for study process, two major categories are usually distinguished: cognitive and metacognitive activities. Pintrich, Smith, Garcia and McKeachie (1991) subdivide the first category into four different types of strategies: rehearsal, elaboration, organization and critical thinking. Metacognitive activities include planning, monitoring and modifying cognitive activities. Several researchers have emphasized that use and knowledge of the latter is a better predictor of good text comprehension than intelligence (Hettinger-Steiner & Carr, 2003; Schraw, 1998; Wang, Haertel & Walberg, 1993).

Wolters and Pintrich (1998) make a distinction between 'depth learning' and 'surface learning'. Depth learners try to give meaning to the text they have to learn. In order to do that, they are critical towards the text and try to elaborate and relate

the information they have to learn to other sources, or to prior knowledge. Elaboration and critical thinking are strategies depth learners use. Ideally, they also use metacognitive activities. Surface learners, on the other hand, spend most of the time repeating, rereading and memorizing text content. Rehearsal and organization are strategies characteristic of surface learners. This surface approach is related to low achievements (Vermunt, 1992).

Schellings et al. (2006) investigated text comprehension activities undertaken by beginning learners in primary school. Based on prior research, they assumed that third-graders were expected to use some of the text-based and prior knowledge-based strategies mentioned above. They also expected pupils of this age to use some metacognitive strategies. The data of the think aloud task used in their research, gave new insight into the used reading comprehension strategies and their interplay with reader characteristics and situational variables. Their study shows that the use of reading strategies is related to reading comprehension as measured by standardized tests. Furthermore, they conclude that the think aloud method is a reliable instrument to determine strategic activities among young readers.

Bimmel and Van Schooten (2004) demonstrated that the degree to which 15-year-old students in general secondary education in The Netherlands master strategic reading activities is related to their reading comprehension.

Having reviewed studies which have investigated good readers and what they were doing while reading, Pressley and Afflerbach (1995) conclude that good readers are very strategic before, during and after reading; good readers interact with the text, and are engaged with the text and the reading process itself.

In a previous study we examined the reading activities of eight 15 to 16-year-old students while reading study texts for school subjects whilst thinking out loud. Half of the students were high achievers, half were low achievers. We found that the high achieving students used more, and more diverse, strategies compared to the low achievers (e.g. making associations, relating information in the text to their own prior knowledge). In addition, only high achievers displayed metacognitive and self-regulating activities. Low achievers tended to use rereading and writing (summarizing or copying text) as their main study strategy (Kniep & Janssen, 2011).

1.1 Observational Learning

We have seen that reading for study is a complex process that arises from and is influenced by, many different factors, such as motivation, cognitive and metacognitive activities, prior knowledge and activities that take place before, during and after the learning process. The question arises how low achieving students can be stimulated to broaden their repertoire of strategies while reading study texts and become ‘depth learners’. One promising instructional approach is ‘learning by observation’, that is: learning by watching peers perform a task. The observational learning method involves much more than imitating models. By observing others at

work, students can develop strategic and process knowledge and might alter their conception of what the modeled task involves. Low achieving students, for instance, may become aware that effective studying involves much more than just copying and memorizing text parts.

According to Bandura's (1986) social learning theory, learning can take place by modeling and self-control processes. In observational learning, four higher order mental activities are involved, namely attention, retention, reproduction, and motivation. Learners have to actively pay attention to crucial details in the model's actions, store the information in memory for retrieval later (retention), have the motivational desire to reproduce the observed behavior, and reproduce this behavior in a successful manner (Schunk & Zimmerman, 1997).

Previous studies on modeling discuss the effects of different model types (age, level, etc.) on different subjects (age, gender, etc.). Zimmerman and Kitsantas (2002), for example, found that college students who observed a coping model who gradually improved her writing technique, surpassed students who had observed a mastery model. Braaksma, Rijlaarsdam and Van den Bergh (2002) found that, among 8th grade students, weak learners learn more from focusing their observations on weak models, whereas better learners learn more from focusing on good models.

It is assumed that similarity (age, level) in observer and model is most effective for learning by observation (Braaksma, Rijlaarsdam & Van den Bergh, 2002). Moreover, observation of models is not so much directed towards imitation of a final product as toward developing a clear image of how a skill should or could be performed. Another important characteristic of this method is that it can strongly stimulate evaluative activities by learners. The observation of others performing a task involves a "natural" step back and thus a more natural type of monitoring, evaluation and reflection on task execution processes than when students learn by practicing and by reflecting on their own performance.

In previous studies, observation was found to be an effective learning tool in various domains, e.g., for writing (Braaksma, Rijlaarsdam, Van den Bergh & Van Hout-Wolters, 2004), reading (Couzijn, 1999), and visual art making (Groenendijk, Janssen, Rijlaarsdam & Van den Bergh, 2013a; 2013b). For example, students who observed strong and weak peers performing a writing task, afterwards wrote better texts than students who wrote texts themselves (Couzijn, 1995; Braaksma, Rijlaarsdam & Van den Bergh, 2002).

The aim of the present study is to examine whether observational learning can also be effective in the domain of text study. We want to find out which instructional approach is more effective for text studying: learning by observation or learning by practicing? Effects are examined regarding students' attitude and use of cognitive and metacognitive strategies, while reading for study purposes.

2. METHOD

In this study a quasi-experimental design was used, involving a pre-test, post-test and control group.

2.1 Participants

The participants of this study were fifty-two 10th grade students in higher general secondary education, from two schools in The Netherlands (male: 48%, female: 52%). Their mean age was 16.1 years (SD = .86). At both schools, two 10th grade classes were randomly divided into two groups, so that each school had one experimental and one control group. The experimental condition (N=28) received the observational learning intervention, the control condition (N=24) received the 'practice' intervention. We found no significant differences between conditions in gender, nor in mean age. We also compared the mean scores on History tests and on CITO[i] between conditions. ANOVA's did not show significant differences.

2.2 Intervention

Both interventions consisted of three lessons of 50 minutes, one lesson a week, pre- and post-test not included. The first author was the teacher for both groups.

In the experimental observational learning condition during the first lesson, the teacher provided a short explanation about metacognition. Then they had to think about their own way of studying and they were asked to write down what they normally do while learning history schoolbook texts. After this, they were shown two videos of about three minutes each, in which peers - three pairs of strong and weak peer models - perform a text study task while thinking aloud. In the video, the peer models were seen to study texts that prepared them for real tests in their History and Economics courses. Students were asked to write down which of the models performed best and why. During the next two lesson participants in the experimental condition watched four more videos and evaluated them in the same manner.

The videos were recorded during a previous investigation of the use of different strategies by weak and strong learners within the same target group (Kniep & Janssen, 2011). During this previous research, students had been observed three times while studying from textbooks in History, Economics and Social Studies. From these videos, we extracted the excerpts used in this research. The selected excerpts were sequenced chronologically in a video, showing the text study process from beginning to end. Selection of the excerpts was based on the extent to which they demonstrated the strategies, and to which these were characteristic for the students involved. Strategies that were used by these models included repeating, writing summaries, making graphic organizers, asking questions, predicting, elaborating with prior knowledge.

Braaksma et al. (2002) argue that students are most effectively helped by observing models that are similar to themselves with reference to ability. This suggestion was used in the design of our intervention. We used two video fragments in each lesson: one of a weak learner, and one of strong one. This allowed all students to identify with the model most closely resembling their own competence level in reading one of the models, the incompetent or the more competent one.

Each excerpt lasted between 3 and 4 minutes and demonstrated the strategies used by the models during reading. These can be broadly categorized in five ways: rehearsal, elaboration, organization, critical thinking and metacognitive self-regulation. Not all the students used the same strategies. Weak students tended to mostly use the same ones: rehearsal and organization. In Figure 1, two descriptions of the videos used are given, one showing a weak performance and one showing a good performance.

Description of a weak performance
Erik is studying a History text about the economic crisis in de USA, 1929.
Erik starts to read the text out loud. First he reads the whole text, he doesn't stop. [1] Then he starts to read the text again. [2] He asks himself a simple question ('Why was everyone really optimistic?') and immediately gives the answer (the next sentence), while staring at the text. [3] In the following fragment, Erik quits reading. The researcher asks him what he is doing. He says he is thinking about a difficult part in the text. He asks: 'What caused the crisis to worsen?'. To answer his own question, he immediately reads the next sentences out loud to answer this question. [4] Then he starts reading again and stumbles across a difficult word. He rereads the previous and the next part to find out what this word means. He repeats a sentence and says that he now understands what the difficult word means. [5] Erik starts rereading again and stops when he doesn't understand a sentence. He rereads again. Then he says: 'Okay'. [6] He starts to underline dates in the text. [7] When he is finished with reading the paragraph he reads the underlined words/data again. He rereads them a second time. [8] He says that he will read repeatedly until he understands the text. [9]

Description of a strong performance:
Kim is studying a History text about the economic crisis in de USA, 1929.
[1] Kim starts to read and at the same time she underlines the important information in the text. [2] Then she tells how she decides what she thinks is most important; she reads the title and connects the information in the text with this title. [3] She then tries to make test questions about the text: "When I read this paragraph I think of my teacher and what he would ask about this part of the text. I think he would ask: How was it possible that the world crises suddenly had its huge impact at that moment and not at another moment? And then I try to give an answer to my own question." [4] "Now I have to learn a list. I count the concepts, there are four. So now I know there are four concepts and I am learning them by heart." [5] "Ahh, I already thought this part would come, about the American president. Because he thinks that people should continue spending money, so the New Deal. I think the next part in the text will tell me more about Roosevelt and the New Deal. I didn't read the text yet, but I just thought this part was coming." [6] "And now I am thinking, when and why did the economic crisis start in the 1930s?" Kim tries to answer to this question.

Figure 1. Two examples of observational learning videos

Students in the observational learning condition were asked to observe, evaluate and reflect on the study approaches of the models in the videos. First, students were asked to write down in silence their evaluation and observations. Their responses were then discussed in class. The teacher's role here was to facilitate the discussion. Both the evaluation ('which observed pupil performed the study task best, and why?') and the reflection ('which one do I most resemble, and what can I change/implement during the learning process?') were of specific interest during this part of the intervention.

For the control group, a 'learning-by-practice' intervention with direct instruction, reflection and evaluation was designed. All of the strategies the other group observed were practiced by this group. Students in the control condition practiced several cognitive strategies on a text used for the study of History. In the first lesson, they practiced rehearsal and organization strategies by playing an envelope game in which they paired concepts they had encountered in the text, see Figure 2.

Task 2 – Envelopes game

We will now practice the strategies 'repeating' and 'organization'. Many students read a text very often and try to get the text 'into their heads'. They learn important concepts 'by heart'. In the next task, you will have to learn the most important concepts from a History paragraph. You don't have time enough to read the text in-depth, but you are allowed to do the task with a classmate.

a. Create pairs.
b. Divide the text you have received (Section 9.2.: "The world's economic crisis") into two parts. You read only one part; your classmate reads the other part. You have 5 minutes to read and learn the text.
c. After 5 minutes the teacher will take the texts back.
d. Open the envelope that your teacher gave you.
e. Make 12 correct pairs from the 24 cards with concepts in the envelope.
f. Ask your teacher for the answer key when you have paired up all the cards. How many pairs were correct?

Figure 2. Example of a task for the strategies rehearsal and organization in the control group

After this, students had to organize the concepts in a self-chosen graphic organizer.

In the second lesson, students had to activate their prior knowledge about the actual economic crisis and relate this knowledge to the content of the History text (about the economic crisis of the 1930s). After that, they had to compare both crises.

In the third lesson students got an instruction and practice in critical thinking and metacognitive strategies. Students were asked to collect arguments for and against a thesis about the same History subject of the first two lessons. Then they had to argue for or against some phrases and comments in the text. They also had to predict the content of an unknown text, based upon a picture, the heading and the first paragraph. The last step was to think of good test questions for the text about the world crisis in the 30s. Each of these three lessons ended with a short discussion about student's own experiences with the strategies they used.

The structure of the interventions was largely kept the same. The only difference is that students in the control condition worked in pairs or groups, and students in the experimental conditions worked individually. Time allocated to practising the various strategies was about the same as the length of the video fragments in the other group. The same holds for the time given for reflection and evaluation on the task afterwards. During the evaluation/reflection phase, participants discussed which strategies had been practiced, and, in turn which students had learned from the practice. Here again, the teacher acted as facilitator.

2.3 Collection of data

A questionnaire was used to measure students' self-reported use of strategies and motivation before and after the intervention. Eight students from each condition were asked to study texts while thinking aloud, before and after the intervention. Attitude towards the intervention and/or the use of strategies was measured by a written learner report after the intervention. The time between pre- and post-test was about four weeks. In figure 3 we present an overview of the measures used in this study.

Experimental condition	Control condition
Pre-test:	
1. Questionnaire on learning strategies and motivation (MSLQ)	
2. Think aloud task (8 participants per condition)	
Intervention:	Intervention:
Observational learning, 3 lessons	Direct instruction and practicing, 3 lessons
Post-test:	
1. Questionnaire on learning strategies and motivation (MSLQ)	
2. Think aloud task (8 participants per condition)	
3. Learner report	

Figure 3. Design of this study.

2.4 Measures

Questionnaire. We used a Dutch translation of Pintrich' Motivated Strategies for Learning Questionnaire (MSLQ, Pintrich & De Groot, 1990; Blom, Hoek & Ten Dam, 2007). The original questionnaire consists of items to measure self-regulation (cognitive and metacognitive strategies), motivation and resource management strategies (Pintrich & De Groot, 1990). Since the goal of the present study is to measure students' strategy use, we omitted items about resource management strategies. We retained items about motivation, because students with a motivational orientation tend to engage in more self-regulated learning, such as metacognitive activities, cognitive strategies, and persistence in task completion (Pintrich & De Groot, 1990).

The terminology used in the items was translated for and adapted to the target group, i.e. Dutch students. To ensure the validity of this research, respondents were asked to keep the school subject History in mind while filling in the questionnaire. This was done because all instruments measured the use of strategies within the same subject (History), and because the intervention revolved around historical texts.

Students' motivation was measured by means of a questionnaire comprising of 23 questions pertaining to the sub-scales intrinsic goal orientation, extrinsic goal orientation, task value and self-efficacy. Respondents indicate the extent to which various statements apply to them on a five point Likert scale. Examples of items are: 'Getting a good grade in this class is the most satisfying thing for me right now' (extrinsic goal-orientation), 'In a class like this, I prefer course materials that really challenge me, so that I can learn new things' (intrinsic goal-orientation), 'I am very interested in the content area of this course' (task value), 'I'm confident I can understand the basic concepts taught in this course' (self-efficacy).

Twenty-seven questions were used to measure the use of strategies on the subscales rehearsal, elaboration, organization, critical thinking and metacognitive self-regulation. Examples of questions for each sub-scale are 'I make lists of important terms for this course and memorize the lists' (rehearsal), 'I try to relate ideas in this subject to those in other courses whenever possible' (elaboration), 'I make simple charts, diagrams, or tables, to help myself organize course material' (organization), 'I treat the course material as a starting point and try to develop my own ideas about it' (critical thinking). Metacognitive self-regulation asked students about planning, monitoring, evaluation and regulating activities in relation to learning. An example question is 'I try to think through a topic and decide what I am supposed to learn from it rather than just reading it when studying'.

We tested the questionnaire for reliability in a pilot research conducted among 26 10th-grade students, each with History as part of their curriculum. After some adjustments, the reliability of the different scales was assessed as acceptable to good (Cronbach's alpha between .63 and .89 at pre-test).

In Table 1, we present the reliability of the questionnaire scales at pre-test and at post-test.

Table 1. Questionnaire: reliability per scale

Scale	Items	Items deleted	pre-test (n=52) (Cronbach's alpha)	post-test (n=45) (Cronbach's alpha)
Rehearsal	5	0	.75	.65
Elaborate	6	1	.77	.72
Organization	5	1	.72	.64
Critical thinking	4	0	.65	.62
Metacognition	5	0	.63	.54
Intrinsic goal orientation	3	2	.80	.73
Extrinsic goal orientation	4	2	.65	.63
Task value	4	0	.74	.56
Self-efficacy	8	0	.89	.86

All scales were less reliable at post-test than at pre-test. In particular the metacognition and the task value scales have low alpha's.

At pre-test some scales were significantly correlated, namely organization and rehearsal (Pearson correlation r = .63), organization and metacognition (r = .64), task value and intrinsic goal orientation (r = .63) and intrinsic goal orientation and self-efficacy (r = .60). No or low correlations were found between other scales. At post-test correlations between the scales were small or non-existent, except for extrinsic goal orientation and task value (.59). Correlations between (the scales of) strategies and motivation in the post-test were absent or low, varying between .29 and .44.

Before the intervention (at pre-test) the students filled in the questionnaire at school, digitally and under the supervision of the researcher. After the intervention (at post-test), students received an email with a link to the digital questionnaire, so that they could fill in the questionnaire in their own time during the exam period. To obtain a realistic representation of their learning strategies, measuring instruments were implemented just before the students' exam period.

Think aloud task. We used a think aloud method to examine individual students' use of strategies while reading and studying for a test. This method was chosen because it may provide rich process-oriented data, and because it gives an insight into the responsive and constructive processes used by the readers. This instru-

ment also makes it possible to trace the thinking steps taken by those readers. Students' statements were recorded, transcribed and analyzed for the strategies used.

In each school, eight students (four from the experimental condition and four from the control group) participated in the think aloud task both before and after the intervention. In order to have a balance between below -, average and above average students, selection of participants took level into account. This was done on the basis of students' CITO scores in reading obtained in the 6th grade. The two conditions did not differ in CITO scores nor scores on History tests.

Before and after the intervention, each student was seated in a quiet room where he or she studied two History texts. Two think aloud sessions of 16 students (32 sessions in total) were organized for this. In selecting the participants, gender and level were taken into account, so as to have balance between girls and boys, and between below-, average and above average students. The reading level of the students was determined on the basis of their CITO scores.

One participant of the think aloud task was excluded from the post-test because he missed the last of the three lessons. In total, 15 participants were included in the analyses of the thinking aloud task, eight (males: 4) for the experimental condition and seven (males: 2) for the control condition.

Two History texts (each between 700 and 1000 words long) were selected from text books by the History teachers. The texts were first handed out on paper in the session, and students had to study them only once during the think aloud task. The texts were given in hard copy, as this most closely resembled the reality of the classroom. Students could choose whether they would use the textbook itself or a copy of the text.

Texts were selected on the basis of their relevance to the learners. None of the texts had been studied at school, but all of them were part of a real, up-coming exam two weeks later. This set-up was advantageous to both student and researcher. The student was enabled to study during school time and a few days before a real exam. This resulted in the most realistic situation for the researcher to investigate.

Students were instructed by means of a short film of a student solving a puzzle while thinking aloud, and a short explanatory text describing a think aloud task. In principle, the researcher remained silent, unless what the student meant was unclear, or unless the pupil seemed to experience problems (sighing, looking back in the text) but did not vocalize it. The researcher always sat next to or opposite the student so as to be able to accurately check the student's responses to the activities.

During the research, the students were instructed to study hard, as if they had to learn for an exam. Students were initially asked to say out loud everything they may think about while reading. This part of the research was filmed so that those responses could be accurately noted afterwards.

Learner report. The Learner Report (LR) was introduced by De Groot (1974) and discussed by Van Kesteren (1993) as "an instrument to identify educational objec-

tives, in particular unruly ones, that have obvious intentions but resist objective measurement" (p. 65). An LR can be used to uncover learning effects that are hardly or not at all measurable by more objective methods.

In this study, an open LR was used as a self-assessment tool. After the intervention lessons, students were asked to report what they had learned from the lessons. This could include generally applicable knowledge, skills, attitudes, self-knowledge and self-understanding. Students were free to write what they wanted and how much they wanted. To help the participants to get started, we used starting sentences, such as I discovered that… and I found out that I…, which participants could use to formulate their own experiences.

3. DATA ANALYSIS

3.1 Questionnaire

We computed the mean scores for all scales for both conditions at pre-test and post-test. We performed univariate tests to analyze the effects of the intervention on the self-reported strategy use and motivation, with condition (two levels) as between-subject factor, post-test scores as dependent variables and pre-test scores as covariates.

3.2 Think aloud task

In total, 30 think aloud protocols (stemming from 15 participants performing two tasks) were filmed, transcribed, segmented and scored for reading and studying strategies used by the students. One coder segmented the protocols and coded the segments. To parse the protocols into segments, each segment had to contain a new cognitive activity (e.g. segment 1, reading; segment 2, elaborating between two text fragments) or the same type of cognitive activity, but with new content (e.g. segment 1, reading normal text excerpt; segment 2, reading notes).

For scoring strategy use, we adapted the items originally included in the four cognitive scales by Pintrich to the students' context. These were rehearsal, elaboration, organization and critical thinking. We distinguished two metacognitive activities, namely checking understanding and evaluation. During coding, we added two more activities, namely consultation and remaining. The first four activities were subdivided into smaller units. This was needed to ensure proper coding. Figure 4 shows a description of the categories used to score the strategies, and the subdivision made within the categories.

Categories	Description and examples
1. Rehearsal	- repeating literal text (aloud) / re-reading - paraphrasing / summarizing text - copying literal text - writing a paraphrase of text
2. Elaborate	- elaborating within the text (between text blocks) - elaborating between text and notes/tasks/class - elaborating with / between other (kinds of) information
3. Organization	- organizing by using highlight / shading / bold words - rearranging into diagrams, charts and/or tables
4. Critical thinking	- evaluating information - comparing information to own ideas
5. Check understanding	- checking whether the text is understood
6. Evaluation	- evaluating learning goals before, during and after learning
7. Consultation	- consulting other sources such as Internet, dictionaries, etc.
8. Remaining	- making superficial comments like 'I'm going to read'

Figure 4. Categories of reading and studying strategies for coding the think aloud protocols

Four protocols were coded by two independent raters to determine the reliability of the coding. The proportion agreement was 83% (Cohen's kappa .68). Coders differed mainly on categories 2 (elaborating within the text) and 7 (consulting other sources). Consulting other sources was seen as elaboration with other sources by the second coder.

In figure 5 we present an excerpt from a protocol of one of the participants. The actions (in square brackets) or the statements of the participant are shown. Each new segment is placed in a different row. After the segment, the number of the code that corresponds with the statement or action is displayed (R = researcher).

I know that time is really important with History, so I write down the dates in my notes. Behind the dates I write down what started in those days. (13: evaluation)
[reads] I don't write down this whole story, that he was overpowered during his holiday. I don't think they are going to ask that in the test. (13: evaluation)
Then I simply write down that there is a conflict between church and state, which is spinning out of control. R: These are your own words? Yes. (4: writing paraphrase of text)
[reads] So, I'm reading one paragraph now, and the most important information I am going to write down in my own words. [writes] (4: writing paraphrase of text)
I think this piece really is about whether the pope or the emperor had absolute power. And I find their ideas rather important as well. The pope sees himself as a successor, because he was identified with Jesus. Therefore, everyone had to listen to him. And the emperor was actually the secular ruler... because of that, he wanted to have much influence as well (2: paraphrase the text)

Figure 5. Excerpt of a protocol of a student studying a History text

To analyze the think aloud data, we computed the mean percentages of all activities for both conditions at pre-test and post-test. We used univariate analyses to analyze the effects of the intervention on the strategy use, with condition as between subject factor, post-test scores as dependent variables, and pre-test scores as covariates.

3.3 Learner report

A total of 48 learner reports were collected, containing 250 statements about what had been learned (so called learning experiences). For four students learner reports were missing because they were absent. We coded the statements using six different categories, based on Janssen (1998): declarative knowledge, procedural knowledge and skills, positive learning experiences, negative learning experiences, metacognitive knowledge and skills, and a category "other experiences".

A typical declarative knowledge statement is "I know that elaboration is linking two different paragraphs". Here, the student knows the meaning of the word 'elaborate' in relation to studying texts. For procedural knowledge, a student may report that he knows how to perform a strategy. An example of such a statement is: "I know how to use the register in my History textbook". Positive learning experiences are statements such as "It is useful to write more while studying". A student wants to do something in the future or feels good about what he had learned. We coded statements like "I don't like to evaluate the way I learn" as a negative learning experience. The student learned something, but concludes he dislikes to think about the experience itself or to change anything in his study habits. For metacognitive experiences, participants had to evaluate their own learning, as is seen in, e.g. "I discovered that I need to think more critically about what I read to be-

come a better learner". We used the category "other experiences" for statements that did not fit into any of the abovementioned categories.

Mean percentages of learning statements were computed per category and per condition. We then used univariate analyses with condition as factor, and percentages of learning statements as dependent variables.

4. RESULTS

4.1 Self-Reported Learning Strategies

The different scales in the questionnaire were significantly correlated between pre- and post-test (between r=.42 and .63), with the exception of the task value scales. In table 2, we represent students' mean scores at pre- and post-test per condition.

Table 2. Scores on questionnaire per scale and per condition, at pre-test and post-test: means and standard deviations (1 = this does not apply to me at all, 5 = this does apply to me a lot)

Category	Moment	Observation		Practice	
		M	*SD*	*M*	*SD*
Rehearsal	Pre-test	3.6	.65	3.6	.72
	Post-test	3.7	.56	3.6	.60
Elaborate	Pre-test	2.9	.62	3.0	.72
	Post-test	3.0	.64	3.1	.64
Organization	Pre-test	3.0	.67	3.0	.79
	Post-test	3.0	.55	2.9	.82
Critical thinking	Pre-test	2.8	.52	2.6	.82
	Post-test	2.8	.47	2.7	.73
Metacognition	Pre-test	2.8	.48	2.6	.82
	Post-test	2.8	.49	3.0	.72
Intrinsic goal orientation	Pre-test	3.0	.85	3.1	1.0
	Post-test	3.0	.66	3.2	.88
Extrinsic goal orientation	Pre-test	3.2	.64	3.1	.78
	Post-test	3.2	.73	3.2	.70
Task value	Pre-test	2.7	.68	2.6	.85
	Post-test	2.9	.60	2.8	.77
Self efficacy	Pre-test	3.4	.51	3.6	.81
	Post-test	3.4	.55	3.6	.85

Table 2 shows that rehearsal is the most reported strategy in both conditions at pre-test and post-test. Metacognition has the lowest mean score compared to the other strategies. For the motivation scales, self efficacy has the highest mean scores for both conditions in pre-test and post-test.

Univariate analyses indicated that there was a condition effect for metacognitive activities (F (1, 43) = 4.63, p = .037). Students in the control condition reported significantly more metacognitive activities used during studying at post-test than students in the experimental condition.

4.2 Observed use of learning strategies

In Table 3, we present the mean percentages of students' strategies during the think aloud task at pre- en post-test.

Table 3. Percentages of learning activities per condition at pre- and post-test (mean percentages and standard deviations)

Category	Moment	Observation (n = 524)		Practice (n = 348)	
		M %	*SD*	*M %*	*SD*
Rehearsal	Pre-test	59,7	13,4	63,3	10,2
	Post-test	52,0	21.5	59,6	11,4
Elaborate	Pre-test	10,2	6,4	6,5	5,3
	Post-test	11,7	10,4	6,3	7,7
Organization	Pre-test	2,3	3,4	3,8	5,2
	Post-test	3,0	3,4	7,7	15,7
Critical thinking	Pre-test	4,1	5,6	0,2	0.6
	Post-test	5,4	6,4	1,3	3,4
Checking understanding	Pre-test	2,2	3,2	5,5	8,6
	Post-test	9,0	5,8	1,9	3,5
Evaluation	Pre-test	8,8	7,4	6,2	5,2
	Post-test	6,1	5,1	7,9	8,7
Consulting	Pre-test	3,2	3,8	2,7	3,5
	Post-test	1,4	1,6	1,0	2,6
Remaining	Pre-test	9,6	6,6	11,8	13,7
	Post-test	11,4	11,0	14,2	12,5

Table 3 shows that students predominantly performed rehearsal activities while studying history texts, in both conditions, at pre-test as well as post-test.

A univariate test indicated that students in the experimental condition were significantly more engaged in 'checking whether the text is understood' (see figure 3) than students in the control condition (F (1, 13) = 6.07, p = .028).

4.3 Students' self reported learning experiences

In the learner reports at post-test, students wrote down between 1 and 9 different learning experiences. It seemed that students in the experimental condition reported more different learning experiences in their learning reports (M = 5.6, SD = 1.9) than students in the control condition (M = 4.7, SD = 1.9). However, this finding did not reach statistical significance (F (1, 47) = 2.54, p = .12). In table 4, the mean percentages of self reported learning experiences per condition and category are shown.

Table 4. Percentages of learning experiences reported at post-test, per category and condition (mean percentages and standard deviations)

Category of learning experiences	Observation		Practice		Total	
	M%	*SD*	*M%*	*SD*	*M%*	*SD*
Declarative knowledge	28.7	25.9	27.8	25.8	28.3	25.8
Procedural knowledge	6.8	11.9	12.2	15.5	9.3	13.8
Positive attitude	19.0	19.5	20.4	25.7	19.7	22.3
Negative attitude	8.6	13.9	5.2	11.1	7.0	12.7
Metacognitive knowledge	32.2	24.6	20.5	21.6	26.9	23.8
Other experiences	4.7	9.1	13.8	22.9	8.9	17.3

As shown in Table 4, most of the learning experiences reported by students in both conditions refer to gains in declarative and metacognitive knowledge. Positive attitudes towards what had been learned were also frequently reported. Relatively few learning experiences could be labeled as knowledge of procedures (or "how to" information) or as negative attitudes. However, the large standard deviations indicate that there are large differences in reported learning experiences within the groups.

Univariate analysis showed that the difference in learning experiences between conditions did not reach significance at the level $p < .05$ However, students in the experimental condition tended to report more metacognitive learning experiences (on average 32.2%) than students in the control condition (on average 20.5%).

5. DISCUSSION

The aim of the present study was to examine whether observational learning can be an effective instructional method in the domain of content area reading and studying. Our main research question was: which instructional approach is more effective for text studying: learning by observation or learning by direct instruction and practicing? Effects were examined on students' attitude and use of cognitive and metacognitive strategies, while studying from school textbooks in a realistic context (preparation for an upcoming history test). We expected that the 'learning-by-observing' intervention would have a more positive effect on the actual use of and attitude towards (the use of) strategies than the 'learning-by-practice' intervention. In particular, we expected an increase in students' use of so-called depth strategies, which are related to high achievements (Wolters & Pintrich, 1998; Vermunt, 1992). We expected that training students by observing strong and weak peer models of their own age would be an effective manner of teaching students how to apply different kinds of strategies and extend their repertoire of approaches to studying texts.

Our research has yielded three results. Students's attitudes and self reported use of cognitive and metacognitive strategies was measured by a MSLQ-based questionnaire. We found that the control group reported significantly more metacognitive activities than the experimental group. Secondly, in the thinking-aloud task the actual use of strategies was measured. Students in the experimental, observational learning condition checked their understanding during studying (one of the metacognitive strategies) significantly more often than in the control group. Finally, learner experiences were again used for students' self-reported use and attitudes towards strategies and studying. We used a learner report only at post-test. No significant differences were found ($p < .05$), but as is shown in table 4, the experimental group tend to report more metacognitive learning experiences than the control group. The experimental group reported more metacognitive learning experiences than the control group.

We can tentatively conclude that our expectations are partly confirmed. The experimental condition used one of two metacognitive activities (checking understanding) significantly more often after the intervention than the control condition. Not only did the experimental group use this strategy more often, they also reported more metacognitive learning experiences after the intervention in their learner reports. In their statements, they reflect and/or evaluate their own way of learning, for example; 'I have chosen a few strategies which suit me well, I think that I am going to apply these ones'. This corresponds with our expectations that the experimental condition would benefit more from the intervention than the control condition, in particular in relation to the use of depth strategies.

The outcome of the questionnaire seems to contradict these results. The control group reported more metacognitive activities than the experimental group. It should be noted, however, that the reliability of the metacognitive scale was not

high; this scale showed the lowest reliability scores, both at pre-test and post-test. Future research should include careful examination of the metacognitive scale of the questionnaire, as, in our opinion, the items measure rather diverse activities. For example, two out of the five items in this scale describe preconditions for learning; one of them deals with study planning, whereas another question is about students taking steps after a class if their notes are unclear. Obviously, these two activities could not be measured during the think aloud task, since this task focused on "online" reading and study behavior. Possibly, the diversity of the scale was responsible for its low reliability. Immediately before or during their studying, students not questioning their notes as far as clarity is concerned may very well reflect the manner in which they study. In future research, the questionnaire should be edited and adapted, so that it is more consistent with the metacognitive strategies that could logically be employed by students during their reading and studying of texts.

Another solution might be to use a different instrument. After all, a questionnaire always involves rating one's own behavior. It is possible that what students say they do is different from what they *actually* do. Our findings seem to indicate that the control group overestimated themselves; the increase in self reported metacognitive strategies does not correspond with the observed behavior during the thinking aloud task. It might be that students in the control condition overestimated themselves more than in the experimental condition, because they had not observed peers in action, and therefore could not compare themselves to others.

Furthermore we noticed a lack of motivation of students to fill in the same questionnaire twice. The second time the students were far less motivated to fill in the questionnaire than the first time. This seems to be borne out by the low reliability scores in the post-test.

Another limitation of this study concerns the generalizability. We limited ourselves to one school subject: history. Furthermore, the sample size was quite small (N=52). The think aloud task was only performed by 16 students and a limited number of History schoolbook texts were used. Our findings showed a large variety in use of strategies within both groups. This problem could be remedied by increasing the number of students in each condition. Due to time constraints it was not possible in this research.

In this research, no distinction was made between weak and strong participants. It would be interesting to investigate whether the interventions affect different type of students (e.g. of different ability and gender) in different ways. However, to examine interaction effects between condition and level of student and/or gender, more participants are necessary.

An explanation for not finding large differences between conditions might be that the control condition in this study was not an untreated control group. In fact, this group received a rather strong intervention. We suggest that future research include a third condition, i.e. one that does not contain an intervention. This would make ascribing possible effects to the interventions more reliable. It is possible that

"growing up" and personal development have an effect on the application of depth strategies in particular.

Despite these limitations we can tentatively conclude that strategy training through observational learning seems to have beneficial effects. Among others, observing comparable (role) models appears to be a promising technique for mastering text study strategies.

REFERENCES

Bandura, A. (1986). *Social foundations of thought and action: a social cognitive theory*. Englewoord Cliffs, NJ: Prentice Hall.

Bimmel P., & Van Schooten E. (2004). The relationship between strategic reading activities and reading comprehension. *L1 – Educational Studies in Language and Literature, 4*, 85-102. http://dx.doi.org/10.1023/B:ESLL.0000033844.88918.e7

Blom, S. & Hoek, D.J., & Dam, G.T.M. ten. (2007) Metacognitieve zelfregulatie, motivatie en perceptie van klassenklimaat. Zijn er sociaal-culturele verschillen? [Metacognitive selfregulation, motivation and perception of class condition. Are there social-cultural differences?] *Pedagogische Studien, 84*, 20-36.

Braaksma, M.A.H., Rijlaarsdam, G., Van den Bergh, H., & Van Hout-Wolters, B.H.A.M. (2004). Observational learning and its effects on the orchestration of writing processes. *Cognition and Instruction, 22*(1), 1-36. http://dx.doi.org/10.1207/s1532690Xci2201_1

Braaksma, M.A.H., Rijlaarsdam G., & Van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology, 94*, 405-415. http://dx.doi.org/10.1037/0022-0663.94.2.405

Brand-Gruwel, S., Schellings, G., & Teurlings, C. (1998). Leren begrijpend en studerend lezen [To learn understanding and studying reading]. In L. Verschaffel, & J.D. Vermunt (red.), *Het leren van leerlingen. Onderwijskundig Lexicon Editie III* (pp. 81-97). Alphen aan den Rijn: Samsom.

Couzijn, M. J. (1999). Learning to write by observation of writing and reading processes: effects on learning and transfer. *Learning and Instruction, 9*(2), 109-142. http://dx.doi.org/10.1016/S0959-4752(98)00040-1

Couzijn, M. J. (1995). *Observation of writing and reading activitie: Effects on learning and transfer.* Unpublished dissertation. University of Amsterdam.

De Groot, A.D. (1974). The problem of evaluating national educational systems. In H.F. Crombag & Data N. de Gruyter (Eds.), *Contemporary issues in educational testing* (pp. 9-27). The Hague: Mouton.

Groenendijk, T., Janssen, T., Van den Bergh, H., & Rijlaarsdam, G. (2013a). The effect of observational learning on students' performance, processes and motivation in two creative domains. *British Journal of Educational Psychology*, *83*, 3-28. DOI: 10.1111/j.2044-8279.2011.02052.x

Groenendijk, T., Janssen, T., Van den Bergh, H., & Rijlaarsdam, G. (2013b). Learning to be creative. The effects of observational learning on students' design products and processes. *Learning and Instruction*, *28,* 35-47. http://dx.doi.org/10.1016/j.learninstruc.2013.05.001

Hettinger-Steiner, H., & Carr, M. (2003). Cognitive development in gifted children: toward a more precise understanding of emerging differences in intelligence. *Educational Psychology Review, 15*(3), 215-246. http://dx.doi.org/10.1023/A:1024636317011

Janssen, T., Ten Dam, G. & Van Hout-Wolters, B. (2002). *Vaardigheden voor zelfstandig leren. Een praktijgericht overzicht van onderzoek*. Koninklijke Van Gorcum, Assen.

Janssen, T. (1998). *Literatuuronderwijs bij benadering: Een empirisch onderzoek naar de vormgeving en opbrengsten van het literatuuronderwijs Nederlands in de bovenbouw van het havo en vwo.* Dissertatie Universiteit Utrecht. Amsterdam: Thesis Publishers.

Kniep, J., & Janssen, T. (2011). Studerend lezen in 4 havo: hoe doen leerlingen dat? [Studying reading in 10th grade: how do students do that?]. *Levende Talen Tijdschrift*, *12*(1), 26-36.

Pintrich, P.R., Smith, D.A.F., Garcia, T., & McKeachie, W.J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).* National Center for Research and Improvement, Washington DC.

Pintrich, P.R., & De Groot, E.V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 22-40. http://dx.doi.org/10.1037/0022-0663.82.1.33

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale NJ: Erlbaum.

Schellings, G., Aarnoutse, C. & Van Leeuwe, J. (2006). Third-grader's think-aloud protocols: Types of reading activities in reading an expository text. *Learning and Instruction, 16*, 549-568. http://dx.doi.org/10.1016/j.learninstruc.2006.10.004

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science, 26*, 113-125. http://dx.doi.org/10.1023/A:1003044231033

Schunk, D.H., & Zimmerman, B.J. (1997). Social origins of self-regulatory competence. *Educational Psychologist, 32*(4), 195-208. http://dx.doi.org/10.1207/s15326985ep3204_1

Van Kesteren, B.J. (1993). Applications of De Groot's "learner report": A tool to identify educational objectives and learning experiences. *Studies in Educational Evaluation*, *19*, 65-86. http://dx.doi.org/10.1016/S0191-491X(05)80057-4

Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs: naar procesgerichte instructie in zelfstandig denken*. Amsterdam: Swets & Zeitlinger.

Wang, M.C., Haertel, G.D., & Walberg, H.J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, *63*, 249-294. http://dx.doi.org/10.3102/00346543063003365

Wolters, C.A., & Pintrich, P.R. (1998). Contextual differences in student motivation and self-regulated learning in mathematics, English and social studies classrooms. *Instructional Science*, *26*, 27-47. http://dx.doi.org/10.1023/A:1003035929216

Zimmerman, B.J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology*, *94*(4), 660-668.

---

[i] *CITO is a standardized – and soon to be mandatory – test used in the last year of primary education in the Netherlands.*