Observational learning in argumentative writing

Braaksma, M.A.H.

Publication date
2002

Document Version
Final published version

OBSERVATIONAL LEARNING
IN ARGUMENTATIVE WRITING

UNIVERSITY OF AMSTERDAM
Graduate School of Teaching and Learning

Cover design: Bart Westerman.
Cover photo adapted from: Gustave Courbet (1819-1877). Interior of my atelier, a real allegory that summarized seven years of my life as artist (1854-1855), Louvre, Paris.
Printed by PrintPartners Ipskamp B.V., Amsterdam.

ISBN 90-9015630-5

This publication was partly supported by grants from the Graduate School of Teaching and Learning, University of Amsterdam, the International Association for the Improvement of Mother Tongue Education (IAIMTE), and the Braaksma family.

ico

OBSERVATIONAL LEARNING
IN ARGUMENTATIVE WRITING

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. mr. P.F. van der Heijden
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit

op dinsdag 21 mei 2002, te 12.00 uur

door

Martine Anne Henriëtte Braaksma

Geboren te Choma (Zambia)

Promotores:	Prof. dr. G.C.W. Rijlaarsdam
	Prof. dr. B.H.A.M. van Hout-Wolters
Co-promotor:	Dr. H.H. van den Bergh
Faculteit:	Maatschappij- en Gedragswetenschappen
	Het promotie-onderzoek is voorbereid aan het Instituut voor de Lerarenopleiding van de Universiteit van Amsterdam

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Over the years I worked on this thesis, many people supported me. I thank them all!

First of all I thank my supervisors. My promotor Gert Rijlaarsdam introduced me to the ‘world of science’, and has always been there for me. I thank him very much for his encouragement, comprehensive help, stimulating feedback, and his personal involvement. I thank my promotor Bernadette van Hout-Wolters for her trust and the opportunities she gave me. I am very grateful to my co-promotor Huub van den Bergh who never tired of answering my questions. Not only was he essential for the statistical analyses, he was also very much involved in the other phases of the project, even in evening hours and weekends.

I thank Michel Couzijn for establishing a research framework that allowed me to start immediately with the first study. Additionally, during the project he provided valuable feedback and helpful advice.

I am grateful to all my colleagues at the Graduate School of Teaching and Learning for the pleasant work climate, especially the PhD students and the other ‘lunch people’ for the enjoyable talks, lunches, and ‘pre-OZO’s’. My special thanks go to the PhD students ‘from the first hour’ who helped me to find my way in the institute and in research: Sabine Severiens, Dirk Hoek, and Hein Broekkamp. I thank my roommates, first Sabine and later Monique Pijls, for being such pleasant company. Monique: I hope we will stay for a long time on the ‘base’. Furthermore, I thank Hein for his great ‘thinking along’ with my project and for his comments on chapters of this thesis.

The combination of research and teaching has been both possible and pleasant due to the cooperation of my colleagues Michel Couzijn, Wilma Groeneweg, Simone Löhner, and especially Else Pistoor to whom, as my teacher mentor, I owe much.

Without the participation of students, this project would not have succeeded. I thank all the students and teachers of the different schools for participating in my studies. Special thanks go to Marianne van der Veer, Maria Pollmann, and the school management of the Open Schoolgemeenschap Bijlmer.

Next, I want to thank a lot of people for helping me in research sessions, typing of protocols, and coding of data: Camilla Meijer, Mariëlle de Reuver, Janneke Jehee, Pluis Draisma, Marianne Braaksma, Annemarie Schumacher, Merlin Rijlaarsdam, Monique Oechies, Patricia Braaksma, and Esther Hartman. I am also very grateful to Petra Wildbret for her support.

Furthermore, my gratitude goes to the languages’ researchers at the Graduate School of Teaching and Learning for the useful discussions and their comments on chapters of this thesis. Especially, the feedback from Tanja Janssen has been of a great value.

I also want to thank my paranymphs Ineke Jongen and Marianne Braaksma for their support and friendship.

I thank my parents Eertze en Ria Braaksma for their enduring interest and encouragement in all that I am doing.

Finally, I thank Bart Westerman, who has stimulated me over the years.

CHAPTER 1

INTRODUCTION

1 LEARNING TO WRITE

The most striking problem in writing is cognitive overload. This has been identified in many studies and many metaphors have been used to describe this phenomenon. Flower and Hayes (1980), for instance, refer to 'juggling with constraints', Van den Bergh and Rijlaarsdam (1999) label this phenomenon as 'coordination problems', others like Bereiter and Scardamalia (1987), Kellogg (1994), Alamargot and Chanquoy (2001) use the term 'cognitive overload' or 'limited memory capacities'. All authors mean that a writer has to carry out too many processes, or has to pay attention to too many different textual characteristics simultaneously and therefore loses track of his own thoughts. An illustrative protocol fragment comes from Bereiter and Scardamalia (1987), who quote a writer who says: 'When I come to a word I can't spell, it throws me off my writing' (p.155).

This phenomenon of cognitive overload applies particularly to writers who are *learning* to write. They have indeed to juggle with constraints; they have to carry out two tasks simultaneously: they have to write a good text and learn from their writing as well. Several instructional attempts have been made to stimulate learners to step back and perform reflective activities, in order to distinguish writing from learning and thereby allowing for application of cognitive recourses to writing and learning successively. A well-known approach is to add a phase of peer feedback and revision into the instructional process (Hillocks, 1986; Rijlaarsdam, 1986, 1987; Rijlaarsdam & Couzijn, 2000a). In general, this approach stimulates writers to step back, and to act as a reader and commenter of texts written by peers, assuming that commenting on other texts transfers to the revision phase of their own texts and to their next writing (Hillocks, 1986). Another approach to promote the use of metacognitive strategies and stepping back is the Self-Regulated Strategy Development approach (SRSD) that Graham and his colleagues developed (Graham & Harris, 1994; Graham, Harris, MacArthur & Schwartz, 1998; Graham, Harris & Troia, 1998). With SRSD students are supported to master higher level cognitive processes associated with successful writing, to promote reflective self-regulation of writing performances and to develop positive attitudes regarding the writing process and themselves as writers. In one of the seven stages of this instructional metascript the teacher models how to use a certain strategy, and learners observe, analyze, and discuss the effectiveness and efficiency of the modeled strategy.

Although Hillocks (1986) in his meta-analysis concludes that peer review is a relatively promising method, the research results are not equivocal (Hillocks, 1986, 219-220; Van den Bergh & Meuffels, 2000). In some studies, positive effects have been shown whereas in others no effects or even negative effects of the manipulations have been assessed. Furthermore, results are not always clearly attributable to the relevant manipulations. In the case of peer review, it is not always clear whether effects are due to giving comments, reading comments, or rewriting. In the SRSD studies, it is not clear whether reported gains are related to gains in control groups with an equal amount of training or on pre skill scores. Of course, the first is a more stringent criterion than the latter. In addition, in the SRSD studies it is far from clear to which of the manipulations (e.g., pre skill development, discussion of the strategy, modeling of the strategy, memorization of the strategy, collaborative practice) the reported effects have to be attributed.

Besides these more technical comments, I would like to stress some more theoretical points. In peer review, the texts and the comments on texts are the central issues. Students have to make inferences about possible improvements in their own writing processes based on these two aforementioned characteristics. As there is no one-to-one relation between textual characteristics and cognitive processes this gap can be quite wide. For the second type of instruction, SRSD, a long program is defined, with many instructional elements, aimed at the development of a feeling of proficiency. These feelings of proficiency, the feeling of being a self-regulated writer and the positive feeling of self-efficacy, will lead to better writing. However, the chain from feelings of proficiency via fluent writing processes to adequate texts is a feeble one, which has not much empirical support (Kuhlemeier, Van den Bergh & Rijlaarsdam, 1997; Van den Bergh, Rijlaarsdam & Kuhlemeier, 1997). Furthermore, the effects of the trained and modeled strategies like brainstorming, planning etc. hardly address causality. Finally, the effects of such strategies on writing are usually not very large (compare Hillocks, 1986).

A method that allows for a distinction between writing and learning to write, and explicitly allows for reflective activities and a direct link between writing processes and the resulting writing product is observational learning. Observational learning has been shown to be effective in writing education (Couzijn, 1995, 1999). This method has at least one feature in common with the two approaches mentioned above: the key learning activities are observation, and the subsequent analysis, evaluation and reflection. However, the important difference is the *lack of writing*. In this approach learners do not write, but instead observe writing processes of a model and the resulting texts. The cognitive effort is shifted from *executing* writing tasks to *learning* (Couzijn, 1999; Rijlaarsdam & Couzijn, 2000a, 2000b).

From a practical point of view, observational learning seems like an ideal candidate for the operationalization of process oriented writing as prescribed in (at least) one of the so-called attainment targets (Ministerie van OCenW, 1998, p. 27):

> 'Students are able to critically consider (own) texts and revise them if necessary. It concerns:
> - correct and judge a (own) text;
> - make suggestions for improvement of a text;
> - rewrite an own text, also based on reactions and suggestions of others;
> - draw conclusions for the performance of next writing tasks.'

This attainment target implies writing education with – among others – the following features: focus on writing processes and writing strategies; reflection on writing context, process, and product; commenting and discussing texts, involving teachers and peer students (Hoogeveen & Bonset, 1998).

2 RESEARCH ON OBSERVATIONAL LEARNING

There is much research about the effects and conditions of observational learning and modeling (Bandura, 1986, 1997; Rosenthal & Zimmerman, 1978; Schunk, 1987, 1991, 1995, 1998). Observational learning, with either teachers, adults or students as models, has proven to be effective with students of various ages and in various school subjects, such as mathematics (Schunk & Hanson, 1985, 1989a, 1989b; Schunk, Hanson & Cox, 1987), reading (Couzijn, 1995, 1999), writing (Couzijn, 1995, 1999; Graham & Harris, 1994; Graham, Harris & Troia, 1998; Schriver, 1992), and speaking and listening (Sonnenschein & Whitehurst, 1984). Observation of models also can raise observers' self-efficacy, or personal beliefs about their capabilities to learn or perform behaviors at designated levels (Bandura, 1986, 1997). Schunk and his colleagues (Schunk & Hanson, 1985, 1989a; Schunk et al., 1987; see also Schunk, 1998, p. 148) reported effects of (various) models on students' self-efficacy, which in turn influenced learning and achievement.

The effectiveness of observational learning depends on a number of instructional factors, for example, model's age, model's competence, and number of models (Schunk, 1991). In general, for students, *peer models* are preferred to teacher models because peer models can give the observers an impression that the observed behavior is within reach ('if he can do it, then so can I'). Two types of peer models can be implemented: coping models and mastery models. *Coping models* initially demonstrate the typical fears and deficiencies of observers but gradually improve their performances and gain confidence in their capabilities. These models illustrate how directed effort and positive self-reflections may overcome difficulties. *Mastery models* on the other hand, demonstrate rapid learning and make no errors. They may also verbalize statements reflecting positive attitudes, and high levels of confidence and ability. In general, *multiple models* are preferred because they, compared with a single model, increase the probability that observers will perceive themselves as similar to at least one of the models (Schunk, 1987, 1991).

However, not only the instructional factors in themselves are important, but also the personal characteristics of observers (e.g., age, sex, competence) may influence the effectiveness of observational learning. For instance, the effectiveness of peer

modeling is hypothesized to depend in part on perceived similarity between model and observer (Schunk, 1987, 1998).
In many studies, observational learning is combined with other didactical interventions. For instance, observational learning is combined with training sessions, or feedback from teachers (Schunk, 1998; Graham & Harris, 1994). Instead, in the current thesis, I have tried to isolate the effects of observational learning, in the domain of argumentative writing.

3 THEORETICAL FRAMEWORK

Figure 1.1 presents the theoretical framework of this study about the effects of observational learning on writing. Observational learning changes the learners' knowledge about writing. These changes have repercussions on the organization of writing processes, which in turn affects the quality of the writing product. As explained in section 2, learner characteristics influence instructional aspects of observational learning.

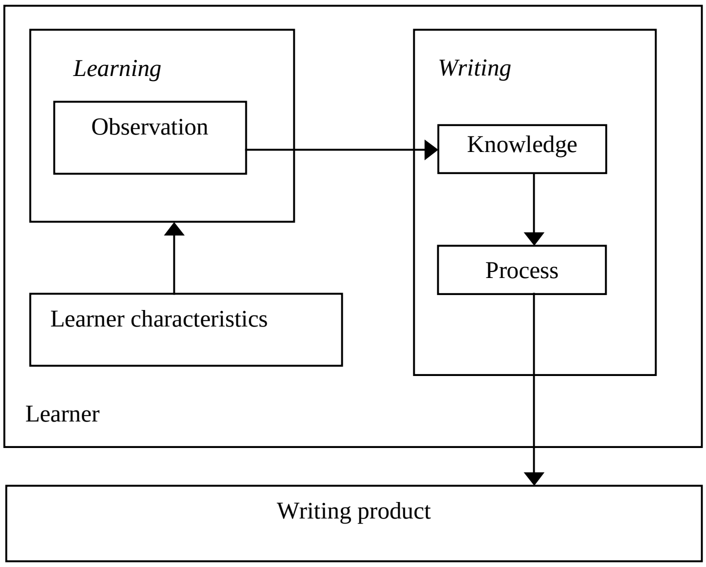


Figure 1.1. Theoretical framework

In the next sections, I will globally explain the relation between observational learning and writing. In the next chapters, the different parts of the theoretical framework will be dealt with more thoroughly.

3.1 Influence of observational learning on knowledge about writing

The knowledge of writers (stored in their long term memory) consists of different kinds of knowledge that ensure different functions within the writing activity: knowledge of genre, audience, linguistics, topic, and task schemas (Hayes, 1996). It is assumed that this knowledge is constructed by regulatory or metacognitive strategies such as personal observations, evaluations, and reactions (Graham & Harris, 1994). These strategies play a key role in the feedback loop in which one learns from the consequence of one's actions (Zimmerman, 1989, 2000; Zimmerman & Schunk, 1989). For instance, students who observed and evaluated their own or others' writing by using criteria (e.g., in the form of questions) appeared to internalize some of these criteria, and this new knowledge was then used to guide the production of future compositions (Hillocks, 1986).

Thus, using metacognitive strategies, self-regulated writers gain information that changes what they know and do (Graham & Harris, 1994, p. 209). A requirement for this self-regulated learning is 'that learners have reached developmental levels sufficient to be able to step back and consider their own cognitive processes as objects of thought and reflection' (Brown, 1987, p. 68).

Assuming that metacognitive processes feed the knowledge about writing, an important question is which instruction method stimulates the use of metacognitive strategies and enables writers to step back. I expect that observational learning is such an instruction method. In observational learning students are stimulated to use and address metacognitive strategies explicitly, because the observation of the performance of others involves a 'natural' step back and thus a natural type of monitoring, evaluation, and reflection on task execution processes (Rijlaarsdam & Couzijn, 2000a). Furthermore, in observational learning students can relatively more easily step back and use metacognitive strategies than in more traditional ways of instruction, where cognitive effort is directed to writing instead of learning to write. When students learn to write, they have to juggle with two tasks: a writing task and a learning-to-write task. In a way, they have to attain two goals simultaneously: they must produce an adequate text, and acquire new understanding about writing. When students learn to write by observation, cognitive effort shifts from *executing* writing tasks to *learning* (Couzijn, 1999; Rijlaarsdam & Couzijn, 2000a, 2000b). In observational learning, students do not have to juggle with the dual task of both writing and learning. They can focus on the learning task, creating a learning opportunity to enlarge their knowledge about writing.

3.2 Influence of knowledge on writing processes

Imagine a student who has to write an argumentative problem-solution text about a measure to obligate every adult to be an organ donor. If this student is a skilled writer, he is likely to have developed knowledge about writing. This includes knowledge about writing processes, as well as knowledge about text structures. Skilled writers know schema structures and their ingredients (cf. Overmaat, 1996; Steehouder et al., 1999) and can make use of these schemas. Therefore, this type of knowledge influences the student's writing process. For instance, protocol fragments like the following will be observed. Please note the large amount of planning activities in the following fragment.

> 'I can use the "measure structure" for my text. I will build my text according the following questions: (a) What is the measure? (the obligation for every adult to be an organ donor), (b) Why do we need that measure? (because of the lack of donors and because otherwise people forget to fill in a form), (c) How will this measure be realized? (automatically every adult is a donor. Only if one really has objections, one can ask for an exception). I will end with describing the effects of the measure.'

Thus, knowledge about writing may affect writing processes (see Scardamalia & Bereiter, 1986, p. 784 for more examples). This assumption is supported by several studies. For genre knowledge for instance, it is supposed that familiarity with a genre influences writing by providing access to an organized scheme in long term memory (McCutchen, 2000, p.17), although Rijlaarsdam, Van den Bergh and Zwarts (1992) could not empirically verify this point of view. Torrance (1996), however, gives a more plausible interpretation of the effects of genre knowledge on writing, by explicitly linking genre and topic knowledge; expert writers are experts in a specific field. He states that (genre) knowledge plays an important role in writers' cognitive processes; writing expertise is often highly dependent on knowledge of genre (Torrance, 1996, p. 5). Scardamalia and Bereiter have also stressed the importance of discourse schema knowledge that writers bring to their writing (Scardamalia & Bereiter, 1986, p. 783-785).

The influence of genre knowledge on writing has been examined in several studies. These studies showed that poor writers differ from skilled writers in 'discourse knowledge' (Ferrari, Bouffard & Rainville, 1998), 'general knowledge of text structures or genres' (McCutchen, 2000), 'text structure knowledge' (Englert, Stewart & Hiebert, 1988), and 'knowledge about structure of composition' (Schoonen & De Glopper, 1996). In addition, Hillocks (1986) speaks for the influence of knowledge on writing. He concluded that students who were given many opportunities to apply information about scales and criteria in judging writing, appeared to gain procedural or operational knowledge of the criteria that influenced their own independent writing (Hillocks, 1986, p. 224).

3.3 Influence of writing processes on writing products

The organization of writing processes appeared to influence the quality of writing products. Studies that relate writing processes to the quality of the written outcomes, showed that weak and good writers rely on the same cognitive and metacognitive activities, but differ in the way they distribute these activities over the writing process (Breetvelt, Van den Bergh & Rijlaarsdam, 1994; Rijlaarsdam & Van den Bergh, 1996; Van den Bergh & Rijlaarsdam, 1999, 2001; Van der Hoeven, 1997). Weak writers are likely to start a certain activity too early or too late. For instance, they tended to be too early in setting goals (i.e., formulating task demands derived from the assignment or self-devised) and in revising text, and too late with reading and interpreting the topic information and documentation, and with evaluating mental text production) (Breetvelt et al., 1994).

4 RESEARCH ISSUES

At the start of this chapter, I wrote that observational learning has proved to be effective in different school subjects, including writing where I referred, among other studies, to the studies of Couzijn (1995, 1999). The current thesis is a continuation of Couzijn's studies and the research issues in this thesis result mainly from Couzijn's outcomes.

Couzijn (1995, 1999; see also Couzijn & Rijlaarsdam, 1996) conducted several experimental studies to assess the effects of observational learning. He compared observational conditions in which students observed writing processes and products with conditions in which students performed writing tasks themselves (learning-by-doing). He showed that observational learning was more effective for writing than learning-by-doing. However, Couzijn's studies about observational learning can be characterized as 'input-output' research. He studied the effects of observational learning on writing products and did not focus on underlying processes. In the current thesis, my main goal is to get more insight in these underlying processes. I will focus on three issues (these issues largely coincide with the boxes and arrows as presented in Figure 1.1):

- *The observational process*. There is a need to know how observation contributes to learning gains. How do students process observation tasks? Which elements in observation tasks contribute to learning gains?
- *Influence of learners' characteristics on the effectiveness of the instructions for observational learning*. The effectiveness of observational learning depends on a number of instructional and personal factors (Schunk, 1987, 1991, 1998). Couzijn focused on a relatively homogeneous group of students (relatively high ability) and examined one type of instruction for observational learning. Students observed pairs of models and had to reflect on the weaker model of each pair. One could ask whether all students learn from reflecting on the weaker model, especially if they are good students. Thus, the issue examined is which type of instruction is effective for which type of student.
- *Effect of observational learning on the mediating variable 'writing processes'*. Couzijn found effects of observational learning on writing products. He assumed – as in the theoretical framework in Figure 1.1 – that these effects were caused by changes in writing processes. In his studies, this assumption was not studied. My contribution to the theoretical framework is to study the effect of observational learning on writing processes.

5 STUDIES IN THIS THESIS

This thesis consists of six chapters, and describes four different studies about observational learning in argumentative writing. Chapters 2 to 5 address the studies that examine the issues mentioned in the previous section. Each chapter has been submitted as an article to an international journal[1]; one has been published (chapter 2), one is in press (chapter 3), the others are under review.

Chapter 2 focuses on the first issue, the observational process, and addresses the question of which elements in observation tasks contribute to learning gains. This question will be answered in a study in which a post-hoc analysis is performed on Couzijn's data.

Chapter 3 regards issue 2, the influence of learners' characteristics on the effectiveness of the instructions for observational learning. The chapter describes an experimental study in which a heterogeneous group of students participated (mixed ability). In that way it was possible to examine which type of instruction (two types of observational learning or learning-by-doing) was effective for which type of student.

In chapter 4, I will pay attention to issue 3 and will study the effects of observational learning on writing processes. To examine that issue an experimental study was set up in which participants learned to write argumentative texts by observation or by doing. To measure writing processes, the participants performed writing posttests while thinking aloud. In that way, possible effects of instruction on writing processes could be assessed.

Chapter 5 also focuses on the first issue, the observational process. Instead of a post-hoc analysis (chapter 2) I carried out a case study to examine in more detail how learners process observation tasks as implemented in the studies reported in the previous chapters. Students thought aloud while carrying out observation tasks and were afterwards interviewed about their behavior. Special attention is paid to find out whether differences in observational conditions elicited different learning behaviors.

Finally, in chapter 6 the results of the studies are synthesized and discussed. Directions for future research and implications for educational practice complete this thesis.

1 Writing a thesis in articles has both advantages and disadvantages. Clear advantage is that each chapter can be read separately. A disadvantage is that there is sometimes overlap between chapters and that the consistency in, for instance, names of conditions is not maximized.

CHAPTER 2

EFFECTIVE LEARNING ACTIVITIES IN OBSERVATION TASKS

A post-hoc analysis

Abstract

On repeated occasions, observational learning has proved itself to be an effective instruction method. Experimental studies have shown to be effective for complex tasks such as reading and writing for both teachers and students as models. The problem when interpreting the results of such research is that, in observation tasks, several mental activities play a simultaneous role. In this study we therefore set out to identify the effective elements of observation tasks. We focused on two elements of the observation tasks, both aimed at stimulating monitoring activities: evaluation of the model's performance and elaboration on this evaluation. We have also distinguished between elaboration on the observed products (the models' written answers), and elaboration on the observed processes (the models' verbalizations of their mental activities).

The data were subjected to a LISREL analysis. First of all, it was observed that subjects who performed 'evaluation' and 'product-elaboration' better, and 'process-elaboration' more often in one lesson, also performed these activities better or more often in the subsequent lesson. Next, we observed an effect of aptitude on the learning activities: pre skill scores influence 'evaluation' and 'product-elaboration'. The most important finding is that 'evaluation' and 'product-elaboration' contribute positively to argumentative writing skills. It is discussed that these findings confirm the importance of the monitoring, evaluative and reflective activities when learning complex tasks as writing.

1 INTRODUCTION

The most severe problem for learners when learning complex tasks such as writing is switching between carrying out the writing task itself and learning from doing so. Learning to write requires the learner to become so closely involved in the writing process that hardly any cognitive energy is left for learning from that process. Below, we suggest that this problem might be overcome, and learning thereby be greatly facilitated, if, rather than actually performing the task, the performance of that task was observed.

Braaksma, M. A. H., Van den Bergh, H., Rijlaarsdam, G., & Couzijn, M. (2001). Effective learning activities in observation tasks when learning to write and read argumentative texts. *European Journal of Psychology of Education,* 1, 33-48.

When a learner performs a writing task, three cognitive components are involved. Firstly, the *executional component*, consisting of orientation, executing and text revising activities. This results in a solution to the set task, i.e., a text which is a proper reflection of the assignment (Kellogg, 1994). The second component is the *monitoring component*, which consists of observing (i.e., monitoring), evaluative and reflective activities (Hayes, 1996). These activities are aimed at acquiring online knowledge of one's actual task behavior, and are guided by the goals the learner-writer set at the beginning and during the execution of the writing task. The third component is the *regulative component*. The regulator controls the effective temporal organization of the monitoring and executional component, and creates new strategies based on learning outcomes (Figure 2.1).

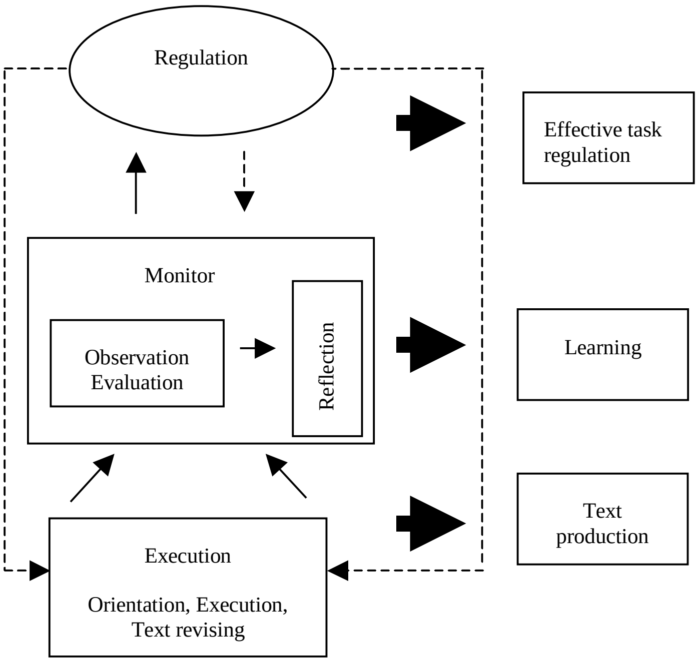


Figure 2.1. Three components of performance of writing (based on Couzijn, 1999, p. 111) Note. Straight arrows indicate the flow of information between activity categories, and dotted arrows indicate activation prompts.

Schunk and Zimmerman (1994, 1998; Zimmerman & Schunk, 1989; Zimmerman, 1989, 2000) stress the importance of monitoring, evaluative and reflective activities during the learning of complex skills such as mathematics and essay writing. The monitoring component provides learners with opportunities to become aware of the writing strategies they have used. Learners identify and conceptualize these strategies and their effects (they 'label' them), and add negative and/or positive evaluations of these effects. The output of the monitor (i.e., writing experiences and evaluations) provides the input for reflection, and thus learning (Figure 2.1). The learner re-conceptualizes and re-evaluates the writing behavior and working-method, adds situation-specific knowledge to the writing strategy (i.e., techniques and rules which are the object of practice), and refines strategies for use in new tasks (Couzijn, 1995; Oostdam & Rijlaarsdam, 1995; Rijlaarsdam & Van den Bergh, 1996).

Writing a text and learning to write a text go hand in hand. We assume that both processes rely on the same set of cognitive activities: execution, monitoring and regulation. Metaphorically, one might say that learners have to juggle with two agendas: a writing agenda and a learning-to-write agenda. In a way, they have to attain two goals simultaneously: they must produce a text that is adequate in terms of communication, and they must acquire new understanding about writing (Rijlaarsdam & Couzijn, 2000a).

In a learning-to-write situation, a learner has to regulate many cognitive activities, both for the writing agenda and for the learning agenda. For each of the agendas, he must alternate effective executive, monitoring and regulative activities. And he must pay special attention to the identifying and evaluating activities of the writing agenda, because they result in learning. However, in educational practice, the writing process dominates the learning process. Not only is it difficult for learner-writers to pay attention to both agendas simultaneously, neither is such attention stimulated in the present writing curriculum. Students whose goal is to write an effective text will not have enough cognitive (regulation) capacity to learn from the writing activity. Due to the cognitive cost of paying attention to both agendas, most learner-writers in a 'learning-by-doing environment' have to focus on (short-term) writing task performance, rather than on (long-term) learning performance. These learners consider their task as a 'job' that needs to be finished: a text has to be produced, and that is it. Therefore, while their attention is focused on the writing product, or sometimes on the writing process, they will not (and maybe cannot) invest cognitive resources in attaining learning goals. So it is not really possible to establish the learning opportunity the writing assignment is intended to be. The question is therefore whether other instructional methods might provide a learning environment which required less executive and more monitoring, evaluative and reflective activities.

1.1 Learning-by-observation

If monitoring, evaluative and reflective activities are important for learning complex skills, whereas cognitive resources are used for task-execution activities (i.e., writing) rather than for learning, instructional methods such as practicing are not the most effective methods. 'Learning-by-observation' might therefore be an instruction method that is better focused on these monitoring, evaluative and reflective activities.

'Learning-by-observation' is the learner-oriented counterpart of modeling. Modeling and observational learning have been studied extensively by Bandura (1986, 1997), Schunk (1987, 1991, 1995, 1998), and Rosenthal and Zimmerman (1978). Observational learning through modeling occurs when observers display new behaviors that, prior to modeling, had no probability of occurrence, even when motivational inducements were offered (Schunk, 1995). Nowadays, cognitive psychologists transfer principles of modeling and 'learning-by-observation' from social psychology (Bandura, 1986) to the field of *cognitive modeling*, which incorporates explanation and demonstration by a model who verbalizes his thoughts and reasons for performing given actions (Meichenbaum, 1977). New trends in instructional science, such as cognitive apprenticeship, rely heavily on demonstration and observation as the key instructional and learning activities (Collins, Brown & Newman, 1989).

Various forms of 'learning-by-observation' have been implemented and studied. Sometimes errors are built into the modeled demonstration to show students how to recognize and cope with them (Schriver, 1991). Self-reinforcing statements (e.g., 'I'm doing well') seem to be useful, especially for students with learning difficulties and for students who have doubts about their capabilities to perform well (Schunk & Hanson, 1985).

Observational learning, with teachers, adults or students as models has shown to be effective for different school subjects, and with different ages of students. For example, Schunk and Hanson (1985, 1989a, 1989b); Schunk, Hanson and Cox (1987) (all mathematics); Graham and Harris (1994); Graham, Harris and Troia (1998) (all writing); Couzijn (1995, 1999); Couzijn and Rijlaarsdam (1996) (reading and writing). See also for an overview Schunk and Zimmerman (1997) and Schunk (1987, 1991, 1995, 1998).

Finding that observational learning is effective is one thing, explaining why is another thing. At least two factors hinder a detailed explanation of the effectiveness of observational learning: a) molar causal laws, and b) absence of process measures. Below, we will elaborate on these two factors.

In experimental studies, the observed effect can often be the result of multiple causes (i.e., molar causal laws) (Cook & Campbell, 1979, p. 33). As observation is a complex learning activity, the same can also be true for studies on the effects of observation as an effective learning activity. An observation task can be simple (e.g., just an observation with implicit induction of rules) or complex (involving explicit evaluation, elaboration, explicit rule formation). Even a simple observation task consists of several cognitive activities, such as perception, comparison, evaluation, abstraction, etc.. This problem also affects the interpretation of Couzijn's study (1995), in which a moderately complex observation task was successfully implemented: almost 95% of all observation tasks were completed by the students. The question now is which specific elements of the observation task contributed to the learning gains Couzijn found.

Another characteristic of most experimental studies is the absence of a description of what subjects are factually doing – that is, process measurements. For this reason, Lysynchuk, Pressley, d'Ailly, Smith and Cake (1989) stress the importance of the measurement of the processes occurring during instructional studies: 'Conclusions about instructional effects can be made with greater confidence when the kind of processing that occurs in each condition is documented' (Lysynchuk et al., 1989, p. 465). One possibility to measure the actual behavior of subjects is online data gathering (e.g., think-aloud protocols, online observation of learning activities and time on task). But even in cases where online measures were not implemented, post hoc analysis of students' workbooks could provide information about the learning process.

In the present study, we performed such a post hoc analysis of Couzijn's (1995) data in order to assess the relationship between process indicators and aptitude scores (pretest) and output scores (posttests). We used workbooks from the subjects involved in Couzijn's study. In these workbooks, we focused on observation tasks, aimed at stimulating monitoring activities, specially evaluation. With this analysis we try to detect effective elements of 'learning-by-observation'. This insight might then help us refine the theoretical background of observation as a learning tool, and guide us in the development of effective observation tasks.

2 METHOD

In this study we used materials (workbooks, pretest and posttests) from subjects in the observational learning conditions in Couzijn's study (1995). Participation in the experiment took place in two four-hour sessions over two days. On the first day, subjects made a pretest and followed the first part of an experimental course on argumentative reading and writing. On the second day, the course continued with the last part and posttests were administered.

2.1 Subjects

Subjects were 84, ninth-grade students (intermediate and high levels) randomly assigned to the observational learning conditions of Couzijn's study (1995). The average age was 15.5 years. 60% of the participants were female, there was an almost even distribution of boys and girls over the conditions. The subjects received a small financial reward for their participation.

2.2 Materials

We analyzed the workbooks of the subjects. These workbooks are part from a course on argumentative reading and writing Couzijn developed. This course consisted of four lessons. Several theoretical aspects were introduced sequentially: the dialectic nature of argumentation, argumentative text structures, hierarchical structures in written texts, and means for the presentation of text elements (Appendix A).

In the lessons, the subjects read theory and applied this theory in observation tasks. In these observation tasks, the subjects observed models who performed (while thinking aloud) reading or writing exercises, applying the theory that had just been presented. For instance, if a part of theory addressed the writing of an attractive ending, the subjects then observed models who were thinking about an attractive ending. After observing the models, the subjects received the product of the writing or reading task on paper. In terms of Figure 2.1 (see page 10) the observing process was as follows: because the models were thinking aloud while executing the writing or reading task, the subjects gained full insight into the orientation, execution and text revision activities (i.e., the writing/reading process) and into the task solution (i.e., the writing/reading product) of the models. Besides, some models also gave the subjects insight into the monitoring activities of these models. Some models showed good examples of these activities, monitoring, evaluating and reflecting their performance of the task. The subjects themselves were engaged in these monitoring activities because they had to identify ('how does the model apply the theory in the writing or reading task?'), to evaluate ('is this application going well?') and to elaborate ('why is this application going well?'). By monitoring the execution of the models, the subjects generated input for their learning task: i.e., learning to write and read argumentative texts.

In each observation task, the subjects observed either one model or two, and, after this observation, were required to evaluate the performance of the model/models, and to elaborate on that evaluation. This was done by means of questions: 'Which model performed better?' or 'Did the model perform well?' ('evaluation') and 'Explain why you think the less good model performed less well.' or 'Explain why the model performed well/did not perform well.' ('elaboration'). The subjects were free to choose the focus of elaboration: they could comment on the (reading/writing) processes, and/or on the products (see Table 2.1 for some examples). The subjects' answers to the evaluation and elaboration questions were interpreted as an indication that the subjects had performed a mental evaluation or a mental elaboration activity.

2.3 Procedure

2.3.1 Material selection

For our analysis, we selected materials from the second half of the course Couzijn developed. With respect to content, this was an obvious choice, as the course is built up cumulatively: in each lesson, the content of the previous lesson is incorporated. The first half of the course, which is an introduction to the second half, explains to the subject the meaning of the concepts of standpoint and arguments, and describes the features of argumentative texts. The second part of the course consists of two lessons (lesson A and lesson B), in which the subject learns how to read and write an argumentative text on the basis of a given structure for argumentation (lesson A) and how to analyze and add means for presentation of text elements (lesson B). By selecting the second half of the course, we obtained representative information on the theoretical content explained in all lessons. In the selected lessons, we coded the subjects' evaluation and elaboration activities in all observation tasks. In total we coded eight observation tasks (six tasks in lesson A and two tasks in lesson B).

2.3.2 Variables

We distinguished three different learning activities in observation tasks: 'evaluation', 'product-elaboration' and 'process-elaboration' (Table 2.1).

Table 2.1. Learning activities

Learning activity	Description	Example
Evaluation	Evaluation of performance of the model/models.	'The performance of model 1 was better.'
Product-elaboration	Focus on product when elaborating on the evaluation.	'Model 2 has connected the standpoint and arguments with "because", and that is wrong when you start with the arguments.'
Process-elaboration	Focus on writing or reading process when elaborating on the evaluation.	'Model 1 checked if he had included all arguments in his text.'

Each learning activity was measured in lesson A and in lesson B. We coded the learning activities in both lessons in a dichotomous way: 'correct' and 'incorrect' for 'evaluation' and 'product-elaboration', and 'presented' and 'not-presented' for 'process-elaboration'. Mean correlation among learning activities was .05 within lesson A, and -.01 within lesson B, so all learning activities could be distinguished empirically. The learning activity 'process-elaboration' was scarcely performed in both lesson A and lesson B (Appendix B). Although not explicitly invited to do so, subjects elaborated more on the products of the observed models than on the observed processes (for possible explanations, see Braaksma and Rijlaarsdam, 1997). The learning activities 'evaluation' and 'product elaboration' were performed at a very high level in both lessons (Appendix B); this is not surprising, as observation tasks are part of a *learning* program, and the aim is mastery, not selection. However, this can lead to ceiling effects and makes analysis cumbersome, as the relations between variables are influenced by an artificial restriction of range.

Table 2.2. Measurement of reading and writing skills

Skills	Sub skills	Description	Pretest	Posttest
Reading	Argumentation structure	Identification of standpoint and (subordinate) arguments.	x	x
Writing	Argumentation structure	Writing argumentative texts based on argumentation structures.		x
	Text structure	Organizing argumentative texts using a standard structure: introduction, body and ending.		x
	Presentation	Applying verbal means to enhance the presentation of an argumentative text.		x

Furthermore, we selected from Couzijn's data four key variables, which are to measure both the reading and writing skills taught in the selected lessons (Table 2.2). Mean, standard deviation, and reliability of these variables are presented in Appendix B. Almost all the dependent scores are moderately reliable (Appendix B); although the reliability of 'writing: presentation' and 'writing: text structure' is low, it is nonetheless sufficient for research purposes.

2.4 Analysis

In order to test whether a learning activity contributed significantly to the explanation of the learning gains, the data were subjected to LISREL (Bollen, 1989). For each learning activity, we expected to reject the 'no-effect-model', in which no relationship is assumed between a learning activity and pretest and posttests, in favor of the 'lesson-effect-model' in which a relationship is assumed between a learning activity in lesson A and the same learning activity in lesson B. Furthermore, we expected to reject the 'lesson-effect-model' in favor of the 'effect-model', in which relations are assumed between a learning activity and pretest and posttests.

2.4.1 Model A: 'no-effect-model'

Figure 2.2 shows an example of the 'no-effect-model'. For reasons of clarity, we indicated only one posttest variable in this figure.

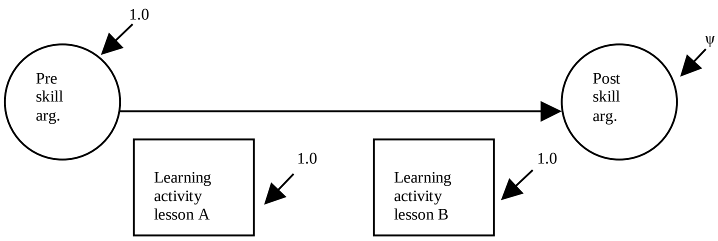


Figure 2.2. 'No-effect-model'
Note. 'Pre skill arg.' indicates a pretest variable for skills with regard to argumentation structure; 'Post skill arg.' indicates a posttest variable for the same skill.

In this model it is assumed that the performance on a posttest variable is only influenced by the performance on the similar pretest variable. Writing ability and reading ability are related within skills, so pretest 'argumentation structure: reading' is related to the posttests 'argumentation structure: reading' and 'argumentation structure: writing'. The pretest has no effect on a particular learning activity, and neither does a particular learning activity have an effect on posttests. This means both that the learning activities performed by subjects are independent of the scores those subjects are awarded for the pretest, and that the scores for the posttests are independent of the scores for the learning activity. It does not make any difference for the performance of the learning activity if the subject is good or weak in a particular skill, and neither does it not make any difference for the performance of the posttests if the subject has performed a particular learning activity well or badly.

2.4.2 Model B: 'lesson-effect-model'

In this model, effects from a particular learning activity performed in lesson A are added to the same learning activity in lesson B (Figure 2.3). Performance of the learning activity in lesson B depends on performance of the same learning activity in lesson A; in other words, if a subject often performs a particular learning activity correctly in lesson A, he will also perform the same learning activity correctly in lesson B. This model reflects the similarity between learning activities in lesson A and learning activities in lesson B. Although the learning *content* differs between the two lessons, there is consistency at the level of learning *activities*.

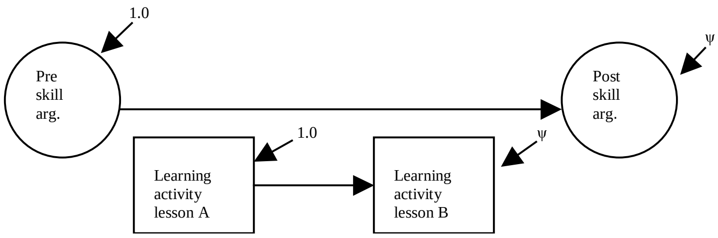


Figure 2.3. 'Lesson-effect-model'

2.4.3 Model C: 'effect-model'

In addition to the effects in the previous models, this model also allows for effects from the pretest on a particular learning activity and/or effects from a particular learning activity on posttests. This means that the performance of a learning activity is dependent on pretest performance. For instance, subjects who have a high score for the pretest will perform a particular learning activity more often, or better than subjects who have a lower score for the pretest. This effect indicates that the learning activities are not totally strange to the subjects, but that subjects are already performing these activities on their own, or that they are able to learn them easily.
Furthermore, in this model, the posttests are also partly dependent on the scores for the learning activities. This indicates a learning effect. For instance, subjects who correctly performed a learning activity on a greater number of occasions will have better results in the posttests than subjects who performed the learning activity less well. There are only paths from learning activities measured in lesson B to the posttests, in accordance with the cumulative structure of the course (Figure 2.4).

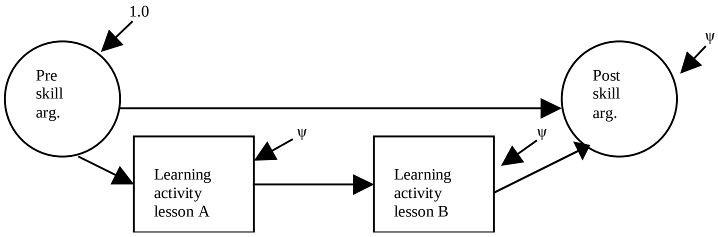


Figure 2.4. 'Effect-model'

3 RESULTS

In this section, for each learning activity, we present a comparison between the three statistical models. Furthermore, we discuss which model best fits each learning activity, and we present the significant relations within that best-fitting model.

For each learning activity, we compared the fit of model A with the fit of model B, and the fit of model B with the fit of model C. If the difference in χ^2 is significant, the model with the fewest degrees of freedom will be chosen because it describes the data significantly better. If the difference in χ^2 is *not* significant, the model with the largest degrees of freedom will be chosen (Jöreskog & Sörbom, 1993).

Table 2.3 shows a different best fit for the three learning activities. For the 'evaluation' and 'product-elaboration' learning activities, model C ('effect-model') best fits the data, while model B ('lesson-effect-model') best fits the data for 'process-elaboration'.

Table 2.3. Comparison-of-fit between three statistical models for learning activities in observation tasks

Learning activity	Model	Fit			Comparison				Preferred model
		χ^2	*df*	*p*		$\Delta\chi^2$	Δdf	*p*	
Evaluation	A	36.65	18	.00					
	B	30.65	16	.02	A-B	6.00	2	.05	
	C	12.73	8	.12	B-C	17.92	8	.02	Model C
Product-elaboration	A	47.37	18	.00					
	B	34.08	16	.01	A-B	13.29	2	.00	
	C	6.75	8	.56	B-C	27.33	8	.00	Model C
Process-elaboration	A	32.15	18	.02					
	B	24.10	16	.09	A-B	8.05	2	.02	
	C	16.36	8	.04	B-C	7.74	8	.46	Model B

Table 2.4 shows the different types of effects which occur in the different models. The regression weights reported are those under the best-fitting models (Table 2.3). The effects will be discussed in the following sections.

Table 2.4. Overview of regression weights and percentage explained variance (between brackets). For 'evaluation' and 'product-elaboration' these regression weights and explained variance are estimated under model C and for 'process-elaboration' estimated under model B

Type of effect			
I	*Pre skill*	*on*	*Post skill*
	Argumentation structure: reading		Argumentation structure: reading .38 (14.4%)
			Argumentation structure: writing .54 (29.1%)
II	*Learning activity lesson A*	*on*	*Learning activity lesson B*
	Evaluation		.26 (5.9%)
	Product-elaboration		.38 (11.1%)
	Process-elaboration		.30 (9%)
IIIA	*Pre skill*	*on*	*Learning activity lesson A*
	Argumentation structure: reading		Evaluation .35 (12.2 %)
			Product-elaboration .48 (23.1%)
IIIB	*Learning activity lesson B*	*on*	*Post skill*
			Presentation: writing
	Evaluation		.21 (4.1%)
	Product-elaboration		.21 (3.7%)

3.1 Effect I: pre skill on post skill

According to Table 2.4, the skills reflected in the pretest 'argumentation structure: reading' have positive effects on the posttests 'argumentation structure: reading' and 'argumentation structure: writing'. The pretest explains 14.4% of the variance in the posttest scores for 'argumentation structure: reading', and 29.1% of the differences of 'argumentation structure: writing'. Subjects who made correct identifications of standpoint and argument in the pretest were also good at the identification of standpoint, arguments and subordinate arguments when they were reading and writing argumentative texts in the posttests.

3.2 Effect II: learning activity lesson A on learning activity lesson B

Table 2.4 also shows that, for all learning activities, learning activity in lesson A has an effect on the same learning activity in lesson B. These are positive effects: subjects who perform the learning activity more often (correctly) in lesson A, will also perform the same learning activity more often (correctly) in lesson B. This means that the learning activities are stable activities. For instance, if a subject often comments on the process in lesson A, he will also be likely to comment on it often in lesson B. Furthermore, the effects tell us something about the measurement of the learning activities: the measurement is consistent, a ‘product-elaboration’ in lesson A is the same as a ‘product-elaboration’ in lesson B.

3.3 Effect IIIA: pre skill on learning activity lesson A

Table 2.4 shows positive effects from the pre skill ‘argumentation structure: reading’ on the learning activities ‘evaluation’ and ‘product-elaboration’. However, it should be noted that this effect does not apply for ‘process-elaboration’, for which model B (which does not contain paths from the pre skill to learning activities) best fits the data.

The table shows that 12.2% of the variance in ‘evaluation’ in lesson A is explained by pretest scores. Subjects who identified the standpoint and arguments well in the ‘argumentation structure: reading’ pretest, were better evaluators during lesson A. It seems that they had already developed good criteria for the identification of standpoint and arguments which made it easier for them to correctly evaluate the execution of the models.

Furthermore, 23.1% of the variance in the ‘product-elaboration’ learning activity is explained by pretest scores. Subjects who identified the standpoint and arguments well in the pretest were better at commenting on the products of the models during lesson A. They were better at the identification of standpoints and arguments, and could use this knowledge when commenting on the products of the models. For them, it is probably easier to explain what is less satisfactory about the products (of one) of the models.

These effects of pretest on learning activities mean that subjects who have a certain level of competence, either already manifest the intended behavior, or learn this type of activity more easily. So, these types of activities are activities that good subjects seem to develop (partly) on their own; these activities are not totally strange to the subjects.

3.4 Effect IIIB: learning activity lesson B on post skill

Like effect IIIA, effect IIIB applies only to ‘evaluation’ and ‘product-elaboration’ activities, it should be borne in mind that model B best fits the data for ‘process-elaboration’, and that this model does not contain paths from learning activities to post skills.

In Table 2.4 positive effects can be seen from the ‘evaluation’ and ‘product-elaboration’ activities to the post skill ‘presentation: writing’. ‘Evaluation’ explains 4.1% of the variance on the ‘presentation: writing’ posttest. Subjects who more often made correct evaluations of the performance of the models turned out to be better in the use of presentation elements when writing argumentative texts. Evaluation seemed to be an effective activity during learning, and subjects who are able to make correct evaluations of observed writing or reading activities may develop their monitoring skills. There seems to be a relationship between evaluating someone else’s actions and executing and evaluating your own. If you are able to make correct evaluations, you probably make correct evaluations of your own performance, thereby bringing about a higher quality of writing.

The table also shows a positive effect for ‘product-elaboration’: 3.7% of the variance on the ‘presentation: writing’ posttest is explained by ‘product-elaboration’. Subjects who elaborated correctly on the products of the models were better in the use of presentation elements when writing argumentative texts. ‘Product-elaboration’ seems to be another effective activity when learning to write argumentative texts. Subjects who make comments on the observed writing or reading products have noticed what is good and less good in the products of the models. This knowledge probably transfers to the subjects' own performance, whereby they use the presentation elements better in their own texts.

4 DISCUSSION

This post hoc study was undertaken in order to identify which elements of observation tasks are effective when students are learning to read and write argumentative texts. We focused on two elements of the observation task: evaluation and elaboration on this evaluation. Three different learning activities were distinguished: ‘evaluation’, ‘product-elaboration’ and ‘process-elaboration’. For each learning activity, we tested which model best fitted the data. With respect to all learning activities, we expected that we would be able to reject model A (the ‘no-effect-model’) in favor of model B (the ‘lesson-effect-model’), and that we would in turn be able to reject model B in favor of model C (the ‘effect-model’).

As well as the effects of pretest on posttests, all learning activities undertaken in lesson A were found to have an effect on the same learning activity in lesson B. As expected, we could reject model A in favor of model B (for ‘process-elaboration’) and model C (for ‘evaluation’ and ‘product-elaboration’). In both models, a relationship is assumed between the learning activity in lesson A and the same learning activity in lesson B. Subjects who performed these learning activities better (‘evaluation’ and ‘product-elaboration’) or more often (‘process-elaboration’) in one lesson also performed these activities better or more often in the other lesson. This effect means that, independent of their actual educational content, the learning activities are stable measured. This characteristic facilitated the interpretation of the effects of learning activities on the posttests.

For the ‘evaluation’ and ‘product-elaboration’ activities, aptitude proved to be related to the learning activities. Subjects who were better in the identification of standpoint and arguments in the pretest performed each of these learning activities better. This means that good subjects either show early evidence of the intended behavior, or learn this type of activity more easily. These activities are therefore activities that good subjects seem to develop (partly) on their own.

We also observed positive effects of ‘evaluation’ and ‘product-elaboration’ activities on a posttest. Subjects who correctly evaluated the peer models turned out to have learnt the use of presentation elements when writing argumentative texts. The same was true for ‘product-elaboration’: subjects who more often commented correctly on the argumentative texts and structures the models produced were better in the use of presentation elements when writing argumentative texts. These positive effects from ‘evaluation’ and ‘product-elaboration’ to argumentative writing, stress the importance of the acquisition of criteria (Hillocks, 1986). It is assumed that subjects who evaluated and elaborated correctly on the performance of the models developed criteria for effective texts and effective writing processes. They acquired a concept of a ‘good’ text that might guided their own process of writing.

These effects of pretest on ‘evaluation’ and ‘product-elaboration’, and of ‘evaluation’ and ‘product-elaboration’ on a posttest, confirm our expectations. As expected, model C, in which relations are assumed between learning activities and pretest and posttests, best fitted the data.

With regard to the third learning activity, ‘process-elaboration’, no effects of pretest on ‘process-elaboration’ nor from ‘process-elaboration’ on posttests were observed (model B best fitted the data and in that model no relations are assumed between learning activities and pretest and posttests). We expected a positive effect from ‘process-elaboration’ on posttests because we assumed observation of *how* the models perform reading and writing activities to be an added value, and expected that subjects who explicitly commented on the reading or writing processes of the observed models would have better scores on the posttests.

A possible explanation for the lack of ‘process-elaboration’ effects is related to the low frequency (Appendix B) of students who refer to processes when they elaborate on their evaluations. This bottom effect results in a restriction of range, which makes it more difficult to observe statistically significant effects, if any.

Furthermore, the measurement of ‘process-elaboration’ was different from that of ‘evaluation’ and ‘product-elaboration’. Rather than coding whether or not subjects' ‘process-elaboration’ was correct, we coded if the subjects explicitly showed hat they had focused on the observed processes. In contrast with the coding of ‘evaluation’ and ‘product-elaboration’, it is difficult to code whether a subject has correctly elaborated on the process. Thus, with no clear distinction between correct and incorrect, the measurement of ‘process-elaboration’ is relatively weak.

Another type of explanation concerns the distinction between processes and performances. It is possible that there actually was a change in writing and reading processes caused by the learning activity ‘process-elaboration’, but that this change did not materialize in higher posttest performance. In our future studies on effective learning activities in observation tasks, we plan to include process measurements of reading and writing along with performance measurements as dependent variables.

It should be noted that the regression weights and explained variance in this study are not very high. At the same time, however, it should also be borne in mind that the observations we report are underestimates of ‘true’ relations between learning activities and learning gains. First, the indicators of learning activities are indirect: we did not observe learners at work, but coded their learning materials, i.e., the products they had left for us. Second, the nature of implementation scores restricts the options for finding relations between learning and learning gains. In a learning program, the aim is for learners to solve most of the problems, with some support being provided by the materials: the aim is mastery, not selection. But despite these two shortcomings, the methodology of a post hoc study proved to be effective in these circumstances.

This study described the features of observation that are important to learning. The results confirm the importance of the monitoring, evaluative and reflective activities defined in the introduction. The effects of 'evaluation' and 'product-elaboration' on the use of presentation elements when writing argumentative texts show that when subjects evaluate the reading or writing activities of models, and explain why the writing or reading tasks are performed well or badly – it transfers to their own writing. Apparently, by observing the reading or writing processes and products, the subjects thus generate a large input for learning. It seems that subjects who have learned by observation are able to handle the double agenda of task-execution and learning, and can learn complex skills more easily.

APPENDIX A

Theoretical contents of the course on reading and writing argumentative texts (adapted from Couzijn and Rijlaarsdam,1996, p. 259)

Lesson 1: 'Argumentative texts and discussions'

Introduction of five main concepts:

- standpoint (opinion)
- argument (reason)
- argumentative text
- issue
- discussion

Lesson 2: 'The structure of argumentative texts'

Presentation of a rhetorical model, consisting of:

- introduction: request for attention; issue at stake; parties and standpoints
- body: author's standpoint; pro-argumentation; refutation or counter-argumentation
- ending: conclusion; most important arguments; consequence

Lesson 3: 'The argumentation in argumentative texts'

Presentation and discussion of several types of argumentation:

- singular argumentation
- compound argumentation
- subordinate argumentation

and the complex argumentation structures of which these are the constituents.

Lesson 4: 'The presentation of argumentative texts'

Presentation and discussion of three means for clarifying the text structure:

- paragraphing and the rhetorical model
- using verbal structure markers
- argumentative connectors

APPENDIX B

N, mean, standard deviation, and reliability (Cronbach's α) of the variables

Variables	*N*	*M*	*SD*	α
Reading ability: argumentation structure pretest	84	17.15	4.41	.85
Reading ability: argumentation structure posttest	84	27.05	7.00	.88
Writing ability: argumentation structure posttest	84	20.38	6.35	.89
Writing ability: text structure posttest	84	8.30	5.36	.63
Writing ability: presentation posttest	84	10.24	3.24	.65
Evaluation lesson A	84	.84	.17	
Evaluation lesson B	84	.87	.25	
Product-elaboration lesson A	84	.80	.21	
Product-elaboration lesson B	84	.91	.25	
Process-elaboration lesson A	84	.14	.19	
Process-elaboration lesson B	84	.18	.24	

Note. The coding for the learning activities in both lessons is as a follows:
'Evaluation' and 'Product-elaboration'

- *minimal score: 0 (learning activity is not correctly performed),*
- *maximal score: 1 (learning activity is correctly performed).*

'Process-elaboration'

- *minimal score: 0 ('process-elaboration' is not performed),*
- *maximal score: 1 ('process-elaboration' is performed).*

CHAPTER 3

THE EFFECTS OF MODEL-OBSERVER SIMILARITY

An experimental study

Abstract

This study examined the effects of similarity in competence between model and observer on the effectiveness of observational learning in argumentative writing. Participants (*N*= 214, eighth grade, mixed ability) were assigned to one of three conditions: an observation/weak focus, an observation/good focus, or a control condition. The two observational-learning groups observed pairs of peer models performing writing tasks. Participants focused respectively on the noncompetent (weak) model or on the competent (good) model. The control group performed the writing tasks themselves. Results were consistent with the 'similarity hypothesis': weak learners learned more from focusing their observations on weak models, while better learners learned more from focusing on good models.

1 INTRODUCTION

Observing is the key learning activity in learning environments in which learners learn from models. It occurs when observers display new behaviors that prior to modeling are not demonstrated even with motivational inducements to do so (Schunk, 1998). There is a large body of research about effects and conditions of modeling (Bandura, 1986, 1997; Rosenthal & Zimmerman, 1978; Schunk, 1987, 1991, 1995, 1998; Schunk & Zimmerman, 1997). With either teachers, adults or students as models, observational learning has proven to be effective with students of various ages and in various school subjects, such as mathematics (Schunk & Hanson, 1985, 1989a, 1989b; Schunk, Hanson & Cox, 1987), writing (Graham & Harris, 1994; Graham, Harris & Troia, 1998; Schriver, 1992), speaking and listening (Sonnenschein & Whitehurst, 1983, 1984), and reading and writing (Couzijn, 1995, 1999). Observation of models also can raise observers' self-efficacy, or personal beliefs about their capabilities to learn or perform behaviors at designated levels (Bandura, 1986, 1997). Schunk and his colleagues (Schunk & Hanson, 1985, 1989a; Schunk et al., 1987; see also Schunk, 1998, p. 148) reported effects of (various) models on students' self-efficacy which in turn influenced learning and achievement.

Braaksma, M. A. H., Rijlaarsdam, G., & Van den Bergh, H. (in press). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology.*

The effectiveness of observational learning depends on a number of factors (Schunk, 1991, p. 113), including perceived similarity in competence between model and observer. Schunk (1987, 1991, 1998) reviewed studies which investigated how variations in model competence affect children's behavior. Some studies involved comparisons of *mastery and coping models*. Mastery models perform the task correctly. They may also verbalize statements reflecting positive attitudes, and high levels of confidence and ability. Mastery models demonstrate rapid learning and make no errors. In contrast, coping models show their hesitations and errors, but gradually improve their performance and gain self-confidence. These models illustrate how determined effort and positive self-reflections may overcome difficulties.

In three studies, Schunk and his colleagues compared the observation of coping models with that of mastery models. In two of these studies (Schunk & Hanson, 1985; Schunk et al., 1987), the participants were children with learning difficulties, and they were supposed to be similar to the coping models. Schunk and Hanson (1985) reported no differences between the coping and mastery conditions on any measure (perceived similarity in competence to the model, self-efficacy for learning, posttest self-efficacy, and posttest skill). Schunk and Hanson (1985, p. 320) supposed that the lack of differences between the coping and mastery conditions might be induced by the participants' prior successes with the task. In a replication study, Schunk et al. (1987) used a task with which participants had no prior success; they found that observing a coping model led to greater perceived similarity in competence to the model, higher self-efficacy for learning, and higher posttest self-efficacy and skill. In a third study, involving predominantly average performers, Schunk and Hanson (1989a) compared mastery models, coping-emotive models (i.e., coping models, as in the earlier studies), and coping-alone models (i.e., identical to coping-emotive models, but without negative statements about their own ability). Observing coping-emotive models led to the highest self-efficacy for learning, but no differences were found in participants' performances on the posttest. According to Schunk and Hanson (1989a, p. 433) this is due to the fact that –because of their prior experiences– participants judged themselves to be more competent than the observed coping-emotive models. Schunk and Hanson assumed that by judging themselves to be more competent, coping-emotive participants overestimated their learning efficacy. This boost in self-efficacy resulting from comparison with the model causes efficacy to lose some of its predictive utility (Schunk & Hanson, 1989a, p. 433).

Other studies compared models *high and low in competence*. Model competence was inferred from the outcomes of modeled actions (success, failure), but also from symbols that denote competence (e.g., prestige) (Schunk, 1991, p. 114). In studies of observational learning of cognitive skills (see Schunk, 1987, p. 161-162) the observation of competent models led to higher performance than the observation of non-competent models. In other words, models dissimilar in competence to observers exert more powerful effects on learners' behaviors.

Graham and his colleagues (Graham & Harris, 1994; Graham, Harris, MacArthur & Schwartz, 1998; Graham, Harris & Troia, 1998) also argue in favor of competent models. They outlined an instructional model for developing writing and self-regulation strategies among students with learning and writing problems. Part of this Self-Regulated Strategy Development (SRSD) model is a teacher who models good writing strategies (i.e., a competent model). SRSD improved students' writing performance.

Similarity in competence may be important in contexts where children cannot readily discern the functional behavior (i.e., when they lack task familiarity, when there is no objective standard of performance, or when modeled actions are followed by neutral consequences; Schunk, 1987, p. 161). Furthermore, in studies in which model competence served as a source of information for purposes of self-evaluation, children were influenced by models of similar (same-age) competence (see Schunk, 1987, p. 162).

To determine whether similarity in competence between model and observer is essential, it is important to distinguish different types of learning purposes. In situations involving the learning of skills or new behaviors, children tend to emulate competent peers. Modeled competence is necessary for students to learn correctly. Even coping models eventually demonstrate competence. But where social comparisons are employed for self-evaluations, children are most interested in the performances of others who are similar in ability (Schunk, 1987, p. 162). Furthermore, all models, but especially models whom children view as similar in competence, may promote children's self-efficacy for learning observed skills (Schunk et al., 1987, p. 54).

1.1 The 'similarity hypothesis' in observational learning to write

In this study we focused on writing argumentative texts, which was a new task for the eighth grade participants. We assume that the process of observation differs between weak and good students. When weaker students observe and evaluate a model, they rely to a lesser extent on knowledge about 'good writing'. In contrast, good students, who already have access to an – albeit imperfect – knowledge base and a set of criteria (Hillocks, 1986), can apply this knowledge to the evaluation of the model.

When weak writers observe a competent *and* a noncompetent model, it is much easier for them to evaluate the weak model: they can use the performance of the better model as a reference. Besides, the performance of the observed weak model is probably more matched with the cognitive processes of a weak learner. The same can be argued for the observation of good models for good writers: the performance of a good model is probably more matched with the cognitive processes of a good learner. Furthermore, when good writers observe pairs of models, they are able to identify and comment on the qualities of the good model, using their own set of evaluation criteria.

This brings us to the 'similarity hypothesis' that is tested in the present study. This hypothesis states that weak writers learn more from focusing their observations on weak models and that good writers learn more from focusing on good models. In the study, participants were assigned to one of two observational-learning conditions or to a control condition. The two observational-learning groups observed the same pairs of models. The participants focused respectively on the noncompetent (weak) model (condition observation/weak focus) or on the competent (good) model (condition observation/good focus). Participants in the control group (condition direct writing) performed the writing tasks themselves. The addition of this condition enabled us to compare the relative strengths and weaknesses of two types of observational learning with a more traditional type of learning.

2 METHOD

2.1 Participants

The experiment involved 214 fourteen-year-old participants from nine eighth grade classes in a Dutch city school. The achievement level of the participants ranged from low to high. Classes were composed heterogeneously in terms of achievement level and gender. The experiment was part of the regular lesson roster, therefore classes were randomly assigned to conditions. There were no differences in gender and age between conditions. In the condition direct writing, 54% of the participants was female, the mean age was 14.01 years. In condition observation/weak focus, 55% of the participants was female. Mean age was 14.12 years. Finally, in condition observation/good focus, 50% of the participants was female, the mean age was 14.15 years.

Table 3.1. Research design

Session	Content	Direct writing (*n*=76)	Observation/ weak focus (*n*=74)	Observation/ good focus (*n*=64)
1	Pretest 1 (verbal intelligence)	x	x	x
2	Pretest 2 (identification argumentation)	x	x	x
3	Instruction session 1	Writing texts	Observing two models on CD-ROM who are writing texts, focus on the *weak* model	Observing two models on CD-ROM who are writing texts, focus on the *good* model
4	Posttest 1 (writing)	x	x	x
5	Instruction session 2	Writing texts	Observing two ‘live’ models who are writing texts, focus on the *weak* model	Observing two ‘live’ models who are writing texts, focus on the *good* model
6	Posttest 2 (writing)	x	x	x

2.2 Design

The general design of the present study was a pretest-posttest control group design (see Table 3.1). We implemented two experimental instruction sessions. In the first instruction session we used a CD-ROM to show the models and in the second instruction session 'live modeling'. After each of these sessions the effect on performance was measured.

By implementing two instruction sessions we can test the 'similarity hypothesis' in two circumstances: when a task is completely new (instruction session 1) and when a task had become already a bit regular (instruction session 2). Furthermore, implementing two variants of the instruction might enhance the validity of the independent variable, as it allows generalizing effects over different specimens of the independent variable.

Participation in the experiment took place in six sessions, each lasting approximately 30 to 60 minutes, conducted by the researcher and research assistants. Pretests were administered during the first two sessions. The third session was the first instruction session. Posttests were administered during the fourth session. The second instruction session was implemented in session five; some posttests were administered in session six (see Table 3.1). The sessions were distributed equally over a five-week period.

2.3 Materials and procedure

For this study, we developed a short course on an aspect of writing argumentative texts, consisting of two instructional sessions. Students learned to transform argumentation structures into short linear argumentative texts. This learning task was completely new for all participants. The course was adapted with permission from Couzijn (1995, 1999). Students applied the theory presented in short writing tasks (i.e., the control condition direct writing) or in observation tasks. During the latter, participants observed pairs of models performing short writing tasks, which were the same as those the participants in the control condition performed. Participants in both observational-learning conditions observed the same pairs of models (five pairs in instruction session 1; two pairs in instruction session 2), in the same order, with the only difference that instructions in the observation/weak-focus condition required participants to reflect on the performance of the weak model, and those in the observation/good-focus condition to reflect on the performance of the good model.

The *first instruction session* lasted about 60 minutes. Materials for this session consisted of a workbook and a CD-ROM program that was projected onto a large screen in front of the class. Participants were informed of what to do by means of on-screen messages (e.g., reading theory in the workbook, answering a mastery question, writing a short text or observing two models). An on-screen timer indicated how much time was left for each activity; participants were alerted by short beeps when the time had almost elapsed. In pilot sessions, the required time for each activity, and the difficulty of theory, mastery questions, and exercises were examined. The required time was calculated in such a way that it was possible to read each theoretical section once, but participants were allowed to glance backward through the workbook, for instance to reread parts of the theoretical sections when answering a mastery question.

The participants' workbooks consisted of four main sequences (see Figure 3.1).

Writing argumentative texts

Sequence I: Theory and five mastery questions about seven main concepts (all conditions)

- standpoint (opinion) (for instance, the music should not be so loud)
- argument (reason) (for instance, through the loudness I cannot work)
- argumentative texts (text in which one gives his standpoint and argumentation for his standpoint)
- singular argumentation (standpoint with only one argument)
- compound argumentation (standpoint with two or more arguments)
- argumentation structures (structure of standpoint and arguments)
- argumentative connectives (words to connect standpoint with arguments, for instance, 'because')

Example mastery question: 'What is the difference between singular and compound argumentation?'

↓

Sequence II: Applying the theory in five short writing tasks (direct writing) or four observation tasks with the focus on *weak* models (observation/weak focus) or focus on *good* models (observation/good focus). For examples of writing and observation tasks see Conditions, page 34.

↓

Sequence III: Theory and four mastery questions about subordinate argumentation (argument which supports an argument), argumentation structures (structure of standpoint, arguments and subordinate arguments) and argumentative connectives (words to connect standpoint with arguments, and arguments with subordinate arguments, for instance, 'because') (all conditions).

Example mastery question: 'When is an argument a subordinate argument?'

↓

Sequence IV: Applying the theory in two short writing tasks (direct writing) or one observation task with the focus on the *weak* model (observation/weak focus) or focus on the *good* model (observation/good focus). For examples of writing and observation tasks see Conditions, page 34.

Figure 3.1. Sequences of instruction session 1 'Writing argumentative texts'

In the first sequence, participants had to study a theoretical section. To stimulate active reading of the theory, all participants answered five mastery questions. In the second sequence, participants applied the theory in different types of exercises according to the condition to which they had been assigned. This meant that participants in the direct-writing condition wrote short argumentative texts based on the argumentation structures they had been given. Meanwhile, participants in the observation/weak-focus condition used a CD-ROM to observe pairs of videotaped models writing short argumentative texts on the basis of these argumentation structures. In two questions, they were instructed to focus their attention on the performance of the weak model. Participants in the observation/good-focus condition observed the same CD-ROM, with the same pairs of models writing the texts. But in this case they were instructed to focus their attention on the good model. By using videotaped models on CD-ROM, standardized presentation across participants was ensured.

After completing the exercises, the participants went on to the third sequence, which contained theory and mastery questions about subordinate argumentation, argumentation structures and argumentative connectives. Finally, in the fourth sequence, they performed depending on the condition, two writing tasks or an observation task.

Providing equal learning time in all conditions, we had to add two extra writing tasks in the control condition, because observing two models writing a text and answering two questions required more time than to write a text. Therefore, participants in control condition direct writing performed in this instruction session seven writing tasks, and participants in both observational-learning conditions performed five observation tasks. In Appendix A, an overview of instructional time for each sequence is presented. The total learning time in control condition and both observational-learning conditions was almost equal.

The *second instruction session* lasted 30 minutes, starting with a summary of the theory presented in the previous session and some mastery questions. There were opportunities for participants to ask questions, and to discuss their answers with their peers. The discussions were about the correct answers on the mastery questions and why an alternative answer could be correct or not. For instance, one of the mastery questions was about the correct connective between standpoint and argument. One of the correct answers was the connective ‘because’. Some participants asked whether their answer ‘thus’ was correct. That answer was incorrect because with the connective ‘thus’, first the argument should be presented and then the standpoint.

The different groups then performed the following activities: the participants in the direct-writing condition wrote two short argumentative texts based on the argumentation structures they had been given. After writing each text, there was a short discussion between the participants about which texts were possible. For instance, participants discussed which connective should be used to connect the subordinate argument with the main argument. This is an important issue because if one uses a coordinating connective as ‘and’, the subordinate argument becomes coordinative instead of subordinate. This discussion was led by the researcher.

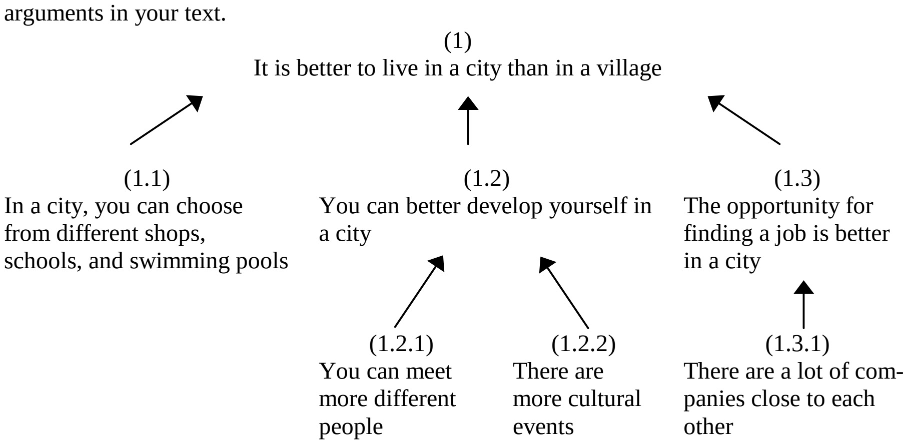
The participants in the observational-learning conditions performed two observation tasks. These tasks and sequencing were similar to the tasks in instruction session 1. However, instead of watching models on CD-ROM, the observers watched ‘live’ models (i.e., a researcher and a research assistant following a script). The use of a script here also ensured a standardized presentation across participants; in order to avoid a researcher-effect, the roles were randomly assigned. As in instruction session 1, participants in both conditions observed the same pairs of models and were instructed to focus on the weak model (observation/weak-focus condition) or on the good model (observation/good-focus condition). After each observation task, the researcher led a short discussion between the participants on the possible answers to the two questions. They discussed why one model was less good (observation/weak-focus condition) or better (observation/good-focus condition) than the other model. For instance, in one of the tasks, the weaker model wrote the subordinated argument as a coordinated argument.

The first part of instruction session 2 (summary and mastery questions) lasted 10 minutes and the second part (performing two writing tasks or two observation tasks) lasted 20 (ten minutes for each task).

2.4 Conditions

Participants in the condition direct writing applied the theoretical material in writing tasks. They wrote short argumentative texts based on argumentation structures (see Figure 3.2).

Below, you will find an argumentation structure. Write a short argumentative text based on this argumentation structure. Make sure that the reader will understand the standpoint and arguments in your text.



Write your text here:
....……………..
……………..……...
(…)

Figure 3.2. Writing task for participants in direct-writing condition

In a little while, you are going to watch on CD-ROM videotape recordings. You will see two models writing a short argumentative text based on the following argumentation structure. The models had to make sure that the reader will understand the standpoint and arguments in their text. Below, you will find the argumentation structure the models received.

(1)
It is better to live in a city than in a village

(1.1)
In a city, you can choose from different shops, schools, and swimming pools

(1.2)
You can better develop yourself in a city

(1.3)
The opportunity for finding a job is better in a city

(1.2.1)
You can meet more different people

(1.2.2)
There are more cultural events

(1.3.1)
There are a lot of companies close to each other

After watching the models, you have to answer the following questions:

1. Which model did less well?
2. Explain briefly what this (less good) model did worse.

When you have observed both models, you may advance to the next page.
Make your notes here, when you observe the models:
....……………..

(…)

(next page booklet)

You saw two models doing the assignment. They wrote the following texts:

Model 1
It is better to live in a city than in a village because in a city you can choose from different shops, schools, and swimming pools, you can better develop yourself in a city and the opportunity for finding a job is better in a city. Because you can meet more different people and there are more cultural events, you can better develop yourself. The opportunity for finding a job is better because there are a lot of companies close to each other. Therefore it is better to live in a city.

Model 2
It is better to live in a city than in a village because firstly in a city you can choose from different shops, schools, and swimming pools. Secondly you can better develop yourself in a city because you can meet more different people and there are more cultural events. Moreover, the opportunity for finding a job is better in a city, as there are a lot of companies close to each other.

1. Which model did less well? Model ....
2. Explain briefly what this (less good) model did worse.

Model ...... did worse, because
………………………………..…………..................

(…)

Figure 3.3. Observation task for participants in condition observation/weak focus

Participants in both *observational-learning* conditions observed the same pairs of models (a competent and a noncompetent model) executing the same writing tasks. Because the models were thinking aloud while executing the writing task, the participants gained full insight into the writing process and into the task solution (i.e., the writing product) of the models. Besides, some models also gave the participants insight into their monitoring activities, demonstrating monitoring, evaluating, and reflecting on performance. The verbalizations of the models' thoughts were directed to (planning and checking) the writing processes, not to achievement beliefs. Our weak models did not verbalize achievement beliefs that reflected low self-efficacy and negative attitudes and neither did our good models verbalize positive statements reflecting high self-efficacy.

The participants were stimulated to make notes when they were observing the models, as these might help them answering two questions after observing the two models. These two questions distinguished between the two observational-learning conditions.

Participants in condition *observation/weak focus* were instructed to focus on the weak model by answering the questions: (a) 'Which model did less well?' and, (b) 'Explain briefly what this (less good) model did worse.' When answering these questions, the participants were stimulated to evaluate the writing processes and the resulting products, and also to reflect on them. Figure 3.3 is an example of an observation task for participants in the observation/weak-focus condition, drawn from instruction session 1. A possible answer on the questions in the Figure 3.3 observation task would be, 'Model one did worse because he wrote first all arguments and then the arguments again with the subordinate arguments. It took too long!'.

Participants in condition *observation/good focus* observed the same pairs of models writing argumentative texts, but answered different questions: (a) 'Which model did well?' and, (b) 'Explain briefly what this (good) model did well.' One of the participants wrote in his answer on the same task as in Figure 3.3 but then for the observation/good-focus condition: 'Model two did well because she used "firstly", "secondly", and "moreover" and she connected the arguments directly with the subordinate arguments.'

The answers of the participants on the observation tasks were scored for accuracy of implementation. We found that in both conditions, participants focused on the intended model 5 of 7 times.

2.5 Testing materials

To indicate which participants were weak and which were good, we measured their aptitude. To ensure that aptitude would not be distorted by the effects of the instruction sessions, this measurement was carried out before the instruction sessions took place. Aptitude was measured by two intelligence tests: a Cognition of Meaningful Units Word List (DAT, 1984) and Verbal Analogies (DAT, 1984). We chose these particular tests because the analysis of argumentation (an important skill when learning to write argumentative texts) can be interpreted as an ability to discern abstract relationships between verbal units (Oostdam, 1991). The tests scores were summed, which resulted in a reasonably reliable measurement (Cronbach's alpha = .69). The mean[2] and also the standard deviation for each condition, plus an example of this measurement of aptitude, are in Table 3.2.

Table 3.2. Number of items, maximum score, means, and standard deviations for each condition, as well as some examples of aptitude and identification of argumentation

			M (*SD*)		
Measurement	Number of items	Maximum score	Direct writing	Observation/ weak focus	Observation/ good focus
Aptitude	22	22	5.36 (2.86)	5.94 (3.06)	6.13 (3.61)
Identification of argumentation	8	8	2.88 (2.08)	3.31 (1.8)	3.57 (1.9)

Example aptitude	*Make a complete and logical sentence:* … is to hear as blind is to… a) ear-light; b) deaf-see; c) deaf-dark; d) sound-eye; e) noise-watch.
Example identification of argumentation	*Underline the argument(s) in the following text:* Your brother was not allowed to go on his own on vacation when he was sixteen. Your sister was not allowed too. You are also not allowed to go on your own on vacation.

To control for a priori differences between conditions on the writing of argumentative texts, all participants performed a pretest that measured their skill to identify argumentation, a skill that is very important to the writing of argumentative texts. The reliability (Cronbach's alpha) of this pretest is .72. The mean and standard deviation for each condition, and an example of an item from this pretest are in Table 3.2. There were no significant initial differences in mean scores between conditions, neither on the measurement of aptitude ($\chi^2_3 = 1.49$; $p = .68$), nor on the pretest on the writing of argumentative texts ($\chi^2_3 = 5.01$; $p = .17$).

Writing performance of each participant after instruction session 1 (posttest 1) and instruction session 2 (posttest 2) was measured by writing tests, in which participants wrote five argumentative texts based on five argumentation structures. The participants had 40 minutes to write their texts in posttest 1 and 25 minutes to write their texts in posttest 2. There was sufficient time to complete the tests: 88% of the participants had no missing items on posttest 1; 82% had no missing items on posttest 2. An example of the type of the assignment can be found in Figure 3.2, with the

2 For all tests, a normalization transformation was applied to fulfill the assumptions required for the analyses.

type of test the same as the writing tasks undertaken by participants in the direct-writing condition.

The tests were coded to measure whether the participants correctly identified the standpoint, arguments, and subordinate arguments in the texts. This measurement was performed by examining whether the participant used appropriate connectives to link standpoints with arguments, and arguments with other arguments or subordinate arguments. For instance, if a standpoint is written firstly and a participant connected standpoint and argument with the connective 'thus', this is coded as incorrect. For the code 'correct', the participant had to use the connective 'because'. The aspect completeness was measured by coding missing items as incorrect. We coded strictly: the participant had to use explicit connectives. If they connected a standpoint with an argument by using a comma, this was coded as incorrect.

The tests given after the two instruction sessions were of similar type, the only difference being one of content. To avoid possible effects of content, a fully counterbalanced design was used: half of the participants received writing test A as posttest 1, and writing test B as posttest 2. The other half received writing test B as posttest 1 and writing test A as posttest 2. The items from text four and five (the more difficult texts) of both tests were used for the measurement of writing performance. The reliability (Cronbach's alpha) of posttest 1 was .68, and that of posttest 2 was .73.

2.6 Analyses

Suppose, Y_{ij} is the score of participant j (j = 1, 2, 3, …, j) at occasion i (i = 1, 2) in the direct-writing condition. Let T_{1ij} and T_{2ij} be dummy variables which are turned on (equal 1) at occasions 1 and 2 respectively, and otherwise turned off (equal 0). Further to this, APT_{0j} stands for the aptitude of individual j and $PERF_{1j}$ for previous performance of this individual. The model[3] to be analyzed (the 'aptitude-instruction-model') for this condition could thus be written as:

$$Y_{ij} = \quad T_{1ij}\,(ß_1 + ß_2{*}APT_{0j} + u_{10j}) \qquad (1)$$
$$T_{2ij}\,(ß_3 + ß_4{*}APT_{0j} + ß_5{*}PERF_{1j} + u_{20j}).$$

In this model, $ß_1$ and $ß_3$ indicate the intercept of first and second measurement occasion, respectively. The influence of aptitude on both measurement occasions can be evaluated by means of $ß_2$ and $ß_4$, these parameters are the regressions of both posttest scores on aptitude scores. The fifth regression weight ($ß_5$) was introduced to the model in order to take into account differences due to the first experimental instruction session. In a more traditional analysis, this parameter would be called a covari-

3 We did not specify classroom level in the model because the unit of analysis is the individual: the intervention was applied to individuals, not to classes. Hypotheses about differential effects on classes – all composed as mixed ability classes – were not plausible. Furthermore, 'only' nine classes were involved. The class level coefficients in the model would never reach significance.

ate. The last two parameters, u_{10j} and u_{20j}, are residual scores that index the differences between the mean score and the writing score of participant *j* in the first and second posttest, respectively. The residuals are assumed to be normally distributed with a variance of S^2u_1 and S^2u_2.

This model, which is given for one condition, can easily be extended to the case at hand, in which there are three conditions to be distinguished. In this case (3 conditions multiplied by 5), 15 regression weights need to be estimated.

The effect of individual variables can easily be tested because the ratio of the parameter's estimate and its standard error is student-t distributed. Differences between conditions can be tested by means of a contrast procedure (Goldstein, 1995).

In this model, all explanatory variables (i.e., aptitude and performance) are centered around the grand means. So, if differences between conditions in the intercept of the first measurement occasion are assessed, the writing score of participants with a mean aptitude score differs between conditions. For the intercept of the second measurement occasion, the exact interpretation is the mean of participants' writing scores with a mean aptitude score as well as a mean performance score.

Differences between conditions in effects of aptitude can be assessed on both measurement occasions. This allows for a direct testing of our 'similarity hypothesis'. If participants benefit from a similarity with the models, a relatively more positive regression weight is expected in the condition observation/good focus than in the condition observation/weak focus.

3 RESULTS

Table 3.3 shows all estimates for the parameters in our model. The first two rows concern posttest 1, and report the intercepts for posttest 1 and the regression weights of posttest 1 on aptitude. The other rows concern posttest 2. Next to the intercepts for posttest 2, two regressions are reported: the regression weights of posttest 2 on aptitude, and the regression weights of posttest 2 on posttest 1. In Appendix B the means and standard deviations for both writing measurements in each condition are reported.

Table 3.3. Estimates for all parameters in the 'aptitude-instruction-model'

Parameters	ß (*se*)		
	Direct writing	Observation/ weak focus	Observation/ good focus
Intercept for posttest 1	6.13 (0.22)	6.20 (0.23)	6.38 (0.24)
*Regression of posttest 1 on aptitude	0.29 (0.08)	*ns*	0.18 (0.07)
Intercept for posttest 2	7.06 (0.21)	6.63 (0.24)	6.47 (0.25)
*Regression of posttest 2 on aptitude	*ns*	*ns*	0.14 (0.07)
*Regression of posttest 2 on posttest 1	0.55 (0.10)	0.35 (0.12)	0.47 (0.14)

*Note. All *variables are centered around the mean.*

The *regression of posttest 1 on aptitude* was *not* statistically significant ($p > .05$) in condition observation/weak focus (see Table 3.3). In other words, the aptitude of these participants made no difference to their writing scores for posttest 1. However, in the two other conditions (i.e., direct writing and observation/good focus) the regression was statistically significant ($p < .05$) and did not differ between these two conditions ($\chi^2_1 = 1.08$; $p = .30$). In these conditions, aptitude influenced writing performance in a similar fashion: participants with a higher aptitude benefited more from the instruction. Thus, the effect of instruction session 1 in both the direct-writing and the observation/good-focus condition differed according to the level of aptitude. Figure 3.4 is a schematic representation of the interactions.

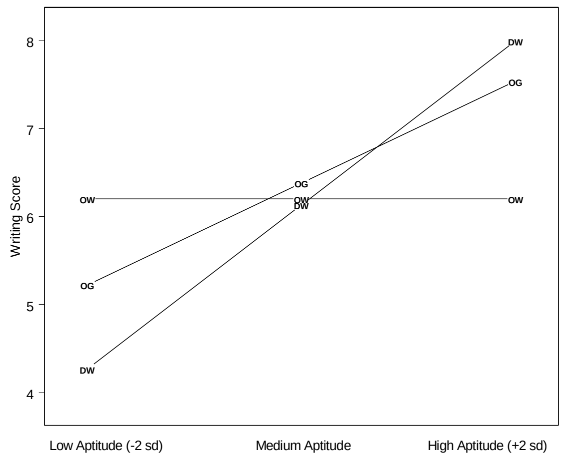


Figure 3.4. Effects of instruction for different levels of aptitude (posttest 1)
Note. 'DW' stands for condition direct writing, 'OW' stands for condition observation/weak focus, and 'OG' stands for condition observation/good focus.

When we tested the *intercept scores for posttest 1* (i.e., the mean writing scores for posttest 1 for participants with a medium aptitude) we found no significant differences between conditions (for all comparisons $\chi^2_1 < .60$; $p > .44$). Thus, for a participant with a medium aptitude condition did not affect performance. But it is better for participants with a low aptitude to be assigned to the observation/weak-focus condition, and for participants with a high aptitude to be assigned to direct-writing or the observation/good-focus condition (see also Figure 3.4).

The *regression of posttest 2 on aptitude* was *not* statistically significant in the direct-writing and observation/weak-focus condition ($p > .05$; Table 3.3). Thus, the aptitude of these participants made no difference to their writing scores for posttest 2. In the observation/good-focus condition, the regression of posttest 2 on aptitude was statistically significant ($p < .05$). Aptitude influenced the writing scores: participants with a higher aptitude benefited more from instruction session 2 (see Figure 3.5).

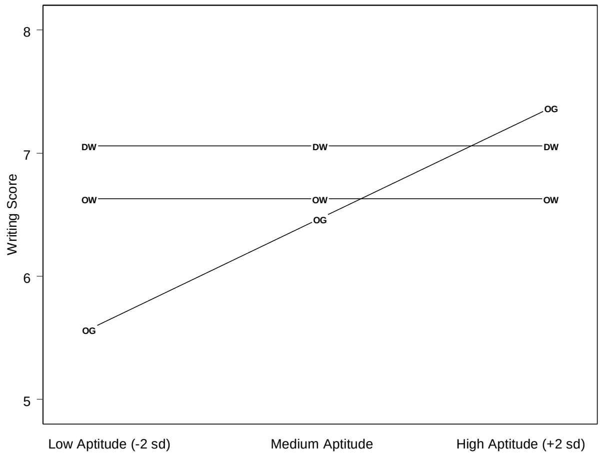


Figure 3.5. Effects of instruction for different levels of aptitude (posttest 2). Conditional on writing performance at posttest 1

It should be noted that the regression of posttest 2 on aptitude was estimated simultaneously with the regression of posttest 2 on posttest 1. This *regression of posttest 2 on posttest 1* was statistically significant in all conditions ($p < .05$) and did not differ between conditions (for all comparisons $\chi^2_1 < 1.63$; $p > .20$; Table 3.3). This means that the effects of aptitude on posttest 2 can be maintained for each level of performance. Thus, considering all effects on posttest 2, we can conclude that, for all conditions, there is an interaction between writing performance at posttest 1, and instruction in session 2. Writing performance at posttest 1 influences writing scores for posttest 2 in all conditions: participants with a good writing performance learned more than participants whose writing performance was less good.

On the other hand, the interaction between aptitude and instruction in session 2 differs. Whereas there was no interaction between aptitude and instruction for the conditions direct writing and observation/weak focus, an effect of aptitude on writing scores for posttest 2 was found for the observation/good-focus condition. Thus, in the latter, a good writing performance is not enough to gain higher writing scores for posttest 2: to benefit from this performance, participants also have to be smart.

When we tested the *intercept scores for posttest 2* (i.e., the mean writing scores for posttest 2 for participants with a medium aptitude and a medium writing performance at posttest 1), we found no significant differences between conditions (for all comparisons $\chi^2_1 < 3.44$; $p > .06$). Thus, for a participant with a medium aptitude and a medium writing performance condition did not affect performance. But it is better for participants with a low aptitude and low performance to be assigned to the observation/weak-focus or direct-writing condition, and for participants with a high aptitude and high performance to be assigned to the observation/good-focus condition (see Figure 3.5).

4 DISCUSSION

Results confirmed the 'similarity hypothesis': both instruction sessions showed that model-observer similarity in competence facilitated learning. Furthermore, results after instruction session 1 showed that weak students benefited more from observational learning (focusing on weak models) than by performing writing tasks. We assume that they profited from observational learning because their cognitive effort is shifted from *executing* writing tasks to *learning* from writing processes of others (Braaksma, Van den Bergh, Rijlaarsdam & Couzijn, 2001b; Couzijn, 1999; Rijlaarsdam & Couzijn, 2000a, 2000b). They can thus focus on the learning task, providing themselves with a learning opportunity to acquire new understanding about writing. Good students benefited not only from observational learning (focusing on good models) but also from performing writing tasks. They are probably able to divide their attention between writing task and learning task, and thus generate enough input for their learning by evaluating their own performance. Instruction session 2 showed that weak participants benefited from performing writing tasks as much as they profited from focusing their observations on weak models' writing. Perhaps familiarity with the task played a role here: because they had already experienced successes with the tasks, they had a chance to build a knowledge base about good writing, and thus became better equipped to learn from executing the writing task.

The writing tasks selected for our study were constrained. What the participants were called upon to do (and what the models they observe were doing) was order and connect ideas we gave them. The participants did not have to come up with new ideas and compose thoughts, as in the larger, less structured writing tasks which have been used in other writing studies in which modeling was a key factor (e.g., Couzijn, 1999). In future studies on the effects of model-observer similarity on writing, we thus recommend other, less structured and larger writing tasks.

Until this study, the similarity claim was restricted to a group of weak students (Schunk et al., 1987). When a group of average performers was studied, no effects of similarity between model and observer were found (Schunk & Hanson, 1989a). Our study, however, involved a heterogeneous group of participants (i.e., participants with mixed abilities), which enabled us to extend the similarity hypothesis to good learners.

Another extension regards the operationalization of similarity. In most studies, participants in the 'good-model' condition observed good models, and participants in the 'weak-model' condition observed weak models. In our case, however, all participants were exposed to pairs of a good and a weak model. The only difference between conditions was the instruction to focus their reflection on the weak *or* on the good model. So, observers watched both models at work and compared them instead of observing just one type of model. Thus, the similarity claim also holds when the exposure to models is the same but the instruction is similarity-inducing.

Because of the different operationalization of similarity, we made other choices concerning the execution of the models and the dependent variables. Although our models did verbalize their thoughts, they focused on the execution of the writing process, not on achievement beliefs and did not reflect (high and low) self-efficacy.

Because of that different focus, we didn't include self-efficacy measures as dependent variables. However, in future studies it might be advisable to include these measures because self-efficacy is hypothesized to influence choice of activities, effort expenditure, persistence and achievement (Bandura, 1986, 1997; Schunk, 1998). Perhaps the present outcomes are mediated by changes in participants' self-efficacy.

Finally, interaction effects played a central role in our study. We examined the interaction between characteristics of the participants, and characteristics of instructions. With interaction effects, one can examine which learning environment is the most effective for which student. With this idea, our results are in line with Cronbach (1957, p. 681):

> 'In general, unless one treatment is clearly best for everyone, treatments should be differentiated in such a way as to maximize their interaction with aptitude variables. Conversely, persons should be allocated on the basis of those aptitudes which have the greatest interaction with treatment variables.'

Observational learning may not be the best learning environment for all students, and not all types of observational learning are equally effective for different types of students.

APPENDICES

APPENDIX A

Overview of instructional time (in minutes) for each sequence in each condition

Sequence	Direct writing	Observation/ weak focus	Observation/ good focus
I (theory and mastery questions)	14.5	14.5	14.5
II (writing / observation tasks)	21.0	22.5	22.5
III (theory and mastery questions)	11.5	11.5	11.5
IV (writing / observation tasks)	12.0	13.0	13.0
Total learning time	59.0	61.5	61.5

APPENDIX B

Mean and standard deviation for both writing measurements in each condition

	Posttest 1		Posttest 2	
Condition	*M*	*SD*	*M*	*SD*
Direct writing	6.04	2.09	7.09	2.10
Observation/weak focus	6.19	1.83	6.65	1.79
Observation/good focus	6.43	1.74	6.56	1.88

CHAPTER 4

EFFECTS OF OBSERVATIONAL LEARNING ON WRITING PROCESSES

An experimental study

Abstract

This study examined why observational learning has positive results in learning outcomes of writing tasks. It is hypothesized that observational learning affects task processes differently than 'learning-by-doing'. The study focused on the effects on orchestration (i.e., temporal organization) of writing processes. An experiment was set up in which participants (*N*= 52, eighth grade) were assigned to an observational-learning condition or an executional-learning condition. To measure their orchestration of writing processes, the participants performed posttest writing tasks under think-aloud conditions. The results largely confirmed the hypotheses, that is, compared to 'learning-by-doing', observational learning resulted in a process pattern with more metacognitive activities (Goal-orientation and Analysis) in the beginning and more executional activities (Writing and Re-reading) in the second part of the writing process. Over the whole writing process, writers who learned by observation showed more Planning activities. In contrast, 'learning-by-doing' resulted in a pattern with more executional activities (Writing) at the start and more Formulating activities during the whole writing process.

1 INTRODUCTION

Observing is the key learning activity in learning environments in which learners learn from models. Observational learning occurs when observers display new behaviors that, prior to modeling, had no probability of occurrence, even when motivational inducements are offered (Schunk, 1987, p. 150). There is a large body of research about effects and conditions of observational learning and modeling (Bandura, 1986, 1997; Rosenthal & Zimmerman, 1978; Schunk, 1987, 1991, 1995, 1998). With teachers, adults or students as models, observational learning has proved to be effective with learners of various ages and in various school subjects, such as mathematics (Schunk & Hanson, 1985, 1989a, 1989b; Schunk, Hanson & Cox, 1987), reading (Couzijn, 1995, 1999), writing (Braaksma, Rijlaarsdam & Van

Braaksma, M. A. H., Rijlaarsdam, G., Van den Bergh, H., & Van Hout-Wolters, B. H. A. M. (submitted). *Observational learning and its effects on the orchestration of writing processes.*

den Bergh, in press-b; Couzijn, 1995, 1999; Graham & Harris, 1994; Graham, Harris & Troia, 1998; Schriver, 1992), and speaking and listening (Sonnenschein & Whitehurst, 1984).

In all these studies, the effects of observational learning are described in learning *results*. It is still unclear how observational learning influences task execution, and it is this relation that we want to examine in the current study. We will focus on the domain of writing. Equally, the effects of observational learning on writing are only described in terms of learning products, and nothing is known about the effects on writing *processes*. In the next sections, we will elucidate that observational learning might effect orchestration of writing processes differently compared to the more traditional 'learning-by-doing'. First, we will define 'orchestration' and explain why orchestration is important and how differences in orchestration are caused.

1.1 Orchestration of writing processes

Several authors (Graham & Harris, 2000; McCutchen, Covill, Hoyne & Mildes, 1994; McCutchen, 2000) use the term 'orchestration' to emphasize that writing processes must be activated and coordinated by a control structure, such as the monitor in the Hayes and Flower model (1980). Temporal management of writing processes plays a central role in writing activity: writers have to continually shift between planning the main ideas, content translating and text revising. This continuous shifting gives the writing activity its cyclic nature and depends on a strategic and recursive ordering of different sub-processes (Alamargot & Chanquoy, 2001, p. 125).

Why is orchestration of writing processes important? Studies that modeled writing processes and related these processes to quality of writing products, showed that orchestration of writing processes is a decisive factor contributing to text quality. It appears that good and weak writers rely on the same cognitive and metacognitive activities, but that they differ in the way they distribute these activities over the writing process. Breetvelt, Van den Bergh and Rijlaarsdam (1994) compared writing processes related to better texts (good writers) with writing processes related to weaker texts (weak writers). They showed that what is decisive is not how often or how much writers execute (meta) cognitive activities, but the temporal organization of these activities over the writing process. Weak writers in their study started a certain activity too early or too late compared to good writers. For instance, weak writers tended to be premature in setting goals (i.e., formulating task demands derived from the assignment or self-devised) and in revising text, and too tardy with reading and interpreting topic information and documentation, and with evaluating text production.

Similarly, Rijlaarsdam and Van den Bergh (1996) found that weak writers were engaged in the process of structuring ideational elements at the end of the process, while good writers paid attention to this process at the beginning of the writing process. Focusing on sequences of different processes of generating ideas, Van den Bergh and Rijlaarsdam (1999) found different patterns for weak and good writers. At the beginning of the process, weak writers exploited translating-generating patterns (i.e., producing text), thereby resorting too early to a knowledge-telling process (Bereiter & Scardamalia, 1987); good writers showed this pattern only half way through the process, relying at the beginning of the process on their analysis of the information from the task at hand. Van den Bergh and Rijlaarsdam (2001) focused on effective patterns of occurrence of task representation activities and formulating activities. They found that task representation activities were positively correlated with text quality only during the initial phase of writing, whereas for formulating activities the correlation changed from negative at the start to positive at the end.

The previous results are found for fifteen-year old students, but also with younger children (twelve years) it is established that the temporal organization of cognitive activities is a significant factor for text quality. Van der Hoeven (1997) found that the correlation between generating text and text quality proved to be time dependent. The positive correlation between generating and text quality diminishes as the writing process passes, and is even negative in the last part of the writing process.

In line with Breetvelt et al. (1994), Graham and Harris (2000), McCutchen et al. (1994), McCutchen (2000), Rijlaarsdam and Van den Bergh (1996), Van den Bergh and Rijlaarsdam (1999, 2001), and Van der Hoeven (1997), we define *orchestration of writing processes* as 'the temporal organization of cognitive activities' and we use this term to emphasize the importance of the distribution of cognitive activities over time.

How are differences in orchestration caused? Most probably, they are due to differences in the writers' knowledge base. According to Hayes (1996) the long-term memory (LTM) comprises a set of different kinds of knowledge that ensure different functions within the writing activity: the type of text (genre knowledge), the addressee of the text (audience knowledge), the linguistic components necessary for the realization of the text (linguistic knowledge), the area of the content of the text (topic knowledge), and the procedures to guide and control the effective realization of the text production (task schemas). For genre knowledge for instance, it is assumed that familiarity with a genre influences writing by providing access to an organized scheme in LTM (McCutchen, 2000, p.17). The influence of genre knowledge on writing has been examined in several studies.

A study by Ferrari, Bouffard and Rainville (1998) indicated that qualitative differences in knowledge are crucial. These researchers found that good and weak writers differ in their discourse knowledge, in this case knowledge about how to construct a comparative text. Good writers showed a deeper grasp of that text structure than did poor writers (Ferrari et al., 1998, p. 483). McCutchen (2000) makes the same finding, arguing that: 'Skilled writers seem to access a macrostructure for the text on which they are working, and such macrostructures are derived from their general knowledge of text structures, or genres' (p.18). Englert, Stewart and Hiebert (1988) found that text structure knowledge was related to students' abilities to structure and organize their writing. For better writers, sensitive to the pattern of superordinate and subordinate ideas in text, text structures set up strong expectations for specific text forms and details (Englert et al., 1988, p. 150). Similarly, Schoonen and De Glopper (1996. p. 96) found a relation between writing performance and knowledge about writing. When describing how to write, poor writers focused on presentation, mechanics and grammar aspects, good writers focused on text structure.

1.2 Observational learning and the orchestration of writing processes

When we know that due to a richer knowledge base, different ways of orchestration result in different quality of text, and know that observational learning results in a better text quality than the more traditional 'learning-by-doing' in writing instruction, the missing link is in what way observational learning affects the knowledge base in a different way than 'learning-by-doing'. And thus, why observational learning influences the orchestration of writing processes differently than learning by doing.

We assume that in observational learning students can pay more attention to learning than in traditional ways of instruction, where cognitive effort is paid more to writing, than to learning to write. When students learn to write, metaphorically, one might say that learners have to juggle with two agendas: a writing agenda and a learning-to-write agenda. In a way, they have to attain two goals simultaneously: they must produce an adequate text, and they must acquire new understanding about writing. When students learn to write by observation instead of executing writing tasks, the cognitive effort is shifted from *executing* writing tasks to *learning* from writing processes of others (Braaksma, Van den Bergh, Rijlaarsdam & Couzijn, 2001b; Couzijn, 1999; Rijlaarsdam & Couzijn, 2000a, 2000b). In observational learning, students do not have to juggle with the 'dual agenda': they don't write! They can focus on the learning task, providing themselves with a learning opportunity to acquire new understanding about writing and to enlarge their knowledge base.

We suppose that the knowledge base of writers is fed by regulatory or metacognitive strategies such as observation, evaluation, and reflection. In the enactive feedback loop in which one learns from the consequence of one's actions, personal observations and judgments are indispensable (Zimmerman, 1989, 2000; Zimmerman & Schunk, 1989). By using metacognitive strategies, writers gain information that changes what they know and do. They re-conceptualize and re-evaluate the writing behavior and working method, and refine strategies for use in new tasks (Braaksma et al., 2001b; Couzijn, 1995, 1999; Oostdam & Rijlaarsdam, 1995). Students who learn by observation are stimulated to use and address these important metacognitive strategies explicitly, because the observation of the performance of others involves a 'natural' step back and thus a natural type of monitoring, evaluation, and reflection on task execution processes (Braaksma et al., 2001b; Butler, 1998; Graham & Harris, 1994; Rijlaarsdam & Couzijn, 2000a).

In sum, when learning by observation, through moving cognitive effort to learning to write, and practicing of metacognitive strategies that feed the knowledge base, students are supposed to build a richer knowledge base about writing. And as we proposed before, a richer knowledge base about writing influences the orchestration of writing processes.

For instance, when students have a richer knowledge base about writing, they know different types of writing tasks, they know the ‘macrostructures’ of the texts or tasks they have to work on. Suppose that these students have to write an argumentative text, then, because they know the structure and components of such text type, they can plan the several steps they have to take, and they can analyze the several parts an argumentative text consists of (i.e., standpoint, connectives, and (subordinate) arguments). This leads to a different orchestration of writing processes: the students will (re)define the assignment, and then retrieve the planning scheme associated with the typical genre or task, and then plan the steps of the task and analyze, at least to some extent, some parts an argumentative text consists of. After that, they will write (parts of) the actual text. In contrast, students who’s knowledge base lacks a ‘macrostructure’ that can guide them through the execution of the task, will show more ‘knowledge-telling processes’ (producing text) without planning and analyzing.

Another important point is that models provide students with orchestration examples in real time. These examples can trigger ‘slumbering’ powers from the observing students to perform such orchestration as well. This process could be strengthened when students observe pairs of peer models who represent the ‘zone of proximal development’.

1.3 The present study

In this study we will focus on the orchestration of writing processes. We hypothesize that observational learning affects the orchestration of writing processes differently than ‘learning-by-doing’. Based on Breetvelt et al. (1994), Rijlaarsdam and Van den Bergh (1996), Van den Bergh and Rijlaarsdam (1999, 2001), and Van der Hoeven (1997), and the examples in the previous section, we specify the resulting differences in orchestration. Observational learning will result in a process pattern with more metacognitive activities such as Goal-orientation, Planning, and Analysis in the beginning of the writing process and more executional activities such as Formulating, Writing, and Re-reading in the second part of the writing process than ‘learning-by-doing’. In contrast, ‘learning-by-doing’ will result in a process pattern with more executional activities during the whole writing process, but especially at the start.

We set up an experiment in which participants were assigned to an observational-learning condition or a control condition. For reasons of generalization, we implemented two observational-learning conditions (see for more details Design section, page 50). To measure their orchestration of writing processes, the participants performed posttest writing tasks under think-aloud conditions.

2 METHOD

2.1 Participants

Participants were students taken from nine eighth grade mixed ability classes in a Dutch city school, with an average age of fourteen years. Because the experiment took place at school and was part of the regular lessons in Dutch, we had to assign whole classes to conditions instead of individual participants. The assignment of classes to conditions was semi-random, in the sense that stratification was applied according to which teacher taught Dutch in more than one class, and also according to the lesson roster. Further assignment within the strata was random. As a result, three classes took part in each condition. There were no differences in mean age and gender between conditions.

We asked the teachers to select from each class six students for participation in the present study, which required thinking aloud when writing argumentative texts in a posttest. In a short letter, the teacher was asked to select students from his class who were able to think aloud. It was emphasized that this does not imply that these students should be good in the subject Dutch, the only importance was that the students were able to put their thoughts into words.

Two participants were excluded from this initial sample. One participant wrote the wrong text under thinking aloud conditions, and the other participant was absent during the instruction session and thus was not able to write the specific text type while thinking aloud. Thus, in total 52 participants were included in the analyses.

2.2 Design

For reasons of generalization, in addition to a control condition doing writing (DW) in which participants performed short writing tasks, we implemented two observational-learning conditions. We choose for two different conditions because observational learning depends on a number of factors (Schunk, 1991, p. 113). One of these factors is the operationalization of the type of model. Studies show several variations in model competence, model age, number of models and model sex (see Schunk, 1987 for a review). Concerning the different operationalizations of the type of models, it would be desirable to generalize our hypothesis that observational learning results in different patterns of orchestration than ‘learning-by-doing’.

In our study, participants in both observational-learning conditions observed the same pairs of models, in the same order. The one and only difference was the model to focus on: instructions in the observation-focus-weak-models condition (WM) required participants to reflect on the performance of the weak model, and those in the observation-focus-good-models condition (GM) to reflect on the performance of the good model.

In each condition, an almost equal amount of participants took part: 18 participants in condition doing writing, 17 participants in condition observation focus weak models, and 17 participants in condition observation focus good models.

Participation in the experiment took place in three sessions. In the first session pretests were administered which measured verbal intelligence and skill in ‘identification of argumentation’. After one week, in the second session the instruction session took place in which participants learned how to write short argumentative texts. A special writing task was used: students learnt to transform argumentation structures into short linear argumentative texts (see below: Instruction session). This writing task was completely new for all participants. The posttest session took place after two weeks and consisted of the measurement of writing processes. All sessions were conducted by the researcher and research assistants.

Analyses of variance showed no initial differences between the three conditions on a pretest which measured 'identification of argumentation' (F (2, 48) = 0.78, p > .05) and on a pretest which measured verbal intelligence (F (2, 49) = 0.71, p > .05).

2.3 Instruction session

For this study, we developed a short course on an aspect of writing argumentative texts. The course was adapted with permission from Couzijn (1995, 1999). Materials for this session consisted of a workbook and a CD-ROM, projected onto a large screen in front of the class room. Participants were informed of what to do by means of on-screen messages, e.g., reading theory in the workbook, answering a mastery question, writing a short text or observing two models. An on-screen timer indicated how much time was left for each activity; participants were alerted by short beeps when the time had almost elapsed.

The participants' workbooks consisted of four main sequences. In the first sequence, participants had to study a theoretical section about main concepts concerning theory of argumentation (e.g., standpoint, argument). To stimulate active reading of the theory, the participants answered mastery questions. In the second sequence, participants applied the theory in different types of exercises according to the condition to which they had been assigned. Participants in the control condition transformed given argumentation structures into linear texts: they wrote short argumentative texts based on the argumentation structures they had been given. Meanwhile, participants in the observational-learning conditions observed pairs of videotaped peer models writing the short argumentative texts based on argumentation structures required in the control condition. These models were thinking aloud while they were writing. In two questions: (a) 'Which model did less well?' and, (b) 'Explain briefly what this (less good) model did worse.', the observers in the observation-focus-weak-models condition were instructed to focus their attention on the performance of the weak model. Participants in the observation-focus-good-models condition observed the same pairs of models, but were instructed to focus on the good model, by answering two questions: (a) 'Which model did well?' and, (b) 'Explain briefly what this (good) model did well.' By using videotaped models on CD-ROM, standardized presentation across participants was ensured.

After completing the exercises, the participants went on to the third sequence, which contained theory and mastery questions about subordinate argumentation, argumentation structures and argumentative connectives. Finally, in the fourth sequence, depending on the condition, they performed two writing tasks or an observation task.

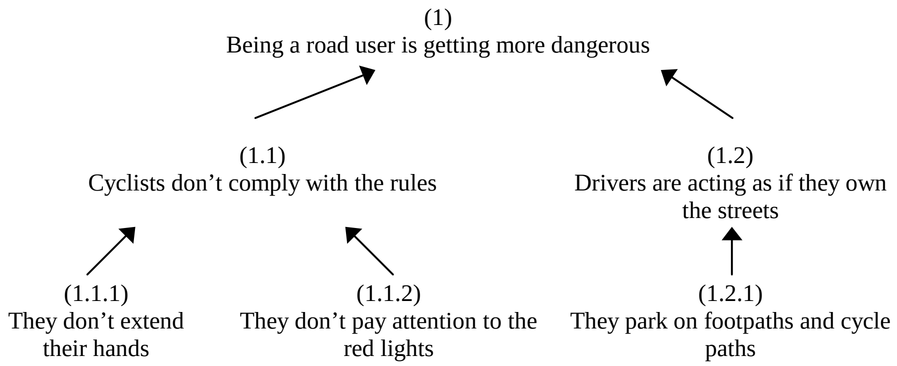
In total, in the instruction session participants either performed seven short writing tasks (during 33 minutes), or observed five pairs executing these writing tasks (during 35.5 minutes). In all conditions the total time spent on reading theory and answering mastery questions was 26 minutes. Thus, total learning time in control condition and observational-learning conditions was almost equal (59 minutes in control condition and 61.5 minutes in the observational-learning conditions).

2.4 Measurement of writing processes

In the posttest session, which took 40 minutes, the participants performed two posttest writing tasks under think-aloud conditions. The writing tasks consisted of transforming a given argumentation structure into a linear argumentative text. In each case, participants completed first a simple, and then a more complex task. Figure 4.1 shows a sample of a writing task.

In total, 104 think-aloud protocols (52 participants, two tasks) were audio taped, typed, fragmented and scored. For scoring the cognitive activities, we adapted categories from an instrument used in our previous writing process studies (Breetvelt, Van den Bergh & Rijlaarsdam, 1994, 1996; Rijlaarsdam & Van den Bergh 1996; Van den Bergh, Rijlaarsdam & Breetvelt, 1994; Van den Bergh & Rijlaarsdam, 1996, 1999, 2001). These categories are based on the writing process model of Hayes and Flower (1980).

Task: Write a short argumentative text based on the following argumentation structure. Make sure that the reader will understand the standpoint and arguments in your text.



Write your text here:

………………………………………………………………….
………………………………………………………………….
(...)

Figure 4.1. Writing task (traffic)

For report reasons, we distinguish two main types of activities: metacognitive activities, and executional activities. All activities dealing with some sort of planning the development, and checking of planning are grouped into the super category of metacognitive activities (Goal-orientation, Planning, Analysis, Evaluation, and Meta-analysis). Activities needed for text production (forward or backward activities) are grouped as executional activities (Formulating, Writing, Re-reading, and Revision). The activity of Pausing was also placed in this group, because of the unclear cognitive status of pausing. Table 4.1 shows a description of the categories, along with some protocol examples.

To illustrate the tasks and the way in which the protocols were coded, the example in Figure 4.2 shows the execution of a writing task. Here, one of the participants, Sanne, is presented with a hierarchical argumentation structure on paper (see Figure 4.1), which she has to transform into linear text. The writing strategy and the sequence of activities of Sanne can be followed via the coded activities in the third column.

Table 4.1. Cognitive activities and their descriptions

Cognitive activity	Description
Goal-orientation	Reading (parts of) the argumentation structure in the assignment and (re)definition task (for instance, ‘So, I have to write a text’)
Planning	Local planning. Planning of a step in the process. The participant coaches himself through the task, usually by uttering a temporal indication (for instance, ‘First, I'll start with the first argument’)
Analysis	Analyzing the argumentation structure of the text, labeling elements as standpoints, arguments, and so forth (for instance, ‘Y is the standpoint’)
Evaluation	Evaluating (parts of) formulations, analyses, or already written text (for instance, ‘That is not a nice sentence’)
Meta-analysis	Monitoring and regulating task execution: e.g., stopping the process for some reason; re-generating information on how to handle this kind of task; making a remark on the level of difficulty of the task or on own ability (for instance, ‘This is rather a difficult one’)
Formulating	Formulating (parts of) texts before writing it down
Writing	Dictating (parts of) texts while writing it down
Pausing	Silence and/or sounds indicating thoughtfulness (for instance, ‘Ehm, ehm’)
Re-reading	Re-reading (parts) of already written text
Revision	Revising aspects of text by addition, deletion, or transposition

Sanne first reads the argumentation structure and then analyses and formulates the different ingredients of the argumentation structure (standpoint, arguments, and subordinate arguments). Next, she thinks about the place she will put the standpoint. Deciding to end with it, she then writes the entire text without analyzing the ingredients until she comes to the standpoint. At that point, she plans the standpoint and writes it down.

Two coders fragmented the protocols and coded the fragments. Each fragment contained a new cognitive activity (e.g., fragment 1: Analysis standpoint, fragment 2: Formulating standpoint) or the same type of cognitive activity but with a new 'content' (e.g., fragment 1: Analysis standpoint, fragment 2: Analysis argument). To determine inter coder reliability, a contingency coefficient was calculated over 18 randomly selected protocols coded by both coders. Cohen's kappa, corrected for chance, equaled 0.802 (*se* 0.025).

No	**Utterance in thinking-aloud process**	**Cognitive activity**
1.	1 Being a road-user is getting more dangerous, 1.1 Cyclists don't comply with the rules, 1.1.1 They don't extend their hands, 1.1.2 Cyclists don't pay attention to the red lights, 1.2 Drivers are acting as if they own the streets, 1.2.1 They park on footpaths and cycle paths	Reads argumentation structure (= Goal-orientation)
2.	The standpoint is	Analyses standpoint (= Analysis)
2.	Being a road-user is getting more dangerous	Formulates standpoint (= Formulating)
3.	Argument 1.1 is	Analyses argument (= Analysis)
4.	Cyclists do not comply with the rules	Formulates argument (= Formulating)
5.	And that is because	Analyses subordinate arguments (= Analysis)
6.	The cyclists don't extend their hands and cyclists don't pay attention to the red lights	Formulates subordinate arguments (= Formulating)
7.	And the argument 1.2 is	Analyses argument (= Analysis)
8.	Drivers are acting as if they own the streets	Formulates argument (= Formulating)
9.	And that is because	Analyses subordinate argument (= Analysis)
10.	The drivers park on footpaths and cycle paths	Formulates subordinate argument (= Formulating)
11.	And now I am thinking if I will first take the argument and then the standpoint	Plans (= Planning)
12.	Then, it must be thus	Analyses connective (= Analysis)
13.	Cyclists do not comply with the rules	Formulates argument (= Formulating)
14.	Because they don't extend their hands	Formulates subordinate argument (= Formulating)
15.	Extend their hands	Dictating subordinate argument while writing (= Writing)
16.	And they don't pay attention to the red lights	Dictating subordinate argument while writing (= Writing)
17.	Further, drivers are acting as if they own the streets	Formulates argument (= Formulating)
18.	because they park on footpaths and cycle paths	Formulates subordinate argument (= Formulating)
19.	Further, drivers are acting as if they own the streets	Dictating argument while writing (= Writing)
20.	because they park on footpaths and cycle paths	Dictating subordinate argument while writing (= Writing)
21.	And then the standpoint	Plans standpoint (= Planning)
22.	Thus, being a road-user is getting more dangerous	Analyses standpoint (=Analysis)
23.	Thus, being a road-user is getting more dangerous	Dictating standpoint while writing (= Writing)
24.	Ok.	Stops/starts new cognitive activity (= Meta-analysis)

Figure 4.2. Example thinking-aloud protocol writing task (traffic)

2.5 Analyses

There are several reasons to use multilevel models (see Van den Bergh & Rijlaarsdam, 1996). The most important one is that multilevel models take the hierarchy of the data into account. In our case this means that the observations of the occurrence of processes are considered as nested within writers; observations 'belong' to a specific writer. This quality of multilevel models makes it possible to analyze the occurrence of processes for individual writers. There is no need to aggregate over all processes of each writer (resulting in frequencies per writer).

Suppose, Y_{ij} is a response variable indicating whether a certain process did occur ($Y_{ij} = 1$) or did not occur ($Y_{ij} = 0$). In order to analyze the distribution of occurrences –and the differences between individuals in probabilities of occurrence– we have to relate this response variable to the moment in the writing process. Polynomial models are very convenient for this purpose. That is, the response variable is modeled as powers of time in the writing process (i.e., $t_{ij}^0, t_{ij}^1, t_{ij}^2, \ldots, t_{ij}^k$). Such polynomials can take almost any shape, dependent on the size and the number of coefficients. Besides, the interpretation of this type of growth-curve is relatively easy compared to other types of model. Because the dependent variable is dichotomous, we prefer a logit[4] transformation:

$$\text{Logit}\,(P_j \mid Y_{ij} = 1) = \beta_{0j} * t_{ij}^0 + \beta_{1j} * t_{ij}^1 + \beta_{2j} * t_{ij}^2 + \ldots + \beta_{kj} * t_{ij}^k. \quad (1)$$

In fact, for each individual *j* the proportion of occurrence of a certain process is modeled as a function of parameters of t_{ij}, a variable indicative of the moment in the writing process (compare Breetvelt et al., 1994). The number of coefficients in the polynomial can be considered as an empirical question. That is, for each process the model needs to be as sparse as can be (the smallest number of coefficients) while still describing the observations adequately. A second constraint is that higher powers of t_{ij} are not taken into consideration if a lower one does not reach significance (Goldstein, 1979).

The model in Equation (1) describes different regression lines for each writer, which may or may not partly coincide. So, there are as many regression lines describing the occurrences of a process during writing as there are individuals. This would be a rather clumsy way of analysis. Therefore, we consider each regression coefficient of each individual writer as a deviation from a mean coefficient:

$$\begin{aligned} \beta_{0j} &= \beta_0 + u_{0j} \\ \beta_{1j} &= \beta_1 + u_{1j} \\ \beta_{2j} &= \beta_2 + u_{2j} \\ &\ldots \\ \beta_{kj} &= \beta_k + u_{kj}. \end{aligned} \quad (2)$$

4 Remember logit (Pj) = Log (pj / 1 – pj) = Log (Fj / Nj – Fj), in which pj denotes the relative proportion, or probability of occurrence, and Fj and Nj denote the frequency of occurrence and the number of observations respectively.

Substitution of Equation (2) in Equation (1) gives the model to be estimated:

$$\text{Logit}\,(P_j \mid Y_{ij} = 1) = B_{0j} * t_{ij}^{0} + \beta_{1j} * t_{ij}^{1} + \beta_{2j} * t_{ij}^{2} + \ldots + \beta_{kj} * t_{ij}^{k} + [u_{0j} * t_{ij}^{0} + u_{1j} * t_{ij}^{1} + u_{2j} * t_{ij}^{2} + \ldots + u_{kj} * t_{ij}^{k}]. \quad (3)$$

The model according to Equation (3) consists of two parts: a fixed part and a random part (between square brackets). The fixed part of the model describes the mean, or inter-individual pattern in occurrences of the process analyzed. The first coefficient (β_0) indicates the intercept, the second (β_1) the linear change with time, the third (β_2) the quadratic change with time, etc.

The random part of the model summarizes the inter-individual differences as deviations from the mean changes during the writing process (i.e., u_{0j} is indicative of differences in intercepts between writers, u_{1j} shows differences in linear change during the writing process, etc.). These residual scores are assumed to be normally distributed with an expected mean of zero and a variance of S^2u_{oj}, …, S^2u_{kj}. As the variance of a residual score reaches significance, this is interpreted as indicating that this coefficient differs between writers. So, if for a certain process only the variance of the intercept (S^2u_{oj}) would reach significance, this would mean that for all writers the orchestration of this process is relatively equal, all regression lines show a parallel change over time, only the mean value differs between writers. Please note that in the random part of the model for each writer residual scores (for linear effects, quadratic effects, etc.) are estimated. So, for each writer different regression lines are approximated.

The model in Equation (3) does not distinguish between the three conditions (doing writing (DW), observation focus weak models (WM), and observation focus good models (GM)). The model however can easily be extended in order to estimate differences between conditions in either the fixed and/or the random coefficients. Let DW_{ij}, WM_{ij}, GM_{ij} be three dummy variables which are turned on (equal 1) only if an observation took place in that condition and are turned off (equal 0) otherwise. Now the model can be written as:

$$\begin{aligned}\text{Logit}\,(P_j \mid Y_{ij} = 1) = {} & DW_{ij}\,(\beta_0 * t_{ij}^{0} + \beta_1 * t_{ij}^{1} + \beta_2 * t_{ij}^{2} + \ldots + \beta_k * t_{ij}^{k}) + \\ & WM_{ij}\,(\beta_0 * t_{ij}^{0} + \beta_1 * t_{ij}^{1} + \beta_2 * t_{ij}^{2} + \ldots + \beta_k * t_{ij}^{k}) + \\ & GM_{ij}\,(\beta_0 * t_{ij}^{0} + \beta_1 * t_{ij}^{1} + \beta_2 * t_{ij}^{2} + \ldots + \beta_k * t_{ij}^{k}) + \\ & [\,DW_{ij}\,(u_{0j} * t_{ij}^{0} + u_{1j} * t_{ij}^{1} + u_{2j} * t_{ij}^{2} + \ldots + u_{kj} * t_{ij}^{k}) + \\ & WM_{ij}\,(u_{0j} * t_{ij}^{0} + u_{1j} * t_{ij}^{1} + u_{2j} * t_{ij}^{2} + \ldots + u_{kj} * t_{ij}^{k}) + \\ & GM_{ij}\,(u_{0j} * t_{ij}^{0} + u_{1j} * t_{ij}^{1} + u_{2j} * t_{ij}^{2} + \ldots + u_{kj} * t_{ij}^{k})].\end{aligned} \quad (4)$$

So, in principle, differences between conditions in mean changes in occurrences of a process as well as differences in variance can be estimated.

It makes sense to study the differences between conditions in terms of individual writers, as this shows how many writers are responsible for an observed difference between conditions. Therefore, in the random part of the model according to Equation (4) a distinction between conditions is made. As a result, the individual regression lines are allowed to differ between conditions.

The difference of two nested models, i.e., the one in Equation (3) (which does not allow for differences between conditions) and the one in Equation (4) (which allows for differences due to condition) can be tested by means of the difference in the respective -2 log likelihood. This difference is chi-square distributed with the difference in number of estimated parameters as degrees of freedom. With this test only a general difference between conditions can be assessed. In order to show processing differences between conditions one has to take the corresponding parameter estimates per condition into account. Therefore, changes in processing during task execution will be presented visually.

3 RESULTS

3.1 Effects of conditions on orchestration

For all cognitive activities (except Pausing and Revision) it appeared that the orchestration of writing processes differs systematically between conditions as was hypothesized. For Goal-orientation, Planning, Analysis, Evaluation, Meta-analysis, Formulating, Writing, and Re-reading, it was found that the model in Equation (4) with different parameters for different conditions had a better fit than the more simple model with the same parameter estimation for all conditions in Equation (3). In all cases, the drop in χ^2 was larger than 28 with 7 degrees of freedom (p = .00) except Pausing and Revision[5] for which no differences between conditions could be shown.

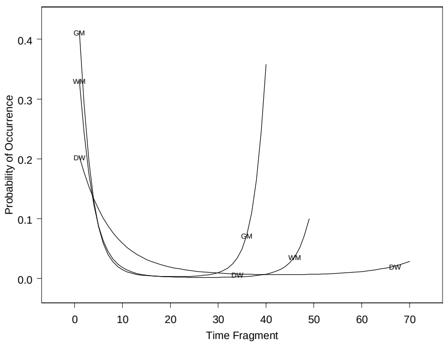
With the comparison of model fit, we can confirm our first hypothesis: observational learning affects the orchestration of writing processes differently than 'learning-by-doing'. The descriptive information in Figure 4.3 shows how the orchestration is influenced by conditions. In this figure, for each cognitive activity, the curves of the orchestration of writing processes for the different conditions are presented. The metacognitive activities are presented first, and the executional activities secondly. The horizontal axis represents time over the writing process in the form of time fragments. On the vertical axis, the probability of occurrence of a cognitive activity is represented. This probability is calculated back from logit scores to proportions by means of an expit transformation (the reverse of a logit transformation). Note that the scale of the vertical axis differs for the different cognitive activities. Parameter estimates (random part and fixed part) of the cognitive activities are given in Appendix A.

To make a more clear interpretation of the curves in Figure 4.3, we will explain the patterns of two cognitive activities: Goal-orientation and Analysis. Then, we will test our theoretical claims by summarizing the information in Figure 4.3.

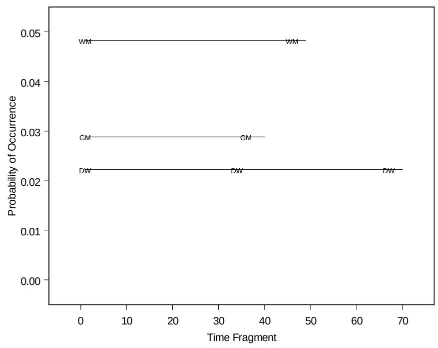
---

5 For Pausing the drop in χ^2 was 0.32 with 3 degrees of freedom (p = .96), and for Revision the drop χ^2 was 0.98 with 3 degrees of freedom (p = .81). Both are not statistically significant thus there are no differences between conditions.

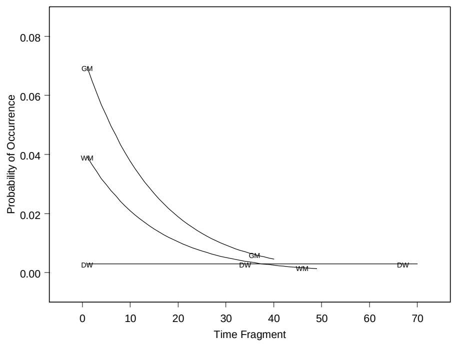
Figure 4.3. Orchestration of cognitive activities per condition.
'DW' stands for condition doing writing, 'WM' for condition observation focus weak models, and 'GM' for condition observation focus good models



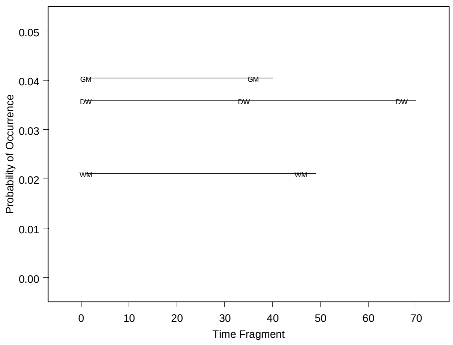
Goal-orientation



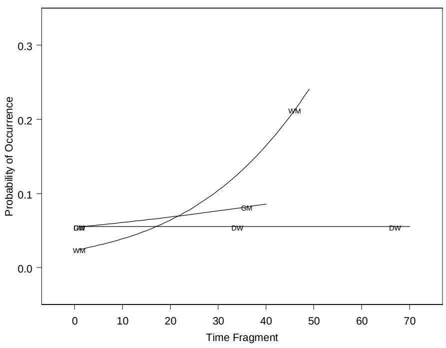
Planning



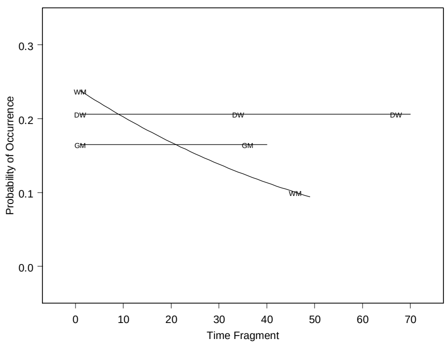
Analysis



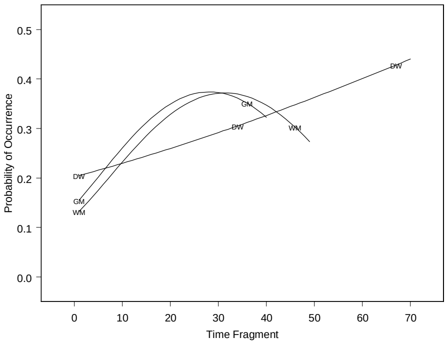
Evaluation



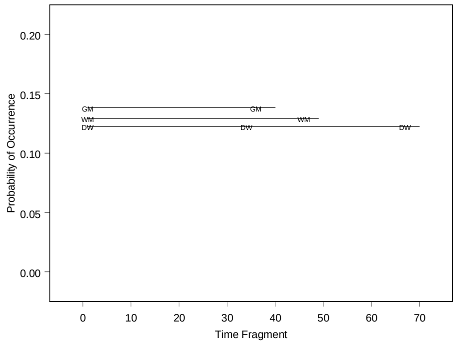
Meta-analysis



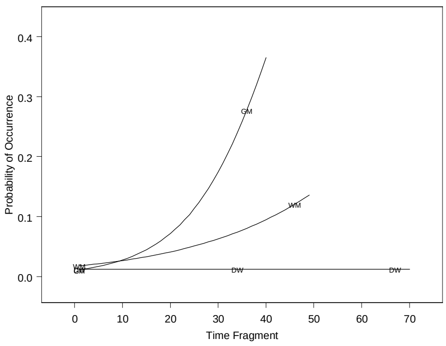
Formulating



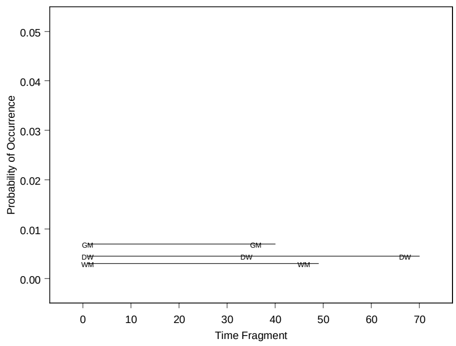
Writing



Pausing



Re-reading



Revision

Goal orientation. Figure 4.3 shows that the orchestration of Goal-orientation differs between conditions. Writers in condition observation focus weak models, and especially in condition observation focus good models, are more likely to engage in Goal-orienting processes at the start of the writing process than writers in control condition doing writing. Furthermore, Figure 4.3 clearly shows that the probability of occurrence of Goal-orienting activities changes over time. In the middle of the writing process, in all conditions almost no Goal-orienting activities are likely to occur. At the end of the writing process, writers in both observational-learning conditions, but again especially in focus-good-models, are more likely to perform Goal-orienting processes (i.e., there to check if their text reflects the argumentation structure).

Analysis. In Figure 4.3 large differences between conditions in orchestration of Analysis can be observed. In contrast with both observational-learning conditions, writers in control condition doing writing are not likely to perform Analyzing activities. Writers in both observational-learning conditions (and especially in condition focus good models) show a change in probability in occurrence of Analyzing activities over time. In the beginning of the writing process, they are likely to engage in Analyzing activities, but as the writing process progresses, the probability of performing Analyzing activities in these conditions decreases.

3.2 Testing theoretical claims

We hypothesized that observational learning will result in a process pattern determined by metacognitive activities such as Goal-orientation, Planning, and Analysis in the beginning and more executional activities such as Formulating, Writing, and Re-reading in the second part of the writing process. In contrast, we hypothesized that 'learning-by-doing' will result in a pattern with more executional activities during the whole writing process, but especially in the first part. In this section, we will test these hypotheses by comparing the orchestration of writing processes in the control condition with condition observation focus weak models. Then, we will examine if we can generalize our findings to the other observational-learning condition.

Table 4.2 summarizes the curves in Figure 4.3 and presents a comparison in orchestration of writing processes between condition observation focus weak models and doing-writing condition. This is done over three parts of the writing process: start, middle and end. Like Figure 4.3, the metacognitive activities are presented first, followed by the executional activities.

Table 4.2 shows that in the beginning of the writing process, writers in condition observation focus weak models perform relatively more metacognitive activities as Goal-orientation and Analysis than writers in the control condition. Furthermore, writers in condition observation focus weak models perform at the start relatively less Writing activities, but more Formulating activities than writers in condition doing writing. In the middle of the writing process, writers in condition observation focus weak models perform still relatively more Analyzing activities but also more executional activities as Writing and Re-reading than writers in doing-writing condition. Furthermore, in the middle of the writing process they perform less Formulating activities but they start to perform a lot of Meta-analyzing activities: in the first part of the middle of the writing process they perform still less Meta-analyzing activities but in the second part, they perform more Meta-analyzing activities than writers in doing-writing condition. At the end of the writing process, the increasing Meta-analyzing activities continue for writers in observation-focus-weak-models condition. Writers in that condition also show more Re-reading activities in the last part of the writing process than writers in control condition doing writing. Moreover, writers in observation-focus-weak-models condition show relatively more Goal-orienting activities at the end of the writing process than writers in the control condition. Finally, writers in observation-focus-weak-models condition show during the whole writing process a constant orchestration of relatively more Planning activities, but less Evaluation activities than writers in doing-writing condition.

Table 4.2. Pattern for orchestration of cognitive activities over start (time fragment 0-10), middle (time fragment 10-30) and end (time fragment 30-70) of the writing process in condition observation focus weak models related to condition doing writing

	Observation focus weak models		
Cognitive activity	Start	Middle	End
Goal-orientation	+		+
Planning	+	+	+
Analysis	+	+	
Evaluation	–	–	–
Meta-analysis	–		+
Formulating	+	–	–
Writing	–	+	
Pausing			
Re-reading		+	+
Revision			

Note. *'–' indicates less cognitive activities compared to condition doing writing, and '+' indicates more cognitive activities in comparison with condition doing writing. No sign indicates the same amount of cognitive activities as in condition doing writing.*

The results we have just described apply for the comparison of condition doing writing with condition observation focus weak models. But do they also apply for the other observational-learning condition? Is it possible to generalize the results? Table 4.3 presents a comparison in orchestration of writing processes between condition observation focus good models and the doing-writing condition. When the comparison with doing writing differs from the comparison of observation focus weak models with doing writing, an asterisk is presented in the table.

Table 4.3 shows that there are only six 'unique' differences from condition observation focus good models with doing writing. This occurrence is so rare, that we may generalize the previously presented results to both our observational-learning conditions.

Results showed that we may largely confirm our second hypothesis: writers in both observational-learning conditions performed relatively more metacognitive activities (Goal-orientation and Analysis) in the beginning and relatively more executional activities (Writing and Re-reading) in the second part of the writing process than writers in doing-writing condition. Over the whole writing process, writers in both observational-learning conditions showed more Planning activities than writers in doing writing. Writers in the control condition performed relatively more executional (Writing) activities at the start and more Formulating activities during the whole writing process than writers in the observational-learning conditions.

Table 4.3. Pattern for orchestration of cognitive activities in condition observation focus good models related to doing writing. An asterisk is presented when the comparison to doing writing differs from the comparison of observation focus weak models with doing writing

	Observation focus good models		
Cognitive activity	Start	Middle	End
Goal-orientation	+		+
Planning	+	+	+
Analysis	+	+	
Evaluation	*	*	*
Meta-analysis	*	+*	+
Formulating	–*	–	–
Writing	–	+	
Pausing			
Re-reading		+	+
Revision			

3.3 Individual patterns

Overall effect sizes might not be very informative because differences between conditions depend on the moment in the writing process. So, overall effect sizes will be an underestimate of the differences between the three conditions. Instead we will focus on changes in the writing process of individual writers (compare Equation (2) and Equation (3)). This way it can be evaluated as well whether differences between conditions are a result of only some participants or a general change in the writing process in a specific condition.

For Formulating and Re-reading the between-subject variance proved to be equal across conditions; the differences between subjects are the same in the three writing conditions (see Appendix A for random parameters estimates).

For Goal-orientation, Planning, Analysis, Evaluation and Writing the between-subject variance changes due to the condition the subjects participated in. That is, in at least one of the conditions these writing processes deviate more from the general pattern. For Goal-orientation, for instance, the intercept variance is not significant in the focus-good-models condition. Apparently, focus-good-models condition urges writers to engage in Goal-orienting activities only in the beginning of the writing process (it has to be noted that the raise in the mean curve for Goal-orientation at the end of the writing process is due to only one participant). Writers in the control condition and in focus-weak-models condition show a deviation in amount of Goal-orienting activities (note that the pattern is the same because there is only variance for the intercept estimates). Some writers show a high amount of Goal-orienting activities in the start and middle of the writing process and other writers show a lesser amount.

For Meta-analysis the between-subject variance failed to reach significance in all three conditions. This lack of differences between writers can be explained in at least three ways. First, there might not be any meaningful difference between writers in pattern of Meta-analysis. Second, our measurement of this activity might be incapable of detecting differences between writers. The third and most plausible possibility is that Meta-analysis occurs more or less at random during the writing process.

In Figure 4.4 the individual differences in patterns for Evaluation and Writing are illustrated. In this figure each line represents one writer, except of course if the patterns of two or more writers coincide.

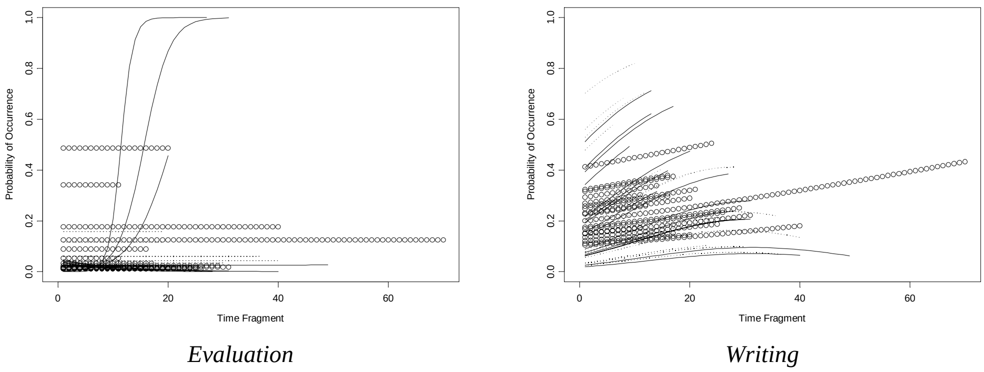


Figure 4.4. Orchestration of cognitive activities per individual writer
Note. 'oo' stands for individual writers in condition doing writing. '–'stands for individual writers in condition observation focus weak models; and '…'stands for individual writers in condition observation focus good models.

For Evaluation three writers in the focus-weak-models condition show a deviant pattern. These writers show a strong increase of Evaluating activities in the middle of the writing process. The writers in the other two conditions do not deviate from the general pattern, the variance in these conditions is not statistically significant. However, the deviation of some writers in condition focus weak models is not large enough to change the general condition pattern (see the flat line in Figure 4.3).

For Writing, in the doing-writing condition the differences between writers proved to be negligible. In both observational-learning conditions, however, the writers show a deviation in amount of the general condition pattern. There are writers who show a large amount of Writing activities at the start of the writing process, but there are also writers in these conditions who show a small amount. Thus, it seems that the deviating pattern for the observational-learning conditions, compared to the control condition we found for Writing, is supported by some individual writers.

The number of writers responsible for differences in either mean pattern or differences in variance differs from process to process. For example, for Evaluation we saw that the vast majority showed the same curve, but for Writing we saw that in both observational-learning conditions only a few individuals were responsible for the difference in orchestration with doing writing. Thus, concerning individual patterns we might conclude that observational learning affects the orchestration of some writers instead of a whole population as a simple comparison of means suggests.

4 DISCUSSION

The purpose of this study was to demonstrate that the effects of observational learning on learning products can be attributed to effects on task processes. We focused on the domain of writing and on the effects of observational learning on orchestration of writing processes. Two hypotheses were examined: (a) observational learning affects the orchestration of writing processes differently than 'learning-by-doing' and (b) the orchestration will be affected in a specific pattern: observational learning will result in a pattern with more metacognitive activities such as Goal-orientation, Planning, and Analysis in the beginning and more executional activities such as Formulating, Writing, and Re-reading in the second part of the writing process. In contrast, 'learning-by-doing' will result in a pattern with more executional activities during the whole writing process but especially at the start.

We reported an experiment in which participants were assigned to one of three different conditions: two observational-learning conditions in which participants observed writers at work, or a control condition in which participants performed writing tasks themselves. To measure their orchestration of writing processes after the instruction session, the participants performed posttest writing tasks under think-aloud conditions.

Results largely confirmed both our hypotheses. First, comparison of model fit showed that observational learning affected the orchestration of writing processes differently than 'learning-by-doing'. Second, results showed that the orchestration largely differed in the expected pattern. Compared to writers who learned by doing, writers who learned by observation performed relatively more metacognitive activities (Goal-orientation and Analysis) in the beginning and relatively more executional activities (Writing and Re-reading) in the second part of the writing process. Over the whole writing process, writers who learned by observation showed more Planning activities than writers who learned by doing. Moreover, in the middle and last part of the writing process, writers who learned by observation performed increasingly more Meta-analyzing activities, indicating monitoring and regulating processes, than writers who learned by doing. On the other hand, writers who learned by doing performed relatively more executional activities (Writing) at the start and more Formulating activities during the whole writing process compared to writers who learned by observation.

Thus, when we compare task processes resulting from observation with task processes resulting from executional instruction, we observe important differences. Through moving cognitive effort from writing to learning to write, and practicing metacognitive strategies, students in an observational-learning environment may build a richer knowledge base which influences the orchestration of writing processes. Furthermore, it seems very plausible that the observers adopted the orchestration examples the models provided and fitted these in their own knowledge base. Note that these differences have to be the result of the type of learning task, and cannot be attributed to the learning content: in all conditions, exactly the same theory in the same sequence was presented.

Writers who learned by doing showed a writing process in posttest tasks that is monotonous and homogeneous. The orchestration for the activities Analysis, Meta-analysis and Re-reading is monotonous. Whereas writers who learned by observation showed for these activities a changing curve over time, writers in doing-writing condition performed these activities at a constant rate during the writing process. The homogeneous behavior of writers who learned by doing is indicated by the relatively lower variance in doing-writing compared to both observational-learning conditions. It seems that writers in the doing-writing condition showed a 'basic routine orchestration' which could be characterized by the sequence of Formulating – Writing – Formulating – Writing etc, without Planning, Analyzing, Meta-analyzing and Re-reading activities. The Formulating component in this 'routine orchestration' could be interpreted as a form of textual, sentence generation planning, which differs from the 'higher order planning' as planning a new step in the writing process. For instance, the protocol of Dennis (one of the writers in doing-writing condition) shows this basic orchestration of Formulating – Writing – Formulating – Writing:

> 'Being a road user becomes more dangerous?' [Reading sentence in assignment]. 'Mmm, strange sentence' [Evaluation]. 'Being a road uses becomes more dangerous' [Writing]. 'Cyclists do not comply with the rules' [Formulating]. 'Cyclists do not comply with the rules' [Writing]. 'Cyclists don't extend their hands' [Formulating]. 'Cyclists don't extend their hands' [Writing]. 'Cyclists don't pay attention to the red lights' [Formulating]. 'Cyclists don't pay attention to the red lights' [Writing]. […].

The protocol of Dennis consists largely of his saying aloud the words before and during writing them. He jumps straightaway into producing text with apparently little difficulty (compare McCutchen, 2000, p. 20). His protocol reveals the 'knowledge-telling strategy' of Bereiter and Scardamalia (1987). It seems that the writers in our study who learned by doing, performed those text producing processes.

In contrast, protocols of the writers in the observational-learning conditions in our study showed the hard work (the wrestle with ideas and language) McCutchen (2000, p. 20) refers to: more Analyzing, Planning, and Meta-analyzing activities. Moreover, the writers in these conditions show a different way of planning: less Formulating activities as a form of planning, but more 'higher level planning' activities. See for example, a part from the protocol from Jelle, a writer from condition observation focus weak models:

> 'First the standpoint' [Planning]. 'Being a road user becomes more dangerous' [Analysis]. 'Being a road user becomes more dangerous' [Writing]. 'Ehm' [Pausing]. 'Then the argument [Planning]. 'Cyclists do not comply with the rules' [Analysis]. 'And then two supporting arguments for that' [Planning]. 'Cyclists don't extend their hands and cyclists don't pay attention to the red lights' [Analysis]. 'And then the second argument' [Planning]. 'Drivers are acting as if they own the streets' [Analysis]. 'Support for that' [Planning]. 'They park on footpaths and cycle paths' [Analysis]. 'I think, first to start with the first argument' [Planning]. 'Cyclists do not comply with the rules' [Writing]. 'Ehm' [Pausing]. 'They don't extend their hands and they don't pay attention to the red lights' [Writing]. 'Then thus the standpoint' [Planning] […].

Jelle's protocol shows that he has an internal plan that is connected with the macro-structure of the text. He knows the structure and the ingredients of the text type, and thus he can plan and then analyze every sub process of the task. After that, he starts writing the actual text, but here also in parts, planning every new step. This is a different approach than Dennis from doing-writing condition showed. Writers who learned by observation are able to divide the writing task in sub tasks by planning each sub process.

Another observation we mentioned when comparing the data from the observation conditions to the doing-writing condition, was the difference in variances. Within the doing-writing condition, variances were low, which indicate that writers showed same patterns of processes, which were, as we saw above, rather monotonous. However, the variances in both observational-learning conditions were larger, indicating more heterogeneous processes. We interpret these differences as differences between routine processes and learning processes. Almost all writers from the doing-writing condition just followed their routine paths when fulfilling tasks, while (some) writers from the observational-learning conditions were trying to implement new patterns. Now we found that these effects of observational learning seem to be more dependent on some writers than on a general change, the next step is to examine what these writers characterize. It might be that model-observer similarity has played a role. Braaksma et al. (in press-b) showed the importance of model-observer similarity in competence when confronted with new tasks. Weak participants wrote better texts when they observed pairs of models and focused on the weak model than when they focused on the good model. For good participants the reverse was observed: they wrote better texts when they focused on the good model than when they focused on the weak model. When this effect not only plays a role on writing products, but also on writing processes, then we would expect that weak writers will change their processes when focusing on weak models, while good writers change when focusing on good models. This assumption would imply that the deviating patterns in the condition observation focus weak models come from weak writers – they were learning – while the deviating patterns in the focus-good-models condition come from good writers.

We implemented in this study two observational-learning conditions. Both conditions showed the same deviating orchestration compared to 'learning-by-doing'. However, this does not mean that writers in observation-focus-weak-models condition showed exactly the same orchestration as writers in condition observation focus good models. Although almost all cognitive activities show the same curve, sometimes the amount of activities is larger in focus-weak-models condition (Planning and Meta-analysis) and sometimes the amount is larger in focus-good-models condition (Goal-orientation, Analysis, Evaluation and Re-reading). When we focus on the kind of the planning activities we previously discussed, we see that writers in focus-weak-models condition show more process-oriented planning activities than writers in focus-good-models condition. Writers in focus-good-model condition show more Analysis-activities, indicating a more 'text-structure-oriented' kind of planning. Furthermore, when we look at the variances in both conditions, we find that in general writers in condition observation focus weak models show a more homogenous group than writers in condition observation focus good models. Maybe writers in focus-weak-models condition show faster a new stable task behavior than writers in focus-good-models condition. But this is only an assumption that should be examined in a more detailed study about learning behavior during different forms of observational learning.

A restriction of this study concerns the relation between our theoretical framework and the writing tasks from which the orchestration of writing processes are obtained. The writing tasks in our study are short and structured, while our theoretical framework is based on research on writing processes which are obtained by larger, less structured writing tasks (more 'essay-like'). For instance, Hayes and Flower (1980) instructed their writers to write expository essays, writers in the study of Ferrari et al. (1998) wrote a comparative text (comparing two cities), and writers in the Breetvelt et al. (1994) study wrote two argumentative texts ('Living alone: yes or no?' and 'Children, yes or no?'). For further studies on the effects of observational learning on orchestration of writing processes, we thus recommend other, less structured and larger writing tasks.

Furthermore, for reasons of generalization, it would also be advisable in future studies to cross the 'subject-borders', and to examine for instance the effects of observational learning on the problem solving processes of students who are solving math problems. From studies of Schunk and his colleagues (Schunk & Hanson, 1985, 1989a, 1989b; Schunk, Hanson & Cox, 1987), we know that observational learning effects the math skill (fractions) positively, but nothing is known on the effects on the processes during the solution of fraction problems. It is possible that observational learning causes a richer knowledge of 'task structures' (compare richer knowledge of 'macrostructures' for text) that guides the student through the task execution process.

Now that we know that observational learning affects tasks processes differently, the next step should be to study the relation between knowledge and processes. We assumed that students in an observational learning environment build a richer knowledge base which influences the orchestration of writing processes differently. This assumption should be examined by comparing the extent and quality of the knowledge base from students who learned by observation with the knowledge base of students who learned by doing.

APPENDIX A

Parameter estimates by cognitive activity (in logit scores), indicating change in occurrence of processes during the writing process. Parameters which are not statistically significant (p > .05) are indicated with the addition ^{ns}. The other parameters are statistically significant (p <. 05).

Metacognitive activities

Parameters	DW	WM	GM	Parameters	DW	WM	GM
Random part	GOAL-ORIENTATION			Fixed part	GOAL-ORIENTATION		
S^2t^0	0.30	0.28	0.05^{ns}	$\beta_0 t^0$	-3.17	-4.78	-5.00
$S^2t^0t^1$	-	-	-	$\beta_1 t^1$	-1.26	-2.38	-2.19
S^2t^1	-	-	-	$\beta_2 t^2$	0.21	0.86	1.41
Random part	PLANNING			Fixed part	PLANNING		
S^2t^0	0.16^{ns}	1.70	2.04	$\beta_0 t^0$	-3.78	-3.00	-3.52
$S^2t^0t^1$	-	-	-	$\beta_1 t^1$	-	-	-
S^2t^1	-	-	-	$\beta_2 t^2$	-	-	-
Random part	ANALYSIS			Fixed part	ANALYSIS		
S^2t^0	0^{ns}	2.13^{ns}	7.38	$\beta_0 t^0$	-5.80	-4.05	-3.45
$S^2t^0t^1$	-	0.67^{ns}	2.59^{ns}	$\beta_1 t^1$	-	-0.71^{ns}	-0.71^{ns}
S^2t^1	-	3.14^{ns}	1.22^{ns}	$\beta_2 t^2$	-	-	-
Random part	EVALUATION			Fixed part	EVALUATION		
S^2t^0	0.91^{ns}	1.13^{ns}	0.31^{ns}	$\beta_0 t^0$	-3.29	-3.84	-3.17
$S^2t^0t^1$	-	2.08^{ns}	-	$\beta_1 t^1$	-	-	-
S^2t^1	-	3.50	-	$\beta_2 t^2$	-	-	-
Random part	META-ANALYSIS			Fixed part	META-ANALYSIS		
S^2t^0	0.18^{ns}	0.18^{ns}	0.18^{ns}	$\beta_0 t^0$	-2.84	-3.05	-2.70
$S^2t^0t^1$	-	-	-	$\beta_1 t^1$	-	0.53	0.12
S^2t^1	-	-	-	$\beta_2 t^2$	-	-	-

Executional activities

Parameters	DW	WM	GM	Parameters	DW	WM	GM
Random part	FORMULATING			Fixed part	FORMULATING		
S^2t^0	0.31	0.31	0.31	$\beta_0 t^0$	-1.35	-1.44	-1.62
$S^2t^0t^1$	-	-	-	$\beta_1 t^1$	-	-0.23	-
S^2t^1	-	-	-	$\beta_2 t^2$	-	-	-
Random part	WRITING			Fixed part	WRITING		
S^2t^0	0.14^{ns}	0.44	0.60	$\beta_0 t^0$	-1.16	-1.02	-0.88
$S^2t^0t^1$	-	-	-	$\beta_1 t^1$	0.16	0.54	0.49
S^2t^1	-	-	-	$\beta_2 t^2$	-	-0.15	-0.16
Random part	PAUSING			Fixed part	PAUSING		
S^2t^0	0.25	0.25	0.25	$\beta_0 t^0$	-1.91	-1.91	-1.91
$S^2t^0t^1$	-	-	-	$\beta_1 t^1$	-	-	-
S^2t^1	-	-	-	$\beta_2 t^2$	-	-	-
Random part	RE-READING			Fixed part	RE-READING		
S^2t^0	0.69	0.69	0.69	$\beta_0 t^0$	-4.41	-3.47	-3.27
$S^2t^0t^1$	-	-	-	$\beta_1 t^1$	-	0.45	1.00
S^2t^1	-	-	-	$\beta_2 t^2$	-	-	-
Random part	REVISION			Fixed part	REVISION		
S^2t^0	0^{ns}	0^{ns}	0^{ns}	$\beta_0 t^0$	-5.40	-5.79	-4.96
$S^2t^0t^1$	-	-	-	$\beta_1 t^1$	-	-	-
S^2t^1	-	-	-	$\beta_2 t^2$	-	-	-

Note. 'DW' stands for condition doing writing; 'WM' stands for condition observation focus weak models, and 'GM' stands for condition observation focus good models. S^2t^0 = between-student variance in process in the intercept; $S^2t^0t^1$ = covariance between intercept and linear change at student level; S^2t^1 = variance in linear process frequency; $\beta_0 t^0$ = intercept; $\beta_1 t^1$ = mean linear change in process frequency, and $\beta_2 t^2$ = quadratic change in process frequency. All explanatory variables are centered around the mean

CHAPTER 5

ACTUAL BEHAVIOR IN OBSERVATION TASKS

A case study

Abstract

In many studies, observational learning has proved to be effective with learners of various ages and in various school subjects, including writing. However, little is known about the actual behavior of learners while carrying out observations. Therefore, in this study, students' behavior when processing observation tasks is analyzed: six students thought aloud while observing writers and were interviewed afterwards.

Results showed that observers carried out many (meta)cognitive activities, especially activities that are based on the internalization and development of criteria for effective writing (observing, comparing, evaluating, and reflecting activities). These are precisely the activities that theorists assume play a central role in learning to write. These activities seem to be stimulated naturally in our observation tasks, while they are not very obvious in the usual writing tasks and exercises.

1 INTRODUCTION

Observing is the key learning activity in learning environments in which learners learn from models. Observational learning occurs when observers display new behaviors that they do not demonstrate prior to modeling, despite motivational inducements to do so (Schunk, 1998). There is a large body of research about effects and conditions of observational learning and modeling (Bandura, 1986, 1997; Rosenthal & Zimmerman, 1978; Schunk, 1987, 1991, 1995, 1998). With teachers, adults or students as models, observational learning has proved to be effective with learners of various ages and in various school subjects, such as mathematics (Schunk & Hanson, 1985, 1989a, 1989b; Schunk, Hanson & Cox, 1987), reading (Couzijn, 1995, 1999), writing (Braaksma, Rijlaarsdam & Van den Bergh, in press-b; Couzijn, 1995, 1999; Graham & Harris, 1994; Graham, Harris & Troia, 1998; Schriver, 1992), and speaking and listening (Sonnenschein & Whitehurst, 1984).

Braaksma, M. A. H., Rijlaarsdam, G., Van den Bergh, H., & Van Hout-Wolters, B. H. A. M. (submitted). *Actual behavior in observation tasks. A case study.*

1.1 The effectiveness of observational learning in written composition

In this article, we focus on observational learning in the domain of writing. According to Hayes (1996), the long-term memory of writers comprises a set of different kinds of knowledge that ensure different functions within the writing activity: knowledge of the type of text (genre knowledge), knowledge of the addressee (audience knowledge), of the linguistic components necessary for the realization of the text (linguistic knowledge), of the area of the content of the text (topic knowledge), and knowledge of the procedures to guide and control the effective realization of the text production (task schemas). It is assumed that the knowledge of writers is fed by regulatory or metacognitive strategies such as observation, evaluation, and reaction (Graham & Harris, 1994, p. 209). In the enactive feedback loop in which one learns from the consequences of one's actions, personal observations, judgments and reactions are indispensable (Zimmerman, 1989, 2000; Zimmerman & Schunk, 1989). Using metacognitive strategies, better writers gain information that changes what they know and do. They become aware of writing strategies they have used. They identify and conceptualize these strategies and their effects (they 'label' them), and add negative and/or positive evaluations of these effects. In other words, they create and specify criteria for effective texts and effective processes, an important feature of good writers (Hillocks, 1986).

This kind of learning requires 'that learners have reached developmental levels sufficient to be able to step back and consider their own cognitive processes as objects of thought and reflection' (Brown, 1987, p. 68). This is especially difficult in writing education. Learners have to attain two goals simultaneously: they must produce an adequate text and enlarge their knowledge about writing. One might say that they have to juggle with two tasks: a writing task and a learning-to-write task. Stepping back from the writing task to learn from it, is something very difficult and unnatural for learners, because all cognitive effort is directed to writing, the dominant task. Several instructional measures attempt to stimulate learners to step back and apply metacognitive strategies. A well-known approach is to add a phase of peer feedback and revision into the instructional process (Hillocks, 1986; Rijlaarsdam, 1987; Rijlaarsdam & Couzijn, 2000a). In general, this approach stimulates writers to step back, and to act as a reader and commenter of texts written by peers, assuming that commenting on other texts transfers to the revision phase of their own texts and to their next writing (Hillocks, 1986). Another approach to promote the use of metacognitive strategies and stepping back is the Self-Regulated Strategy Development approach (SRSD) that Graham and his colleagues developed (Graham & Harris, 1994; Graham, Harris, MacArthur & Schwartz, 1998; Graham, Harris & Troia, 1998). With SRSD students are supported to master higher level cognitive processes associated with successful writing, to promote reflective self-regulation of writing performances and to develop positive attitudes regarding the writing process and themselves as writers. In one of the seven stages of this instructional metascript the teacher models how to use a certain strategy, and learners observe, analyze, and discuss the effectiveness and efficiency of the modeled strategy.

The instructional method we focus on (observational learning) has at least one feature in common with the two approaches mentioned above: the key learning activities are observation, and the subsequent analysis and evaluation. However, the important difference is the *lack of writing*. In our approach learners do not write, but observe writing processes and the resulting texts. The cognitive effort is shifted from *executing* writing tasks to *learning* (Braaksma, Van den Bergh, Rijlaarsdam & Couzijn, 2001b; Couzijn, 1999; Rijlaarsdam & Couzijn, 2000a, 2000b). Because learners do not write, they may focus on the learning task, providing themselves with a learning opportunity to acquire new understanding about writing.

Observational learning directs learners' attention to writing processes, and students are stimulated to use and address metacognitive strategies explicitly, because the observation of the performance of others involves a 'natural' step back and thus a natural type of monitoring, evaluation, and reflection on task execution processes (Rijlaarsdam & Couzijn, 2000a). Moreover, by its very nature, observational learning focuses on the internalization and development of criteria.

1.2 Processes in observation tasks

Observational learning is governed by four constituent processes: attention, retention, production, and motivation (Bandura, 1986). *Attention* to relevant events in the learning environment is necessary for the events to be meaningfully perceived. *Retention* requires coding and transforming the modeled information in memory, as well as cognitively organizing and rehearsing information. *Production* involves translating mental conceptions of modeled events into actual behaviors. *Motivation* influences observational learning because if students believe that models demonstrate useful or interesting behavior they are likely to attend to such models closely and to retain what they learn. Bandura (1986, p. 51-70; see also Schunk, 1991, p. 106-108) provides a detailed description of factors that operate within these processes and that influence observational learning.

However, Bandura was focused on natural conditions in which observational learning occurred. Schunk (1998; Schunk & Hanson, 1985, 1989a, 1989b; Schunk, Hanson & Cox, 1987) focused on the educational setting of observational learning and examined the effects of different factors (number of models, model sex, model competence) on observers' learning. Little is known about the actual behavior of learners induced by observing in an educational context. This study aims to fill in this gap. Its purpose is to contribute to the theoretical background of observational learning in the sense that we are concerned with *how students process observation tasks*. By analyzing students' activities when performing observation tasks, we will gain detailed information about how students actual process observation tasks. In doing so, we focus on three main questions.

First, we will search for indications that support or refute our *theoretical assumptions* about the effectiveness of observational learning. We will investigate whether and how observers use metacognitive strategies, internalize and develop criteria for effective writing, and pay attention to learning to write. Special attention is paid to *evaluation and elaboration processes* where the use of metacognitive processes and the internalization and application of criteria are stimulated. In an earlier study by Braaksma et al. (2001b), it was found that evaluation and elaboration activities are important for the effectiveness of observational learning. Students who correctly evaluated peer models and commented correctly on the products the models produced were better in writing argumentative texts. These results indicate that a more detailed look at the evaluation and elaboration processes might be worthwhile.

Second, we would like to know whether *different instructions* for observations influence the behavior of observers. A previous study (Braaksma et al., in press-b) about the effects of observational learning in writing showed that weak learners learned more from focusing their observations on weak models, while better learners learned more from focusing on good models. Thus, the instructions worked out differently. There must be something that induced the specific learning behavior.

1.3 The present study

We set up a case study in which students thought aloud while observing writers as models. To explore whether different instructions result in different activities, two observational-learning conditions were implemented: weak focus and good focus.

Participants in both conditions followed a short course on argumentative writing. They read theory about writing argumentative texts and then applied the theory in observation tasks, observing pairs of peer-writers. Performing these tasks, participants focused respectively on the weak model (condition weak focus) or on the good model (condition good focus). Afterwards, the observers were interviewed about the way they had carried out the observation tasks.

2 METHOD

2.1 Participants

Six students (ninth grade, higher streams) of three secondary schools in the Netherlands took part in the study. They participated voluntarily and received a small financial reward. Participants were semi-randomly assigned to the conditions: we selected two girls and one boy for each condition, each from a different school. Average age of the participants was 14.8 years.

2.2 Procedure and materials

Participation in the study took place in an individual session during approximately two hours. In this session one participant and the researcher were present. The session consisted of three major parts: (a) an introduction of the study, (b) a writing course with observation tasks, and (c) an interview.

2.2.1 Introduction of the study

In the first part (25 minutes) the participants received a written explanation about the aim and method of the study:

> '[…] To find out how our materials work, we want to know what students are thinking when they are reading theory and performing exercises. Therefore, we ask you to think aloud. Everything you say will be audio taped in order to know how you worked. We want to know in which manner you read the theory and how you approached the exercises. […].'

It was emphasized to the participants that they stayed anonymous. Next, they received an instruction for the thinking aloud procedure which was based on Ericsson and Simon (1993). The instruction to think aloud was as follows:

> 'We want you to say aloud *everything* you think. Thus, say everything you *read*, *think*, and *write* aloud. Give as much information about what is going on in your head.'

After reading the instruction, the participants listened to an audio tape with a good example of a student who is thinking aloud while writing an essay. After that, students trained by trying to solve an 'inverse' crossword puzzle (filling in the definitions), while thinking aloud. As they carried out this task, the participants were audio taped and afterwards they listened to their own performance.

2.2.2 Writing course with observation tasks

The second part of the session (65 minutes), took place under think aloud conditions. The participants followed a course in an aspect of writing argumentative texts, namely how to transform argumentation structures into short linear argumentative texts. This learning task was completely new for all participants. The course was part of Couzijn's lesson series (1995, 1999) who based his materials on the pragma-dialectic perspective on argumentation developed by Van Eemeren and Grootendorst (1992). Materials for the course consisted of a workbook and a CD-ROM, displayed on a computer monitor. Participants received instructions by means of on-screen messages, e.g., reading theory in the workbook, answering a mastery question, or observing models. An on-screen timer indicated how much time was left for each activity; participants were alerted by short beeps when time was almost over.

The participants' workbooks consisted of four main sequences. In the first sequence, participants had to study a theoretical section about some main concepts of argumentation (standpoint, argument, singular and compound argumentation, and argumentative connectives). To stimulate active reading of the theory, mastery questions were provided. In the second sequence, participants applied the theory in four observation tasks (see Figure 5.1 for an example). In these tasks, after taking notice of the writing assignment for the models, participants were informed of two questions they had to answer after observing two models. Then, participants observed pairs of videotaped peer models executing the same writing task: writing short argumentative texts based on argumentation structures. These models were thinking aloud while they were writing. One of the models was a relatively good model (correct writing product), the other was a relatively weak model (incorrect writing product). See Appendix A for an overview of performances of the models (writing processes and resulting products).

Participants in both conditions observed the same pairs of models but their focus was different due to the instruction. Participants in weak-focus condition were asked to focus on the *weak* model, and participants in good-focus condition on the *good* model. The participants were stimulated to make notes while observing the models, as these might help them answering the two questions they had to answer after observing the two models. After the observations, the participants received the texts produced by the models and were asked to evaluate the observed performances. Next, they were asked to elaborate on their evaluation. The evaluation question was: 'Which model performed better?' In the weak-focus condition, the elaboration question was: 'Explain briefly what the other (less good) model did less good or worse'; in the good-focus condition this question was: 'Explain briefly what this (good) model did well.'

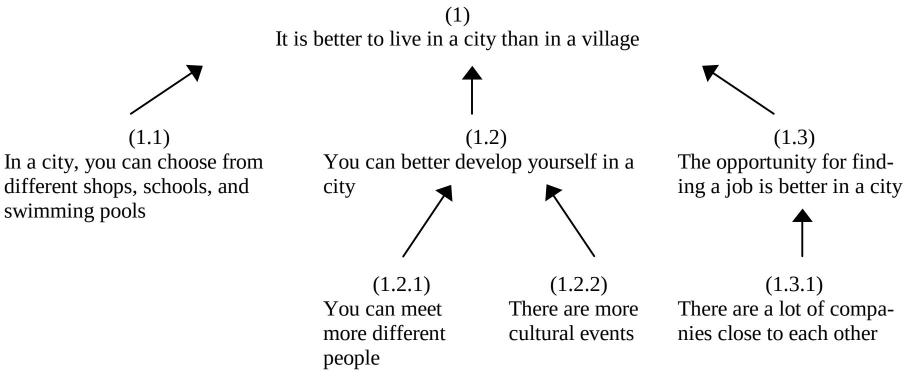
After completing the observation tasks, the participants went on to the third sequence, which contained theory and mastery questions about subordinate argumentation, argumentation structures and argumentative connectives. Finally, in the fourth sequence, they performed a fifth observation task related to this part of theory.

The use of videotaped models on CD-ROM ensured standardized presentation across participants. The five observation tasks ranged from relatively easy to more difficult tasks, starting with singular argumentation (task 1) via compound argumentation to both compound and subordinate argumentation in task 5.

To clarify the content of the observation tasks and the behavior of the models, we will explain in more detail one of the observation tasks. Figure 5.1 presents observation task 5, the most complex task.

The task starts with presenting the observers the argumentation structure the models received. The models had to transform this structure into 'linear' text. Next, the questions the observers had to answer after observing the models are presented. Then, the observers started observing the models. They watched the video recording of models' writing. Remember that the models were thinking aloud during their writing. The models' writing processes can be found in Appendix B. Table B1 in Appendix B provides an overview of the writing process of the first model (a boy). His writing process consisted of various (meta)cognitive activities, planning, analyzing, formulating, evaluating, writing, and re-reading his text. From his writing, the observers could infer that the performance of this model is not entirely correct. Although he correctly identified standpoint, arguments, and subordinate arguments, he repeated the arguments to connect these with the subordinate arguments, and at the end of his text, he repeated the standpoint. This tactic is not to be preferred, making the text unnecessary long and complex. Then, the observers watched the performance of the second model (a girl). Table B2 in Appendix B provides an overview of her writing process. Like model 1, the second model also showed various (meta)cognitive activities (planning, analyzing, formulating, evaluating, writing, and re-reading). However, the performance of model 2 is much better, because she connected the arguments directly with the subordinate arguments. Next, after the models had been shown, the task presents the models' written texts. From these texts, the observers could observe again the performances of both models, but now in the form of written products. Finally, the task presents the evaluation and elaboration question to the observers.

In a little while, you are going to watch on CD-ROM videotape recordings. You will see two models writing a short argumentative text based on the following argumentation structure. The models had to make sure that the reader will understand the standpoint and arguments in their text. Below, you will find the argumentation structure the models received.



After watching the models, you have to answer the following questions:

1. Which model performed better?
2. Explain briefly what the other (less good) model did less good or worse

When you have observed both models, you may advance to the next page.
Make your notes here, when you observe the models:
....……………..

(…)

(next page booklet)

You saw two models doing the assignment. They wrote the following texts:

Model 1
It is better to live in a city than in a village because in a city you can choose from different shops, schools, and swimming pools, you can better develop yourself in a city and the opportunity for finding a job is better in a city. Because you can meet more different people and there are more cultural events, you can better develop yourself. The opportunity for finding a job is better because there are a lot of companies close to each other. Therefore it is better to live in a city.

Model 2
It is better to live in a city than in a village because firstly in a city you can choose from different shops, schools, and swimming pools. Secondly you can better develop yourself in a city because you can meet more different people and there are more cultural events. Moreover, the opportunity for finding a job is better in a city, as there are a lot of companies close to each other.

1. Which model performed better?
2. Explain briefly what the other (less good) model did less good or worse

………………………………………………………………….

(…)

Figure 5.1. Observation task (task 5) for participants in condition weak focus

2.2.3 Interview

During the last 15 minutes of the session, the participants were interviewed about how they performed the observation tasks. The purpose of the interview was to gain inside background information to be used in the analysis and interpretation of the think-aloud protocols. The interview consisted of four open-ended questions. If an answer was unclear, the interviewer asked for an explanation. The questions were about the observer's behavior when (s)he took notice of the argumentation structure the models received, the observer's attention during the observation of the models, whether the observer took the questions into account (s)he had to answer after observing the models, and the way of answering these questions.

2.3 Data sources and analyses

First, we made detailed descriptions of each student's activities during observation tasks. Multiple data sources were used: (a) think-aloud protocols from participants performing observation tasks (in total 30 protocols: six participants who executed each five tasks), (b) participants' workbooks containing notes and answers on evaluation and elaboration questions, and (c) observations of the participants' behavior by the researcher (for instance, 'participant shakes his head when observing the first model').

Think-aloud protocols yield meaningful information about the internal structure of cognitive processes (Smagorinsky, 1989, p. 465). The addition of the other sources of data, especially the notes in the participants' workbooks, were essential because of the complexity of both thinking aloud and observing two writers who are thinking aloud.

To structure the descriptions, the observation tasks are divided in three phases: preparatory phase (taking notice of the argumentation structure the peer models received), observing phase (observing two peer models writing argumentative texts), and post-observing phase (taking notice of the texts the peer models wrote and answering the evaluation and elaboration questions).

Second, the participants' answers to the interview questions about the way they had carried out the observation tasks were an additional source of data for the analyses for the three main questions (indications in support of theoretical assumptions, evaluation and elaboration processes, and influences of different instructions).

3 RESULTS

We focus first on the description of observers' behavior. Afterwards, we will come back to the three questions: (a) indications in support of theoretical assumptions, (b) evaluation and elaboration processes, and (c) influences of different instructions.

3.1 Processing observation tasks

To provide a detailed picture of students' observation processes, we first follow one of the participants more closely during the performance of one of the observation tasks. Next, we compare her behavior to that of the other participants.

3.1.1 Claudia processing observation tasks

Claudia[6] is a small, blond-haired, fourteen-year-old girl that comes from a school in a large city in the Netherlands (ninth grade, high level stream). She has been assigned to condition weak focus. She was enthusiastic and performed her task very seriously. Often, she re-read parts of the theory to be sure. Sometimes she formulated parts of the theory in her own words. We will follow Claudia during her performance of observation task 5, the most complex task (see Figure 5.1).

In the *preparatory phase,* Claudia took notice of the argumentation structure the models received. First, she analyzed the structure by reading the elements of the structure and directly naming their argumentative functions. For reasons of clarity, reading and formulating activities in protocol fragments are printed in italics, and other activities (e.g., planning and analyzing) in regular font.

> '*It is better to live in a city than in a village*, and then you have here *in a city you can choose from different shops, schools and swimming pools*, according to me this is singular argumentation, and then *the opportunity for finding a job is better in the city, there are a lot of companies close to each other*, that is subordinate […]'.

This fragment shows that she did not first read the structure in the form of a structure and only then analyzed the structure, but that she combined the two activities.

After analyzing the structure, she (mentally) formulated (parts of) the text by planning the argumentative connectives between some elements, and then formulating the sentence:

> '[…] *You can meet more different people*, there you can place *and: and you meet more different people and there are more cultural events* […]'.

By performing these analyzing and formulating activities, Claudia showed that she is not a passive observer but instead actively formed a mental image of a possible task outcome solution before she is confronted with the models' solutions.

In the *observing phase*, Claudia observed two models on video. During this observation, we observed from the notes in her textbook that she paid attention to both models. She compared the two models and commented on the models' argumentation (i.e., an analysis of the argumentation in the models' writing) and on the models' approach of the task. She also judged the performance. She wrote:

> 'Model 1 does not take after the argument the two subordinate arguments, and takes not after the third argument the subordinate argument but it is still logical and it is correct. He writes a long time everything a couple of times. Model 2 does it correctly and everything in one time.'

---

6 Claudia is not her real name. For reasons of anonymity, the names of the participants in this article are fictitious.

Here, Claudia identified the different arguments and subordinate arguments in the writing of model 1. She connected the elements in the model's writing with the elements in the argumentation structure. By contrasting the performance of this model with the ideal performance according to the structure (and which would reflect Claudia's own performance?), and the word 'but', she indicated (implicitly) that this performance is not the most preferable. However, despite these remarks, she judged the performance of model 1 as correct and logical. Next, she judged the performance of model 2 as being correct, connecting the arguments directly with the subordinate arguments.

From these observations, it can be inferred that Claudia has criteria for a correct task performance at her disposal with which she can compare the two models, evaluate the performance of the models, and elaborate on that evaluation. Furthermore, the writing of the models invited Claudia to evaluate the performance of the models and to elaborate on that evaluation. The observation of the models' writing was enough for Claudia to comment on their performance. She did not need to read the results of these writing processes (the models' written texts) to judge their performance and to explain that judgment. This conclusion is supported by the activities of Claudia in the next phase.

In the *post-observing phase* (taking notice of the texts the models wrote and answering the evaluation and elaboration question) Claudia started *before* reading the texts with writing down her answers on the evaluation and elaboration question:

> 'I think that model 2 performed better because she, because model 1 he mentioned all sentences, no, some sentences, a couple of times. According to me, the ordering was not very logical and model 2 was clearer. Let me read both texts.'

These utterances demonstrate that Claudia did not need to read the produced texts to answer the questions. For her, observing the models at work was enough to answer the evaluation and elaboration question. Still (to check her answer?), she planned to read both texts. She commented on both models and used a 'general' description, using words as 'the ordering was not very logical', she did not analyze the argumentation in the models' texts as she did in the observing phase.

After reading both texts, she looked back to her answer on the elaboration question. She added:

> 'Although model 2 performed better, model 1 performed also well.'

Next, she read both texts again and thought once more about the performance of model 2 by saying:

> 'Maybe, it is not right, that she [model 2] says *firstly*, that that is more important than *secondly* but I still think, according to me model 2 is still more clear. Oh so!'

This (spoken) addition shows that Claudia was heavily involved with judging the performance of the models. Here, she meant that she doubted about model 2's use of 'firstly', and 'secondly'. She was not sure whether it is correct that argument 1 is more important than argument 2.

When we review Claudia's behavior during observation task 5, we may conclude that Claudia was involved in many different activities. Especially in the post-observing phase, she was very actively checking and improving her elaboration of the evaluation of the models' performance. Furthermore, the analysis of the argumentation structure, (mental) formulations of a possible task outcome solution, and evaluation and elaboration activities in all phases show that Claudia had a standard of good task performance which enabled her to compare, evaluate and explain the performance of the models.

3.1.2 All participants processing observation tasks

When we compare Claudia's behavior in observation task 5 with her performance in the other tasks, and with the behavior of the other participants, we find similarities and differences. We may establish a 'main pattern' of activities (i.e., a lot of filled boxes in the rows) and more 'individual' activities (i.e., some filled boxes in the rows) in the Tables 5.1, 5.2, and 5.4. These tables show for each participant his/her cognitive activities while performing five observation tasks.

In the *preparatory phase* (see Table 5.1), the participants mostly read the argumentation structure 'as structure' (i.e., 'only' reading the elements, without analyzing and formulating activities). Sometimes this reading is followed by an analysis of the structure and/or formulating the text. Robert (3) sometimes showed a different strategy: he did not read the structure but tended to start immediately formulating the text.

Table 5.1. Cognitive activities of participants Claudia (1), Yvonne (2), Robert (3) (all three in condition weak focus), Dorien (4), Karin (5), and Joost (6) (all three in condition good focus) during the preparatory phase

	Task 1		Task 2		Task 3		Task 4		Task 5	
	Participant number									
Cognitive activities in preparatory phase	123	456	123	456	123	456	123	456	123	456
When taking notice of the argumentation structure for the models:										
1 Reading the structure as structure	■■■	■■■	■■□	■■■	□■□	■■■	■■■	■■■	□■□	■■■
2 Analyzing the structure	□□□	□□□	□□□	□□□	■□□	□□□	□□□	□■□	■□□	□□□
3 Formulating the text*	□□□	□□□	■□□	□□□	■□□	□□□	■□□	□□□	■□□	□□□
4 Formulating the text while reading the structure	□□□	□□□	□□■	□□□	□□■	□□□	□□□	□□□	□□■	□□□

Note. Filled boxes represent that a cognitive activity is performed by the participant (■)*, and open boxes represent that a cognitive activity is not performed by the participant* (□)*. *Note. Participant number 1 often formulated two possible texts: first text with 'standpoint, because, argumentation' and second text with 'argumentation, thus, standpoint'.*

In the *observing phase* (see Table 5.2), the participants mostly paid attention to both models. During this observation, the participants often judged the performance of the models:

> 'Model 1: good, model 2: does not sound nice.'(Robert (3), task 3).

Frequently, the participants commented on the models' writing. Most often, they focused on the outcomes of the writing processes; occasionally, they focused on the approach of the models (see Table 5.2). Table 5.3 shows also the different objects they focused on, with some protocol examples.

Table 5.2. Cognitive activities of participants Claudia (1), Yvonne (2), Robert (3) (all three in condition weak focus), Dorien (4), Karin (5), and Joost (6) (all three in condition good focus) during the observing phase

	Task 1	Task 2	Task 3	Task 4	Task 5
	Participant number				
Cognitive activities in observing phase	123 456	123 456	123 456	123 456	123 456
When observing the models:					
1 Paying only attention to model who was mentioned in question*	□□□ □□□	□□□ □□■	□□□ ■□□	□□□ □□□	□□□ □□□
2 Paying attention to both models	■■■ ■□□	■□■ ■■□	□■■ □■□	■■■ ■■□	■■■ ■■□
3 Paying only attention to model who was *not* mentioned in question*	□□□ □■□	□■□ □□□	□□□ □□□	□□□ □□□	□□□ □□□
4 Giving a judgment	■■□ ■□□	■■□ □■□	□■■ □□□	■□■ □■□	■■■ □■□
5 Commenting on argumentation	□□■ ■■□	□□■ ■□■	□□■ □□□	■□□ ■□□	■□□ ■■□
6 Commenting on argumentative connectives	■■■ ■■□	■■■ ■□■	□□□ ■■□	■■■ ■■□	□■□ □□□
7 Commenting on language aspects	□□□ □□□	□□□ □■□	□■■ □□□	□□■ □■□	□□□ □□□
8 Commenting on approach	□□□ □□□	□□■ □□□	□□□ □□□	□□□ □□□	■□■ □■□
9 Comparing the models with own execution (explicit)	□□□ □□□	□□□ □□□	□□□ □□□	□□□ □□□	□□■ □■□

**Note.* *Activities 1 and 3 are not intentional: the participants focused on the first model displayed only*

Table 5.3. Different objects and examples of comments during observing phase

Object	Example	Participant/task
Argumentation	'Model 1: uses *because* and writes first the argument and then standpoint. Model 2: also uses *because* and writes first standpoint and then argument.'	Robert (3), 1
Connectives	'Model 2: argument, *and* + *because*.'	Dorien (4), 2
Language	'Model 2: absolutely not a nice sequence of sentences.'	Yvonne (2), 3
Approach	'The first model was busy improving his text.'	Robert (3), 5

During the observations of the models, Karin (5) and Robert (3) compared the performance of the models explicitly with their own performance. Robert (3), task 5:

> 'Model 1: starts with the standpoint, he likes that more. I also like it to start with the standpoint.'

In the *post-observing phase* (see Table 5.4), the participants mostly read the texts the models wrote. When they did not read the texts, they indicated that their process observation was enough to evaluate the performance of the models and to elaborate on that evaluation. Robert (3), task 5:

> ‘Oh, I know the answers already.’

When the participants answered the evaluation question, they sometimes commented on that evaluation. Joost (6), task 4:

> ‘Who performed better? Model 1, although he also performed not totally correct.’

When they had to answer the elaboration question, the participants sometimes focused only on the model that was mentioned in the question, sometimes they focused on both models. Occasionally, they focused only on the model who was *not* mentioned in the question. Thus, not all observers paid sole attention to the model on which they had to focus.

Answering the elaboration question, observers commented on different objects in the models’ text (see Tables 5.4 and 5.5).

Table 5.4. Cognitive activities of participants Claudia (1), Yvonne (2), Robert (3) (all three in condition weak focus), Dorien (4), Karin (5), and Joost (6) (all three in condition good focus) during the post-observing phase

	Task 1	Task 2	Task 3	Task 4	Task 5
	Participant number				
Cognitive activities in post observing phase	123 456	123 456	123 456	123 456	123 456
1 Writing answers before reading the texts	□□□ □□□	□□□ □□□	□□□ □□□	□□□ □□□	■□□ □□□
When taking notice of the texts written by the models:					
2 Reading both texts	■■□ ■■■	■■■ ■■■	■■■ ■■■	■■■ ■■■	■■□ □■■
3 Giving a judgment after reading the text(s)	■■■ □□□	□■□ □□□	□■□ □□□	■■□ □□□	□■□ □□□
4 Giving a revision proposal	□■□ □□□	□□□ □□□	□□□ □□□	□□□ □□□	□□□ □□□
5 Not reading one text or both texts	□□■ □□□	□□□ □□□	□□□ □□□	□□□ □□□	□□■ ■□□
When answering the evaluation question:					
6 Evaluating without comments	□■■ ■■■	■■■ ■■■	□■■ ■■■	□■■ ■■□	■■■ ■■□
7 Evaluating with comments	■□□ □□□	□□□ □□□	■□□ □□□	■□□ □□■	□□□ □□■
When answering the elaboration question (I):					
8 Paying only attention to model who was mentioned in question	■■■ □□□	■□■ □□□	■■■ □□□	■■□ □■□	□■□ □■□
9 Paying attention to both models*	□□□ ■□■	□■□ ■■■	□□□ ■■■	□□□ ■□□	■□■ ■□■
10 Paying only attention to model who was *not* mentioned in question*	□□□ □■□	□□□ □□□	□□□ □□□	□□■ □□■	□□□ □□□

Table 5.4. (Continued)

Cognitive activities in post observing phase	Task 1	Task 2	Task 3	Task 4	Task 5
	Participant number				
	123 456	123 456	123 456	123 456	123 456
When answering the elaboration question (II):					
11 Commenting on argumentation	■■■ ■□■	□□■ ■■■	□□■ ■■■	■■■ ■□■	□□□ ■■■
12 Commenting on argumentative connectives	■■□ ■■□	■■□ ■■□	■■□ ■□□	■■□ ■■■	□□□ □□□
13 Commenting on language aspects	□□■ □□□	□■□ □□□	■■■ □■■	□□□ □□□	■■■ □□□
14 Giving a revision proposal	■■□ ■■□	■□□ ■■□	■□□ □□□	□□□ □□□	□□□ □□□
15 Comparing answer of the models with own answer	□□□ □□□	□□□ □□□	□□□ □□□	■□□ □□□	□□□ □□□
16 Looking back in notes	□□□ □□□	□□■ □□□	□□■ □□□	□□□ □□□	□□■ □□□
17 Looking back in theory	■□□ □□□	□□□ □□□	□□□ □□□	□□□ □□□	□□□ □□□

**Note. Participant number 3 was conscious of his 'mistake'. After elaborating, he apologized for paying attention to the good model (task 4) and to both models (task 5).*

Sometimes, students commented in the form of a proposal of how to revise the text of the weaker model. Such a revision proposal was usually directed to the argumentative connectors:

> 'Model 1 should have written *thus* instead of *because*.' (Karin (5), task 1)

Table 5.5. Different objects and examples of comments when answering the elaboration question during post-observing phase

Object of elaboration	Example	Participant
Argumentation	'The text of model 2 is not correct because model 2 creates a second standpoint'	Joost (6), task 3
	'The text of model 1 is wrong because the argument is being used as standpoint'	Claudia (1), task 1
Connectives	'The text of model 1 is correct. Model 1 mentions the arguments with *namely*, *besides* and *and*'	Dorien (4), task 3
Language	'The sequence of the sentences is not correct in the text of model 2'	Yvonne (2), task 2
	'Model 1 changes a lot in the text, he writes *I think Frans is originally from Limburg* instead of *It is for sure that Frans is originally from Limburg*'	Claudia (1), task 3

In sum, with regards to how observers are processing observation tasks, we conclude that the observers were generally consistent in their observing processes. The processes were largely task content independent and observers' processes looked like each other. Students mostly read the argumentation structure 'as structure', which is sometimes followed by an analysis of the structure or by formulating the text. When they observed the models, they generally paid attention to both models and most times, they judged the models. Elaborations on the judgment were repeatedly focused on argumentation and argumentative connectives in the models' writing. In the third phase, the observers mostly read both texts and elaborated on their evaluation, frequently by commenting on the argumentation and the argumentative connectors the models used.

Now we have some insight into observers' processes, let us turn to the interpretation of these processes. Our interpretation is guided by the three main questions, introduced in the first section of this article, concerning: (a) indications in support of theoretical assumptions, (b) evaluation and elaboration processes, and (c) influences of different instructions.

3.2 Indications in support of theoretical assumptions

We assumed that observational learning could be effective because in observational learning students are stimulated to use metacognitive strategies, internalize and develop criteria for effective writing, and pay more attention to learning to write.

The overview and examples of the observers' behavior show that the observers were strongly involved in metacognitive activities. During the observing and post-observing phase, the participants *observed* the models' writing, *evaluated* the performance of the models and *reflected* explicitly on the observed performances. Thus, they identified and conceptualized the writing strategies of the models in different ways (e.g., commenting on argumentation, commenting on argumentative connectives) and added negative and/or positive evaluations of these strategies.

Furthermore, the observers' comparisons, evaluations and elaborations in observing and post-observing phase show that criteria for effective writing were internalized, applied, and developed. For evaluating the models' writing, the observers used criteria. These criteria were made explicit when observers elaborated on their evaluation. Remember for instance the elaboration from Claudia (1) in observation task 5 when she explained why the writing of model 1 is not as good as the writing of model 2. By arguing that model 1 did not directly connect the subordinate arguments after the arguments, she applied the criterion 'arguments and subordinate arguments are connected directly'. This criterion could have been provided by the performance of the good model. See also the elaboration of Robert (3) in observation task 2. He applied the criterion 'argument supports standpoint' that could have been derived from the performance of the better model:

> 'Model 2 is not right, the argument is supported by the standpoint instead of the other way.'

In the introduction of this article, we argued that observers do not have to juggle with a dual task because they do not write, and thus could focus on learning processes. However, the observers' behavior shows that observers also wrote! However, these processes differed from the executional processes the models performed. The observers did not actually write. They did not have to pay cognitive effort to the 'physical aspects of handwriting' or to spelling and grammar.

Instead, the observers placed their (mental) performance of the writing task in service of the observation task. They compared their (covert) performance with the (overt) performance of the models. This is demonstrated through observers' activities in all three phases of the observation task. During the *preparatory phase*, students analyzed the structure and formulated the text verbally. For instance, Claudia (1):

> 'I should say: *I don't like the last book of Roald Dahl, because I don't understand the story and it doesn't contain nice illustrations any longer*. Or *I don't understand the story thus*… all right.'

In the *observing phase*, some observers compared their own performance explicitly with the performance of the models. For instance, Karin (5):

> 'Model 1: good structure, but I would put the subordinate arguments immediately below (or after) the arguments.'

In the *post-observing phase*, Claudia (1) showed an explicit comparison of the models' performance with her own performance. After she answered the elaboration question, she formulated her own task outcome solution in two possibilities ('standpoint, because, argument' and 'argument, thus, standpoint') that are literally the same as she formulated in the preparatory phase. Then, she added:

> 'Yes, the performance of model 1 was better.'

Furthermore, from the interviews it appears that all six observers knew how the text should be written:

> 'Before I saw the video, I thought of what I would do. I already knew the answer, but of course there are more possibilities.' (Joost (6))

> 'Before I saw the video, I made the text in my head and compared it with the solutions of the models.' (Robert (3))

3.3 Evaluation and elaboration processes

In this section, we focus on evaluating and elaboration processes. Because it was shown that evaluation and elaboration activities contribute explicitly to the effectiveness of observational learning (Braaksma et al., 2001b), it is worthwhile to have a more detailed look at these processes.

Our data show that the observers started evaluating the models as soon as they started their observations. They judged (15 times) the performance of the model(s), for instance by writing 'correct' or 'incorrect' in their notes. Furthermore, the observers nodded or shook their heads (in approval or disapproval) when observing the model(s).

These judgments were very stable, as appears from a comparison of evaluations in the observing and the post-observing phase. The choice for a better model in the observing phase corresponded all 15 times with the choice for the better model in the evaluation question in the post-observing phase.

Furthermore, the observers' behavior shows that elaborations on the evaluations were already given in the observing phase. Table 5.6 shows the frequencies of different objects of these elaborations during observing phase (and post-observing phase).

Table 5.6. Frequencies of objects of elaboration during observing phase and post-observing phase and number of constant objects

Object of elaboration	Observing phase	Post-observing phase	Constant
Argumentation	12	21	10
Argumentative connectives	18	16	12
Language aspects	5	10	2
Approach	4	–	–

Table 5.6 shows that in the observing phase the participants commented most frequently on the argumentative connectives in the writing of the models, followed by commenting on argumentation. From the interviews with the participants, it also appears that they were focused on the argumentative connectives. On the question: 'To which aspects did you pay attention during the observation of the models?' Karin (5) answered:

> 'I watched carefully whether they used the word *because* or *thus* between standpoint and argumentation. I saw whether a model was good or less good when they used for instance *because* instead of *thus* and that is wrong and thus less good.'

Dorien (4) answered:

> 'I paid attention to *because*, *thus* and other connectives. I also paid attention to the order of standpoint and arguments, whether this was correct. If the first model was wrong, then the second model had to be correct but still I controlled that.'

After reading the models' written texts, in the post-observing phase, most comments were directed towards the argumentation in the text of the models, followed by commenting on argumentative connectives and language aspects (see Table 5.6). All comments in the post-observing phase were thus directed to aspects of the written texts. No comments on approach were made. This is in contrast with the observing phase where observers commented four times on the approach of the models. Apparently, the appearance of the written texts in the post-observing phase was so directing that no comments on approach were made.

The column 'constant' in Table 5.6 shows that the types of comments are more or less stable between the two phases, especially the category 'commenting on argumentative connectives'. From the 18 times that participants in the observing phase commented on argumentative connectives, they commented 12 times also in the post-observing phase on these connectives. Another indication that comments made in the observing phase are used in the post-observing phase is that many times comments made in the observing phase were literally repeated in the post-observing phase. Sometimes, participants also referred to their notes. Robert (3), for instance, referred to his notes three times, connecting his comments on the quality of the models' writing to the elaboration question ('Let's look what I have written in my notes').

Concerning the development of evaluation and elaboration processes it can be concluded that the participants already started evaluating and elaborating during the observation of the models and that these evaluations and elaborations are used for answering the questions in the post-observing phase.

Moreover, the participants indicated that they already knew the answers on the questions in the post-observing phase after they observed the models. Sometimes they skipped the reading of the models' texts (cf. the utterance of Robert (3): 'Oh, I know the answers already'). In the interviews, all six observers indicated that they knew the answers already. The written texts were used:

> 'To find an exact word.' (Joost (6))
>
> 'To re-read if you have forgotten a detail.' (Dorien (4))
>
> 'For sure, to check my answer.' (Karin (5))
>
> 'For sure, to look if I had forgotten anything.' (Robert (3))

3.4 Influences of instructions

In this last section of the results, we focus on influences of instructions on the participants' behavior. Because two observational learning conditions were implemented, (weak-focus condition and good-focus condition) we could explore whether observers who focused on weak models employed a different strategy than observers who focused on good models.

The Tables 5.1, 5.2, and 5.4 show no differences in behavior between observers who focused on weak models and observers who focused on good models, with the exception of one cluster of activities. This cluster concerns the focus on the models when answering the elaboration question in the post-observing phase. Remember that not all observers limited their attention to the model to which they were asked to focus on. Observers in both conditions went off-task. However, the degree of 'off-task behavior' varied in both conditions.

Table 5.4 clearly shows differences between conditions in the degree of 'off-task behavior' in the cluster 'focus on model when answering the elaboration question': observers in the condition weak focus focused only on the model mentioned in the question much more often (11 times) than observers in the good-focus condition (2 times). Related to this, observers in the condition good focus paid more attention to both models (11 times) than observers in the weak-focus condition (3 times).

Perhaps, observers in the condition good focus would like to focus on the weak model as well. This assumption is confirmed by answers from the participants in the interviews. Two observers in the condition good focus indicated that they found it difficult to comment on the good model. Joost (6) (paying four times attention to both models and one time only to the (unintended) weak model) said:

> 'I found it difficult to say what the good model did well. It seems more easy to me to explain what the other did wrong.'

Karin (5) (paying two times attention to both models, one time only to the (unintended) weak model and two times to the (intended) good model) reacted in the same manner:

> 'It seems more easy to me to say what the weak model did wrong.'

The major difference between reflecting on a good model and on a weak model lies in the availability of a frame of reference. Participants who had to explain why the weaker model was weak, could rely on the performance of the better model: the frame of reference was provided. When explaining the weak performance, participants provided a 'rule' – 'model 2 did present the standpoint and the arguments in the wrong order, because standpoint should go first'. This 'rule' describes exactly the behavior the good model showed. However, when a learner has to reflect on the better model, he has to use criteria that he already has at his disposal, an internal 'theory of good writing'.

4 DISCUSSION

This study examined students' actual behavior when performing observation tasks. In our interpretation of observers' processes we focused on three main questions: (a) indications in support of theoretical assumptions, (b) evaluation and elaboration processes, and (c) influences of different instructions.

It is found that during the execution of observation tasks, observers performed many different (meta)cognitive activities including: reading and analyzing the argumentation structure, (mentally) formulating a possible task outcome solution, observing, comparing, evaluating and judging the models, and commenting on the models' writing.

The observers' behavior was largely focused on what they observed. Their activities were directed towards the models' writing. Particularly from the interviews, we know that observers had ideas about the task outcome solution and that they compared this solution with the performance of the models, but in their actual behavior we didn't find many explicit comparisons.

Results supported our assumptions about the effectiveness of observational learning. Observers were strongly involved in metacognitive activities. They observed the models' writing, identified and conceptualized the models' writing strategies, evaluated the performance of the models and reflected explicitly on the observed performances. The performance of these activities suggests that observers have internalized, applied, and developed criteria for effective writing. Furthermore, observers also performed (mental) executional processes and during the observations, they compared their own (covert) performance with the models' actual performances. However, the difference with the models' (overt) writing is that the mental writing of the observers should be seen as a means: they 'write' in the context of the observation task, which is a learning task. They do not write in the context of a (short term) exercise task.

This study brought detailed information about evaluation and elaboration activities. In a previous study, these activities were identified as important for the effectiveness of observational learning (Braaksma, et al., 2001b). The current study revealed that observers started already evaluating and elaborating during the observation of the models and that these evaluations and elaborations are used for answering the evaluation and elaboration questions in the post-observing phase. Moreover, it appeared that the observing of model's writing was sufficient for the observers to evaluate the performance of the models and to elaborate on that evaluation. There was no need for them to read the models' written texts to answer the evaluation and elaboration questions.

Finally, this study showed differences in behavior due to differences in instructions in one cluster of activities. Observers in condition good focus went more times off-task than observers in the weak-focus condition when they had to elaborate their evaluation. Observers in condition good focus often commented on both models. We assume that it was more difficult for these observers to comment *only* on the good model because explaining why a model performed well required well-developed conscious knowledge about 'good writing' and an explicit set of criteria, which observers could apply to this explanation. Indeed, some observers in good-focus condition indicated it may be easier to comment on the weaker model because the performance of the better model can then be used as a reference.

It is important to realize that our study is a case study in which only six students participated, and that the statements we made are thus based on a small group. Furthermore, some activities in this study were only performed by one single participant in a few tasks. Therefore, the Tables 5.1, 5.2, and 5.4 are key features in this article because these tables show which activities were performed in many observation tasks by many observers and which activities were not.

The findings in this study support the importance of multiple indications of observed processes (Afflerbach & Johnston, 1984, p. 319; Smagorinsky, 1994, p. 15). It appeared that almost no observers were able to think aloud when observing two models, but by analyzing the observers' notes, we still obtained information about the observers' activities during observing phase. In addition, the interview about the observers' behavior appeared to be very useful and additive. For instance, because of the answers of the participants on an interview question we know that the students had an idea of the task outcome solution and that they were comparing this solution with the solution of the models. This behavior was not always derived from the think-aloud protocols.

The results from this study may lead to considerations about new instructions for observation tasks and future research. First, because it was found that observers did not need the models' written texts to evaluate the models' performance and to elaborate on that evaluation, one could consider to develop an observation task in which students 'only' observe the models' writing and do not receive the products of that writing. Perhaps, such an instruction elicits comments on the models' approach as well as on their written texts. Moreover, to elicit observers' evaluations and elaborations based on the models' approach, the instruction could be explicitly directed towards approach. For instance, an evaluation question could be: 'Which model approached the writing task in the best way?' and the accompanying elaboration question: 'Explain why.' A combination of an evaluation and elaboration question focused on approach could be: 'Advise the weaker model how to improve his/her approach.'

Second, this study confirmed the importance of evaluation and elaboration activities. In a previous study the importance of evaluation and elaboration activities was inferred by comparing posttest scores from observers who correctly evaluated and elaborated with posttest scores from observers who did *not* correctly performed these activities (Braaksma et al., 2001b). However, in both studies, the observers were instructed to evaluate and to elaborate the models by means of eliciting questions, and they had to *consolidate* their answers by writing them down. One might say that the added instruction, although effective, distorted the 'natural way of observing'. So, the observational processes we analyzed are embedded in an instructional environment. Further studies could aim to study: (a) the effect of consolidation, and (b) the effect of the 'explicit' asking of evaluation and elaboration activities by comparing the observational processes of different learners in a 'simpler', more natural environment.

This study showed that observers carried out many (meta)cognitive activities, especially activities that support the internalization and development of criteria for effective writing (observing, evaluating, and reflecting). These are activities that are not to be expected to appear in traditional instructional settings, in which fulfilling the writing task is the dominant activity of learners. From other studies, we know that observing writing instead of writing has a strong learning effect (Braaksma et al., in press-b; Couzijn, 1995, 1999). Now that we have observed the observing process in more detail, we have added to the theoretical framework around these empirical findings.

APPENDIX A

Description of writing processes[7] and resulting products for weak and good models in five observation tasks

Task[8]	Weak model		Good model	
	Writing process	Resulting product	Writing process	Resulting product
1	Goal-orientation, correct analysis of argument, planning of text, formulation of (incorrect) text, (correct) analysis of argument, (incorrect) formulation, writing (not thinking aloud).	Incorrect text. Text starts with argument, followed by incorrect connective and then standpoint.	Planning of standpoint and argument, correct analysis of standpoint and argument, formulation of (correct) text, writing (not thinking aloud), and re-reading of text.	Correct text. Text starts with standpoint, followed by correct connective, and then argument.
2	Immediately writing of standpoint, (incorrect) connective, argument, (incorrect) connective and argument.	Incorrect text. Text starts with standpoint, followed by incorrect connective, and then arguments (with incorrect connectives).	Goal-orientation, analysis of arguments and standpoint, planning of connectives, formulation of (correct) text, writing (not thinking aloud).	Correct text. Text starts with standpoint, followed by correct connective, and then arguments (with correct connective).
3	Formulation of (incorrect) text, (incorrect) analysis of last argument, writing.	Incorrect text. Text starts with standpoint, followed by correct connective, and then arguments (with incorrect connectives).	Goal-orientation, formulation of (correct) text, writing (not thinking aloud). Analysis of connectives.	Correct text. Text starts with standpoint, followed by correct connective, and then arguments (with correct connectives).
4	Goal-orientation, writing of (incorrect text), re-reading, semantic analysis of argument.	Incorrect text. Text starts with standpoint, followed by incorrect connective, and then argument.	Planning of connective, planning of text, formulation of (correct) text, evaluation, formulation of (correct) text, evaluation, writing.	Correct text. Text starts with argument, followed by correct connective, and then standpoint.
5	Meta-analysis, goal-orientation with analyzing elements, Planning, formulating and writing of arguments. Planning, formulating, evaluating, and writing of subordinate arguments. Meta-analysis, re-reading and evaluating text (see Table B1 in Appendix B).	Incorrect text[9]. Text starts with standpoint, followed by correct connective, and then arguments, subordinate arguments and again standpoint (with correct connectives). Main arguments are *not* immediately connected with the subordinate arguments (see Figure 5.1).	Planning, formulation, evaluation, analysis, and writing (without thinking aloud) of standpoint, arguments and subordinate arguments, re-reading of text (see Table B2 in Appendix B).	Correct text. Text starts with standpoint, followed by correct connective, and then arguments with subordinate arguments (with correct connectives). Main arguments are immediately connected with the subordinate arguments (see Figure 5.1).

7 The coding of the writing processes of the models is based on a coding scheme that is described in Braaksma, Rijlaarsdam, Van den Bergh, and Van Hout-Wolters (2001a).

8 Sequence of the models was as follows: in two instances (task 1 and task 5), first the weak model and then the good model was shown; in the other instances, the good model came first, the weak model second.

9 In contrast with the other products of the weak models, this product is not really incorrect. No incorrect connectives were used but the text is less clear than the text of model 2 because of the repeated arguments.

APPENDIX B

Table B1. Writing process of model 1 (weak model) in observation task 5

	Utterance	Cognitive activity
1	Let's see	Starts cognitive activity = Meta-analysis
2	We have a person and he says: it is better to live in a city than in a village	Reading part of argumentation structure = Goal-orientation
3	And he thinks so because in a city you can choose from different shops, schools, and swimming pools, you can better develop yourself in a city and the opportunity for finding a job is better in a city	Reading part of argumentation structure = Goal-orientation but with analyzing elements
4	He thinks you can better develop yourself in a city because you meet more different people and there are more cultural events	Reading part of argumentation structure = Goal-orientation but with analyzing elements
5	And he thinks the opportunity for finding a job is better in a city because there are a lot of companies close to each other	Reading part of argumentation structure = Goal-orientation but with analyzing elements
6	We can do this two ways	Planning next step
7	We can start with the arguments first or with the standpoints first or make a mix of them	Planning next step
8	I always prefer to start with the standpoint so it is clear immediately what it is about and so we just put that down first	Planning standpoint
9	It is better to live in a city than in a village	Writing standpoint
10	Well then you have to argument that	Planning arguments
11	Because	Formulating connective
12	Then we just put down all the arguments	Planning arguments
13	Because in a city you can choose from different shops, schools, and swimming pools, you can better develop yourself in a city and the opportunity for finding a job is better in a city	Writing arguments
14	Then we have some subordinate arguments	Planning subordinate arguments
15	We can: since you meet more different people, and there are more cultural events better develop yourself in a city	Formulating subordinate arguments and argument
16	Yes, let us do that	Evaluating formulation
17	Since you meet more different people and there are more cultural events, you can better develop yourself	Writing subordinate arguments and argument
18	And what's more, the opportunity for finding a job is better	Formulating argument
19	Yes	Evaluating formulation
20	Because the opportunity for finding a job is better	Writing argument
21	No, nonsense, that because has to move	Evaluating written text
22	The opportunity for finding a job is better because there are a lot of companies close to each other	Writing argument and subordinate argument
23	Therefore it is better to live in a city	Writing standpoint
24	So, let's see if this fits with everything	Starts cognitive activity = Meta-analysis
25	It is better to live in a city than in a village, because in a city you can choose from different shops, schools, and swimming pools, you can better develop yourself in a city and the opportunity for finding a job is better in a city. Because you can meet more different people and there are more cultural events, you can better develop yourself. The opportunity for finding a job is better because there are a lot of companies close to each other. Therefore it is better to live in a city	Re-reading written text

Table B2. Writing process of model 2 (good model) in observation task 5

	Utterance	Cognitive activity
1	Well, we start with the standpoint again	Planning standpoint
2	1, it is better to live in a city than in a village	Formulating standpoint
3	[Writing standpoint, without thinking aloud]	Writing standpoint
4	Because, firstly in a city you can choose from different shops, schools and swimming pools	Formulating argument
5	[Writing argument, without thinking aloud]	Writing argument
6	Here no subjoined, what is it called, subjoined argumentation	Analysis of subordinate argumentation
7	So I can start with 1.2	Planning argument
8	Secondly, you can better develop yourself in a city	Formulating argument
9	[Writing argument, without thinking aloud]	Writing argument
10	Because you can meet more different people and there are more cultural events	Formulating subordinate arguments
11	[Writing subordinate arguments, without thinking aloud]	Writing subordinate arguments
12	Well now only 1.3 and 1.3.1	Planning argument and subordinate argument
13	You can say moreover or thirdly	Planning connective
14	Moreover once again	Planning connective
15	Moreover is the opportunity for finding a job better in a city	Formulating argument
16	No	Evaluating formulation
17	Moreover the opportunity for finding a job is better in a city	Formulating argument
18	[Writing argument, without thinking aloud]	Writing argument
19	Because there	Formulating subordinate argument
20	I just don't know what it is called, subjoined, subordinate argumentation	Analysis subordinate argumentation
21	Well, then you can put because, but for the sake of originality we just put as	Planning connective
22	Moreover the opportunity for finding a job is better in a city as	Formulating argument with connective
23	Let’s look	Starts cognitive activity = Meta-analysis
24	Yes	Evaluating formulation
25	As there are a lot of companies close to each other	Formulating subordinate argument
26	[Writing subordinate argument, without thinking aloud]	Writing subordinate argument
27	It is better to live in a city than in a village, because firstly in a city you can choose from different shops, schools, and swimming pools. Secondly you can better develop yourself in a city, because you can meet more different people and there are more cultural events. Moreover, the opportunity for finding a job is better in a city as there are a lot of companies close to each other	Re-reading written text

CHAPTER 6

DISCUSSION

Abstract

This thesis is aimed at getting deeper insight into different aspects of observational learning in argumentative writing. In four studies I tried to address three issues of observational learning: (a) the observational process itself, (b) the way learner characteristics are involved in instructional variants, and (c) the effects of observational learning on writing processes. In this concluding chapter, the main results of the four separate studies are synthesized, and some of these results are compared with related studies. Then, some reflections upon the studies are offered. Finally, some directions for future research and implications for educational practice are presented.

1 SYNTHESIS OF RESULTS

In chapter 1, I introduced the theoretical framework of this thesis. Figure 6.1 shows this framework again.

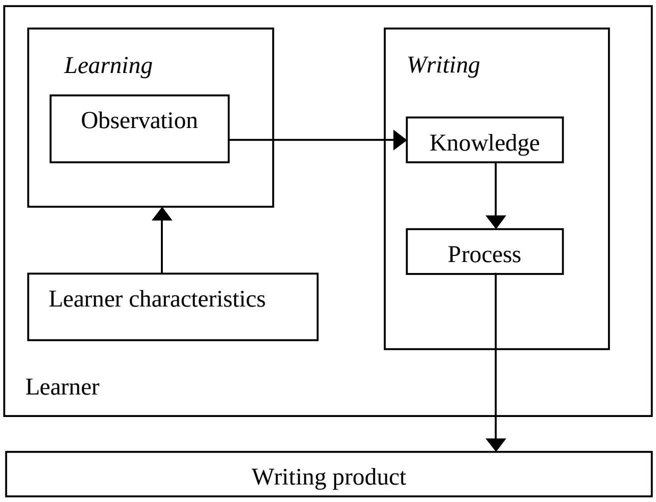


Figure 6.1. Theoretical framework

Figure 6.1 shows the chain of variables involved in explaining the effectiveness of observational learning. In sum, observational learning changes the learners' knowledge about writing. These changes in knowledge affect the organization of writing processes, which in turn affects the quality of the resulting writing product.

In this thesis, three new issues in observational learning were addressed: the observational process itself, the role of learners' characteristics on the effect of different types of observational learning, and the effect of observational learning on the mediating variable 'writing processes'. By studying these three issues, I aimed to contribute to the theoretical framework of observational learning. Using different methodologies – post-hoc analysis, experimental research and a case study – I tried to gain a better insight into the details of observational learning in argumentative writing.

1.1 Processes of observational learning

From the studies reported in chapters 2 and 5, some important insights are inferred on the processes involved in observational learning. We now know that observers are strongly involved in metacognitive activities. They observe the writing of the models, identify and conceptualize the models' writing strategies, evaluate the performance of the models and reflect explicitly on the observed performances. The performance of these activities suggests that observers have developed, applied, and internalized criteria for effective writing.

Two of these activities were found to be crucial for the effectiveness of observational learning: evaluation and elaboration. Now, we know that these activities are indeed already performed when observers are observing the performance of the models, and that observing the models' writing processes is sufficient to perform these activities. This is an important finding, because it refutes the alternative explanation of the effects of observational learning that the effect might also be attributed to the observation and comparison of the resulting texts, not of processes. This explanation would hold that students could skip the process, and only had to evaluate the quality of the text. However, this is not the case: the observations of processes play an important role and enable students to reflect on the observed performances. The writing processes provide observers sufficient information for evaluating the models' writing and to elaborate on that evaluation. The evaluations and elaborations performed after the observers read the models' written texts are mostly repetitions of the evaluations and elaborations performed during the process observations.

Furthermore, we have indications that observers also perform (mental) executional processes ('mental writing') before and during the observations, and that they compare their own (covert) performance or problem solution with the models' actual performances.

These results support the theoretical framework where I hypothesized that metacognitive activities construct the knowledge about writing (Graham & Harris, 1994), and assumed that these activities were stimulated in observational learning. The results confirmed that indeed in observational learning students use important meta-

cognitive strategies: they observe ('Which strategy does model 1 use?'), evaluate ('Is this strategy going well?'), and elaborate ('Why is this strategy going well?').
I have identified these last two activities as crucial. Furthermore, I assumed that writers who learned by observation could more easily step back from their own executional process and use metacognitive strategies more easily because their cognitive effort is directed to learning to write rather than to writing (Couzijn, 1995, 1999; Rijlaarsdam & Couzijn, 2000a, 2000b). It seems that indeed most of the cognitive effort is directed to the *learning* task. The covert writing of the observers costs some cognitive effort, but instead of the all-effort-consuming writing task, this covert writing task is 'nested' in the observation task. Covert writing is a means in the learning task, not a writing task to finish. Because observers 'write' in the context of the observation task, which is a learning task, they are able to pay attention to learning to write.

Two other findings should be mentioned here. In chapter 2, it was revealed that learning activities in observation tasks are rather content independent. Regardless of the actual educational content, students were consistent in their learning activities: students who performed a learning activity rather well in Couzijn's lesson A (about argumentation structures) performed this learning activity relatively good in lesson B (about presentation elements). Furthermore, the learning activities that observation tasks consist of, seem to be already existent to a certain degree in good learners' repertoires. This conclusion is based on the observed effect of aptitude on the performance of learning activities. This implies that good learners could engage in these metacognitive activities during the execution of writing tasks.

A result to conclude with is the difference in learning behavior due to instructional differences in the observation tasks (chapter 5). It seems to be more difficult to focus on a good writer than to focus on a weak writer: focusing on the performance of good writers caused more frequent off task behavior than focusing on weak writers.

1.2 Effects on writing processes

We now know more about the effects of observational learning on writing processes (chapter 4). We are aware that observational learning influences the writing processes differently than learning-by-doing. Writers who learned by observation perform relatively more metacognitive activities (Goal-orientation and Analysis) at the start and relatively more executional activities (Writing and Re-reading) in the second part of the writing process than writers who learned by doing. Over the whole writing process, writers who learned by observation show more Planning activities than writers who learned by doing. Moreover, in the middle and last part of the writing process, writers who learned by observation perform increasingly more Meta-analyzing activities, indicating monitoring and regulating processes, than writers who learned by doing. Furthermore, writers who learned by observation showed for some activities a changing execution over time, whereas writers in doing-writing condition performed these activities at a constant rate during the writing process (i.e., a monotonous process). Moreover, variances in both observational learning conditions were larger than in the control condition, indicating more heterogeneous processes.

These effects from observational learning on writing processes are in line with the theoretical framework. Through adopting the orchestration examples the models provide, moving cognitive effort from writing to learning to write, and practicing metacognitive strategies, students in an observational learning environment may develop richer knowledge about writing. This knowledge influences writing processes in the direction I found: more knowledge about the task, the genre, the procedure, leads to more planning and analysis (cf. McCutchen, 2000).

1.3 Effects of learners' characteristics on the effects of instructions

I have established that the same instructional method is not the most effective for every learner (chapter 3). Moreover, task familiarity seems to play an important role in this finding.

When a task is *new*, weak students benefit more from observational learning (focusing on weak models) than from performing writing tasks as they are less likely to use metacognitive strategies such as monitoring, evaluating and reflecting. Moreover, they profit from observational learning because their cognitive effort is shifted from executing writing tasks to learning from writing processes of others. They can thus focus on the learning task, providing themselves with a learning opportunity to acquire new understanding about writing. However, weak learners benefit *only* from observational learning when they reflect on weak models, they do not profit from reflecting on good models. My explanation of this finding is that it is much easier to evaluate the weak model because the performance of the better model can be used as a reference. Besides, the performance of the observed weak model is probably more matched with the cognitive processes of a weak learner.

When a task is new for better learners they benefit not only from observational learning (focusing on good models) but also from performing writing tasks. They are probably able to divide their attention between writing task and learning task, and thus generate enough input for their learning by evaluating and reflecting on their own performance. However, good learners benefit *only* from observational learning when they reflect on good models, they do not profit from reflecting on weak models. Probably, the performance of a good model is more matched with the cognitive processes of a good learner. Moreover, good learners are able to identify and comment on the qualities of the good model, because they already have their own internal set of evaluation criteria available.

When a task is *familiar*, weak students benefit from performing writing tasks as much as they profit from focusing their observations on weak models' writing. Because the learners had already experienced successes with the tasks, they had a chance to construct knowledge about good writing, and thus were better equipped to learn from the writing task themselves. Reflecting on the performance of the better model however, is still not effective for these weak students.

When a task is familiar, good students profit only from observational learning with the focus on good models, not from focusing on weak models or from performing tasks themselves. Maybe, they need the challenge of explaining why the better model performed well.

The results concerning the observational conditions contribute to the research about the role that model-observer similarity plays in children's behavioral change (see Schunk, 1987). The finding that model-observer similarity in competence facilitated learning corresponds with findings of Schunk, Hanson and Cox (1987). They found that for weak students observing a coping model (i.e., a model that shows his hesitations and errors, but gradually improves his performance) led to greater perceived similarity in competence to the model, higher self-efficacy for learning, and higher posttest self-efficacy and skill.

2 COMPARISON WITH COUZIJN'S STUDIES

My studies are strongly related to and inspired by Couzijn's studies. He conducted several experimental studies in which he compared learning-by-doing with observational learning (Couzijn, 1995, 1999). He showed that observational learning was more effective than learning-by-doing. As opposed to Couzijn, I did not observe a main effect from observational learning on writing products. From study 2 it appeared that 'medium' students learned as much from observation (in two types) as from doing (see chapter 3). However, several circumstances complicate a direct comparison with Couzijn's studies.

First, participants differed. In Couzijn's studies, a group of 15-year-old students who had just finished the ninth-grade (intermediate and higher level) participated. In my study, students (mean age: 14 years) who were halfway the eight-grade (low-high level) participated. Thus, the level of 'my medium' students is much lower than the 'medium' students who took part in Couzijn's studies.

Furthermore, the setting differed. Whereas Couzijn's studies took place in a laboratory situation and students participated voluntarily and received a financial reward, my study is a field study, which took place at school as part of regular Dutch lessons and thus all students in the classes participated.

Moreover, I shortened the amount of learning time, and focused the content of the course on 'argumentative texts'. Students who participated in my study, attended less instructional sessions than students in Couzijn's study did, and the content was mainly restricted to transforming argumentation structures into linear texts. Therefore, the pretests and posttests I used for the analyses also differed from Couzijn's tests.

Last, but certainly not least, I would like to point at my optimization of the procedures for the control group. In Couzijn's study the participants in the control conditions worked during the lessons individually at their own pace in their workbooks with theory and exercises, while the instructions for participants in the observational learning conditions were presented via a personal video player and were paced. Couzijn's effects might be caused by the use of video-instruction and/or the pacing. In my study, therefore, the same procedures were used for control and observational learning conditions. Participants in all conditions were instructed via screen, and the pacing of theory and exercises was the same for all participants.

3 ALTERNATIVE EXPLANATIONS, IMPLEMENTATION AND METHODOLOGICAL ISSUES

In this section, I will reflect on four aspects of the current studies. When possible, some suggestions for improvements or for future research are given.

3.1 Alternative explanations

For the effects from observational learning on writing processes (study 3, chapter 4) and especially on writing products (study 2, chapter 3), one could propose alternative explanations, especially concerning the model-observer similarity effects regarding the weak students, assessed in the second study (chapter 3). One could suggest that these learning effects were caused by a changing self-efficacy. Schunk and his colleagues (Schunk & Hanson, 1985, 1989a; Schunk et al., 1987; see also Schunk, 1998, p. 148) reported effects of (various) models on students' self-efficacy, which in turn influenced learning and achievement. Perhaps the present outcomes are mediated by changes in students' self-efficacy.

However, I chose for a different operationalization of similarity than Schunk and his colleagues. As a consequence I made other choices concerning the execution of the models and the dependent variables. Although the models in my studies did verbalize their thoughts, they focused on the execution of the writing process, not on achievement beliefs and did not reflect (high and low) self-efficacy. Because of that different focus, I did not include self-efficacy measures as dependent variables.

However, in future studies it is advisable to include these measures because self-efficacy is hypothesized to influence choice of activities, effort expenditure, persistence and achievement (Bandura, 1986, 1997; Schunk, 1998).

3.2 Effect sizes

A second point concerns the effect sizes in the studies, which were consistent, but not very large and could not be calculated in every study. Reporting statistical significances of findings do not indicate whether an effect is substantial or not: effect sizes do.

Concerning the third and fourth studies (chapters 4 and 5), it is not wise to calculate sizes. For study 3, where the effects of observational learning and learning-by-doing on writing processes were examined, overall effect sizes would be an underestimate of the differences between the conditions because the differences between conditions depend on the moment in the writing process (see chapter 4). Hence, the effect sizes range from negative values to positive values dependent on the moment in the writing process and are therefore not very informative. The aim of study 4 (chapter 5), a qualitative case study in which six students participated, was to get insight into the observers' behavior. In such a case study where effects are not quantified or tested, effect sizes are beside the point.

In study 1 (chapter 2) the explained variances can be used as indications for the effect sizes. The activity evaluation explained 4.1% of the variance in the writing posttest, and the activity product-elaboration 3.7%. According to Cohen (1988), these effects can be interpreted as medium effects. In the discussion of chapter 2, it was put forward that large effects were not to be expected because the indications of the activities were indirect (I did not observe learners at work, but coded their materials), and because the activities were performed at a very high level (the aim was mastery, not selection).

For study 2 (chapter 3) effect sizes are listed in Table 6.1 for different aptitude values, only in those cases where significant results between conditions were observed.

Table 6.1. Effect sizes for different comparisons, for different levels of aptitude (low vs. high) and two measurement occasions (1st and 2nd)

	Aptitude			
	Low		High	
	1st	2nd	1st	2nd
Direct writing vs. observation/weak focus	.36	*ns*	.28	*ns*
Direct writing vs. observation/good focus	*ns*	.51	*ns*	.18
Observation/weak focus vs. observation/good focus	.10	.26	.30	.07

Hence, from Table 6.1 it can be observed that in general it concerns only small effects on both measurement occasions. However disappointing at first glance, these figures are in a range generally reported (Hillocks, 1986).

A special feature of the results of the second measurement occasion is that the differences between conditions can be compared to the influence of the previous writing scores. So, a more interpretative measure of effect size can be obtained. Comparing direct writing and observation/good focus it shows that for low-aptitude students, condition is about twice as important as previously shown writing skill, for high-aptitude students, the effect of condition is clearly less marked (i.e., only half that of previously observed writing). In the comparison of both observational conditions, condition proved to be as important as previously writing for low-aptitude students and almost twice as important for high-aptitude students. So, from an interpretative point of view the influence of conditions are not to be neglected. However, these effects are at least partly due to the rather low correlations between different writing tasks.

The small effects sizes are perhaps partly due to the relatively short courses. For the comparison of both observational conditions another explanation for the (relatively) disappointing effect sizes can be put forward. In both conditions students observed a pair of models: a relatively good model and a relatively weak model. The only difference between conditions is defined by the instruction the students received. It has been shown that the students did not follow these instructions to the letter: in both conditions some students paid attention to the other model as well.

This inevitably limits the differences between conditions and therefore the effect size.

I would like to mention a second point too that might have limited the differences between observational conditions as well. It concerns the differences between the good and the weak model. It was shown (in chapter 5) that the weak models might not have been that weak, only the final solution of these models were incorrect; their writing processes were not always that weak. With hindsight, it might have been preferable to exaggerate the differences between both models in terms of processes and products.

3.3 Implementation of instructions

As mentioned before, not all students followed the instruction precisely. This observed deviation is inspired by the fact that two models were presented to each student. Differences between conditions in 'instructions-as-set' and 'instructions-as-performed' were assessed. In condition good focus, observers reflect more on both models than observers in condition weak focus do. This result is found in study 4 (chapter 5), but might have influenced the results of study 2 (chapter 3). This raises the question of the interpretation of the reported similarity effects. These effects could partly result from the focus on the 'wrong' model. This issue needs further elaboration by experimental studies in which participants are allowed to observe *only* a good model, or *only* a weak model (cf. Schunk & Hanson, 1985, 1989a; Schunk et al., 1987). When such an instruction form is chosen, one has also more control about what students are really doing.

3.4 Experimental design

A final point I would like to stress concerns the design in the second study (chapter 3) in which two instruction sessions were implemented. After all, the implementation of two instruction sessions has advantages (one can test the 'similarity hypothesis' when a task is new and when a task already is more regular, and two variants of the instruction may enhance the validity of the independent variable; see chapter 3) but other choices could have been made.

In my case, the choice for implementing two instruction sessions was strongly related with compromises one has to make when performing a field study. Originally, I planned only one instruction session (the session with the models on video) and a number of delayed posttests to measure the long-term effects of observational learning. However, when the posttests were administered after that first session, I noticed and felt sorry that some of the students had not learned sufficiently; they had not learned all they could have learned. Because the study took place at school, in their regular Dutch lessons, I felt myself obliged to give up the delayed posttests and instead I implemented a second instruction session. However, the implementation of a second instruction session is in fact more the rule than the exception. In describing an intervention program with observational learning, Schunk (1998) states that:

> 'in a typical intervention, there are several instruction sessions because students usually require multiple sessions to begin to acquire skills and because model treatments are often implemented more than once' (Schunk, 1998, p. 145).

In the second session, the students stayed in the same conditions and followed a second instruction session according to their conditions: students in control condition direct writing performed writing tasks themselves, and students in both observational learning conditions observed pairs of models performing writing tasks with the focus on weak models or on good models. However, there was one difference. Because there was no time to record new models on videotapes, the models were not shown on video but performed 'live', a research assistant and I acting as models following a script. The behavior of the models did not differ from their behavior in the first instruction session, although adults were acting as models, their behavior was of that of students and similar to the first instruction session.

This design made it possible to add the variable 'task newness' into the analyses, but unfortunately this variable was nested in the two operationalizations of 'observation': video and live observation. No conclusions can be formulated about the different effects of these two operationalizations. For this reason a counterbalanced design would have been preferable.

Concluding, although it is unfortunate that no information is obtained of the (similarity) effects of observational learning over a longer period, I feel that I made the best decision giving priority to the students' learning. My solution shows that this does not have to go at the expense of the results.

4 DIRECTIONS FOR FUTURE RESEARCH

This section addresses seven aspects of the current thesis that have to be examined in future studies.

4.1 Further studies in the theoretical framework

The verification of assumed influences in the theoretical framework should be a main and important aspect of future studies. First, is knowledge of writers who have learned by observation indeed larger than the knowledge of writers who learned by doing? To study this, one has to include measurements of knowledge as a dependent variable when one compares observational learning with learning-by-doing. In the future, it will be wise to include measurements of genre knowledge, and measurements of knowledge of procedures to guide and control the effective realization of the text production (task schemas), because observational learning in all likelihood has the most effect on these types of knowledge.

Second, it is necessary to verify the influence of knowledge on writing processes. Until now, studies which assume that (genre) knowledge affects writing processes, do not actually measure writing processes (except for McCutchen, 2000). They 'only' measure effects of knowledge on writing products, and assumed that these effects were mediated by changed processes. In study 3, I showed that observational learning changed the organization of writing processes and I assumed that this change was caused by a change in knowledge. However, the precise influence of knowledge on processes has been left more or less in the dark: if students have extensive knowledge about how to approach a task (i.e., genre knowledge and (more general) task schema knowledge, thus procedural knowledge), does that influence their writing approach? To answer this type of question experimental research in which aspects of knowledge are manipulated systematically is necessary. Only in the latter are causal inferences on the relations between knowledge and (the organization of) writing processes possible.

Finally, the assumed influences of organization of writing processes on quality of writing products have to be verified. Does the type of writing organization found in writers who learned by observation (more Analysis and Goal-orientation in the beginning of the process, more Planning over the whole process etc.) indeed lead to better writing products? I did not examine the effect of orchestration of writing process on writing products in the third study. However, from another study (Braaksma, Rijlaarsdam, Couzijn & Van den Bergh, in press-a) I know that there is a relation between the occurrence of planning and analyzing in the writing process and the quality of the resulting product. But in that study, only frequencies of processes were related to quality of products. No attempts were made to assess the orchestration of writing processes in that study.

4.2 New types of instructions for observational learning

Another important part of future research regards the effects of new types of instructions for observational learning. Because it was found that observers did not need the models' written texts to evaluate the models' performance and to elaborate on that evaluation, one could consider developing an observation task in which students *only* observe the models' writing processes and do not receive the products of that writing. One can contrast this new instruction with an instruction in which students observe writing processes *and* writing products (the current instructions), and with an instruction in which students *only* 'observe' writing products. If students do learn as much, or even more from the 'process-only' instruction, we know that students do learn from the observation of writing processes only, and the finding in the case study that process observations are enough to evaluate the models' performance and to elaborate on that evaluation, is then experimentally confirmed.

Because I found that the observers rarely made comments on approach and because it is wise to stimulate comments in order to deepen students' insight into writing processes, maybe such 'process-only' instruction will elicit more comments on the models' approach. Furthermore, if the goal is to elicit observers' evaluations and elaborations based on the models' approach, the instruction could also be explicitly directed towards approach. For instance, an evaluation question could be: 'Which model approached the writing task in the best way?' and the accompanying elaboration question: 'Explain why.' A combination of an evaluation and elaboration question focused on approach could be: 'Advise the weaker model how to improve his/her approach.'

Other possible types of instructions are related to evaluation and elaboration activities. In study 4 (chapter 5), the importance of evaluation and elaboration activities was confirmed. In study 1 (chapter 2) this importance was inferred by comparing posttest scores from observers who correctly evaluated and elaborated with posttest scores from observers who did *not* correctly perform these activities. However, in both studies, the observers were instructed to evaluate and to elaborate the models by means of eliciting questions, and to *consolidate* their answers by writing them down. One might say that the added instruction, although effective, distorted the 'natural way of observing'. So, the observational processes I analyzed are embedded in an instructional environment. It would be worthwhile to study the effect of consolidation by implementing an instruction in which students don't have to write down their evaluations and elaborations. To study the effect of the explicit asking of evaluation and elaboration activities, one could compare an instruction in which learners observe without evaluation and elaboration questions (a more 'natural' environment) with an instruction in which these questions are added (the current instructions).

4.3 Type of task

Connected to the previous suggestions for new instructions, is the point of what kind of tasks the observation tasks in the current studies precisely are, and whether this type of task is the most effective. Bandura (1986, p. 73) nuances the characterization of Marlatt (1972) that instructions provide rules for generating examples (i.e., deductive in nature), and modeling provides examples for inducing rules (i.e., inductive in nature). Bandura states that this distinction may apply in some instances, but not in all. Instructions are not always generic, and modeling is not necessarily specific. In cognitive modeling, models verbalize the rules and strategies guiding their choice of actions (Meichenbaum, 1977). It is thus stressed by Bandura that whether observational learning occurs inductively or deductively, depends on what is being modeled. He then refers to research from Rosenthal and Zimmerman (1978) who have analyzed a number of studies that have investigated the extent to which children acquire and then generalize the rule-governed behavior to new situations. Three types of instructions are distinguished, children: (a) have been instructed in the rule, (b) have simply observed modeled performances embodying the rule, or (c) have been provided with the rule, along with modeled examples. It was found that with young children, behavioral modeling (from which the rule underlying the responses must be inferred; type b) is comparable or superior to verbal instruction in the rule for producing generalized changes. However, generalized cognitive skills are best imparted by providing rules with demonstrations of how they can be applied in certain situations (i.e., type c) (see Bandura, 1986, p. 73).

The type of observational learning in my studies can be characterized as the last instruction type: in the theory sections, students have been provided with the rule (e.g., when an argumentative text starts with the standpoint, then standpoint and arguments are connected with connectives as 'because'), and the modeled examples show the application of the rule. The models verbalized the rules and strategies that guided their choice of actions (i.e., cognitive modeling). I tend to characterize this type of observational learning as deductive. This is one of the two effective types of instructions Rosenthal & Zimmerman (1978) mentioned. The other effective instruction type was type b, in which learners infer the rule from the modeled behavior. Rosenthal and Zimmerman restricted the effectiveness to younger children. Nevertheless, it would be worthwhile to study type b with students (mixed ability) in the age group I studied. First, this allows studying the effectiveness of observational learning in a more pure form: the effect is not related to explicit theory sections in which the rule is presented. Secondly, when such a study is conducted, one could also examine for which type of student such inductive instruction method is effective. It might be that this method is especially valuable for better learners.

4.4 Sequence of instructions

Related to the previous directions for future research, is a study on the sequence in which the instructions are presented. From study 2 (chapter 3) we know that there is an interaction between familiarity with the task and the type of instruction for different types of student (aptitude). Perhaps, one can develop an ideal sequence for students. With the findings from study 2 such a sequence could be the following: when a learner is weak and the task is new, it is the best to apply the theory in observation tasks with the focus on the weak model. When the task is more familiar to this (however, still weak) student, in addition to observational learning (with the focus on the weak model), he also learns by performing writing tasks. The last step is when the task is still familiar but when the learner has developed into a good student. Then, the most effective instruction for him is to learn by observation and to reflect on the better model. However, it can also be interesting to let the students choose by themselves from what kind of instruction they think they can learn the best. One could consider providing hints ('If you think you are not very good in this task, it is maybe the best…') but within the framework of self-directed learning maybe this is not preferable.

In this 'sequence research' also the other proposed instruction types (e.g., process-directed, more natural, inductive instructions) can be involved to examine what type of instruction is at which moment the best for which type of student.

4.5 Writing tasks

Another direction for future research concerns the writing tasks. In the current studies, these tasks are constrained: what the students were called upon to do (and what the models they observe were doing) was order and connect ideas given to them. The participants did not have to come up with new ideas and compose thoughts. This made the tasks a bit unnatural and certainly less communicative. However, on the other hand, there was also the advantage of presenting ideas to the students. Because content of the standpoint and (subordinate) arguments were already provided to them, the topic knowledge was equal for all students. This fact was especially important for the study on the effects of observational learning on writing processes (study 3). Because topic knowledge was a constant factor, it couldn't disturb the effects from (genre) knowledge on writing organization. All students knew what they had to write about.

For future studies I propose to study the effects of observational learning for larger, less-structured and more communicative writing tasks as they occur in the Dutch language curriculum in the higher years of secondary education, for instance, larger 'real' texts as an opinion article in a newspaper or book reviews. One could even think of studying the effects of observational learning when students compose hypertexts such as web pages or PowerPoint presentations (compare Braaksma et al., in press-a).

4.6 Transfer effects

A sixth issue for future research regards one of the major findings of Couzijn (1995). He also investigated transfer effects from reading to writing and from writing to reading in different conditions. He showed that the transfer effects of observational learning from writing to reading were much larger than the transfer effects from reading to writing. This was one of the reasons for Couzijn recommending observational conditions to others, and observation of writing above observation of reading. For me it was the reason to restrict this thesis to observation of writing, and not to include a concurrent condition in which reading is observed.

Although this design prevents me to compare the transfer effects from writing to reading with the effects from reading to writing, it is possible to assess the transfer effects from writing to reading. Reading measures were included as dependent variables in the data-collection, and in the near future I plan to study these transfer effects.

4.7 Observational learning in regular lessons

As a final direction for future research I suggest to develop a series of lessons in which observational learning plays an important role, and to implement that series in the 'real' curriculum. Although the experimental studies (chapter 3 and 4) took place in regular Dutch lessons, the teaching approach was very restricted. There was no place for spontaneous interactions with the students. I would like to study the place of observational learning in regular lessons and examine how students and teachers act with observational learning. This might give clues for development of optimal observational-learning lessons.

5 IMPLICATIONS FOR EDUCATIONAL PRACTICE

First of all I want to advise teachers to alternate the 'usual' writing tasks with observation tasks so that students are relatively easy occupied with monitoring, evaluation and reflection activities and the focus is more naturally on writing processes. It is certainly not necessary to construct videotapes with peer-models. The teacher can, as an expert-model, or as a coping model, perform writing tasks thinking aloud in front of the classroom. When the teacher uses a script for the performance of the model(s), which is based on real problems of students (like the second instruction session in study 2), he can treat very well the problems with which students struggle.

However, it is absolutely not necessary that a teacher is acting as a model. In the second instruction session of study 2, in which a research assistant and I were acting as models, several students indicated enthusiastically that they wanted to act as models as well, and I am sure that they are able to. From a pedagogical point of view, the instructions should be revised then. Emphasis should not be placed on the faults of the weaker writer, but on contrasting and evaluating the performances, elaborating on suggestions for better performance for one or both models.

Second, I think it is important to differentiate the instructions and adapt the instructions when the task becomes familiar. A couple of times in this thesis I stressed the strengths of model-observer similarity and also here I want to emphasize that it is more effective (and more easy and thus more comfortable) for weak students to reflect on weak models, and for good students it is more effective (and more difficult and thus more challenging) to reflect on good models.

Until we know more about the effectiveness of 'natural' observing (see section 4.2), it is important that when students are observing models, they explicitly evaluate the model's performance, and that they explicitly elaborate on that evaluation. That can be done by asking them eliciting questions as in the materials I developed. The students can answer these questions individually on paper, and then later the answers can be discussed, but it is also possible that the students directly answer the questions verbally and discuss them in pairs or in groups.

Furthermore, I suggest that it will be very informative if students construct scripts for several task performances themselves. Especially the construction of scripts for a good and a weak task performance gives students insight into the criteria for good writing.

Finally, observational learning seems a type of education, which suits the attainment targets of the basic secondary education. This holds for writing, as shown in the introduction, but for other subject matters as well. Besides, it also agrees with guidelines proposed by the Dutch Inspectorate of Education. In the last evaluation of basic secondary education, it was mentioned that possibilities for students to actively process learning materials is underused in the first grades of Dutch secondary education (Inspectie van het Onderwijs, 1999). Observational learning might be a tool to fill up this gap.

REFERENCES

Afflerbach, P., & Johnston, P. (1984). Research methodology on the use of verbal reports in reading research. *Journal of Reading behavior, 16,* 307-322.

Alamargot, D., & Chanquoy, L. (2001). Nature and control of processing. In G. Rijlaarsdam (Series Ed.) & D. Alamargot & L. Chanquoy (Vol. Eds.), *Studies in writing: Vol. 9. Through the models of writing* (pp. 125-154). Dordrecht: Kluwer Academic Publishers.

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

Bandura, A. (1997). *Self-efficacy. The exercise of control.* New York: W.H. Freeman and Company.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley.

Braaksma, M. A. H., & Rijlaarsdam, G. (1997, August). *Learning to write and read argumentative texts by observation. An explorative study on effective learning activities in observation tasks.* Paper presented at the 7th European Conference for Research on Learning and Instruction, Athens, Greece.

Braaksma, M. A. H., Rijlaarsdam, G., Couzijn, M., & Van den Bergh, H. (in press-a). Learning to compose hypertext and linear text: Transfer or interference? In R. Bromme & E. Stahl (Eds.), *Writing hypertext and learning: Conceptual and empirical approaches*. Pergamon.

Braaksma, M. A. H., Rijlaarsdam, G., & Van den Bergh, H. (in press-b). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology.*

Braaksma, M. A. H., Rijlaarsdam, G., Van den Bergh, H., & Van Hout-Wolters, B. H. A. M. (2001a). *Observational learning and its effects on the orchestration of writing processes.* Manuscript submitted for publication.

Braaksma, M. A. H., Van den Bergh, H., Rijlaarsdam, G., & Couzijn, M. (2001b). Effective learning activities in observation tasks when learning to write and read argumentative texts. *European Journal of Psychology of Education, 1,* 33-48.

Breetvelt, I., Van den Bergh, H., & Rijlaarsdam, G. (1994). Relations between writing processes and text quality: When and how? *Cognition and Instruction, 12,* (2), 103-123.

Breetvelt, I., Van den Bergh, H., & Rijlaarsdam, G. (1996). Rereading and generating and their relation to text quality, An application of multilevel analysis on writing process data. In G. Rijlaarsdam, H. van den Bergh & M. Couzijn (Eds.), *Theories, models & methodology in writing research* (pp. 10-21). Amsterdam: Amsterdam University Press.

Brown, A. L. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 65-116). Hillsdale, NJ: Erlbaum.

Butler, D. L. (1998). Metacognition and learning disabilities. In B. Y. L. Wong (Ed.), *Learning about learning disabilities* (2nd ed., pp. 277-307). San Diego, CA: Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, N.J.: Lawrence Erlbaum.

Collins, A., Brown, B. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing and mathematics. In L. B. Resnick (Ed.), *Knowledge, learning and instruction. Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Erlbaum.

Cook, Th., & Campbell, D. T. (1979). *Quasi Experimentation: Design & analysis issues for field settings*. Chicago: Rand McNally College Pub.

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

Couzijn, M. (1999). Learning to write by observation of writing and reading processes: Effects on learning and transfer. *Learning and Instruction, 2,* 109-142.

Couzijn, M., & Rijlaarsdam, G. (1996). Learning to read and write argumentative text by observation. In G. Rijlaarsdam, H. van den Bergh & M. Couzijn (Eds.), *Current trends in writing research: Effective teaching and learning of writing* (pp. 253-272). Amsterdam: Amsterdam University Press.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12,* 671-684.

DAT (1984). *Differentiële Aanleg Test.* [Differential Aptitude Test.] Lisse: Zwets & Zeitlinger.

Englert, C. S., Stewart, S. R., & Hiebert, E. H. (1988). Young writers' use of text structure in expository text generation. *Journal of Educational Psychology, 80,* 143-155.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis. Verbal reports as data. Revised edition.* Cambridge: The MIT Press.

Ferrari, M., Bouffard, T., & Rainville, L. (1998). What makes a good writer? Differences in good and poor writers' self-regulation of writing. *Instructional Science, 26,* 473-488.

Flower, L. S., & Hayes, J. R. (1980). The dynamics of composing: Making plans and juggling constraints. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing: An interdisciplinary approach* (pp. 31-50). Hillsdale, NJ: Lawrence Erlbaum.

Goldstein, H. (1979). *The design and analysis of longitudinal studies: Their role in the measurement of change*. London: Academic.

Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). London: Edward Arnold.

Graham S., & Harris, K. R. (1994). The role and development of self-regulation in the writing process. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications* (pp. 203-228). Hillsdale, N.J.: Erlbaum.

Graham, S., & Harris, K. R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist, 35*(1), 3-12.

Graham, S., Harris, K. R., MacArthur, C., & Schwartz, S. (1998). Writing Instruction. In B. Y. L. Wong (Ed.), *Learning about learning disabilities* (2nd ed., pp. 391-421). San Diego, CA: Academic Press.

Graham, S., Harris, K. A., & Troia, G. A. (1998). Writing and self-regulation: Cases from the self-regulated strategy development model. In D.H. Schunk & B.J. Zimmerman (Eds.), *Self-regulated learning, from teaching to self-reflective practice* (pp. 20-41). New York: The Guilford Press.

Hayes, J. R. (1996). A new framework for understanding cognition and effect in writing. In C.M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 1-27). Hillsdale, N.J.: Erlbaum.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing: An interdisciplinary approach* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum.

Hillocks, G. (1986). *Research on written composition: New directions for teaching*. Urbana, Ill.: ERIC Clearinghouse on Reading and Communication Skills, National Institute of Education.

Hoogeveen, M., & Bonset, H. (1998). *Het schoolvak Nederlands onderzocht. Een inventarisatie van onderzoek naar onderwijs Nederlands als eerste en tweede taal in Nederland en Vlaanderen*. [The subject Dutch examined. An inventory of research on education in Dutch as first and second language in the Netherlands and Flanders.] Leuven/Apeldoorn: Garant.

Inspectie van het Onderwijs [Dutch Inspectorate of Education] (1999). *Werk aan de basis: Evaluatie van de basisvorming na vijf jaar*. [The evaluation of basic secondary education in the Netherlands.] Utrecht: Inspectie van het onderwijs.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equations modeling with the SIMPLIS command language*. Hillsdale, N.J.: Erlbaum.

Kellogg, R. T. (1994). *Psychology of writing*. New York: Oxford University Press.

Kuhlemeier, H., Van den Bergh, H., & Rijlaarsdam, G. (1997, August). *Speaking and writing apprehension: dimensionality, reliability, and track and gender differences*. Paper presented at the 7th European Conference for Research on Learning and Instruction, Athens, Greece.

Lysynchuk, L. M., Pressley, M., d'Ailly, H., Smith, M., & Cake, H. (1989). A methodological analysis of experimental studies of comprehension strategy instruction. *Reading Research Quarterly, 4,* 458-470.

Marlatt, G. A. (1972). Task structure and the experimental modification of verbal behavior. *Psychological bulletin, 78* (5), 335-350.

McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist, 35* (1), 13-23.

McCutchen, D., Covill, A., Hoyne, S. H., & Mildes, K. (1994). Individual differences in writing: Implications of translating fluency. *Journal of Educational Psychology, 86,* 256-266.

Meichenbaum, D. (1977). *Cognitive behavior modification: An integrative approach*. New York: Plenum.

Ministerie van OCenW [Ministry of Education, Culture and Science] (1998). *Kerndoelen basisvorming 1998-2003. Relaties in beeld. Over de relaties tussen de algemene doelen en de kerndoelen per vak.* [Attainment targets basic secondary education 1998-2003. Relations in view. About the relations between the general targets and the attainment targets for each subject.] Zoetermeer.

Oostdam, R. (1991). *Argumentatie in de peiling*. [Assessing Argumentation.] Unpublished dissertation. University of Amsterdam.

Oostdam, R., & Rijlaarsdam, G. (1995). *Towards strategic language learning*. Amsterdam: Amsterdam University Press.

Overmaat, M. (1996). *Schrijven en lezen met tekstschema's. Effectief onderwijs in schriftelijke taalvaardigheid in de bovenbouw van het voortgezet onderwijs*. [Writing and reading with texts schemas. Effective education in written skills in the higher classes of secondary education.] Unpublished dissertation. University of Amsterdam.

Rijlaarsdam, G. (1986). *Effecten van leerlingenrespons op aspecten van stelvaardigheid*. [Effects of peer response on aspects of writing proficiency and processes.] Unpublished dissertation. University of Amsterdam.

Rijlaarsdam, G. (1987, March). *Effects of peer evaluation on writing performance, writing processes, and psychological variables.* Paper presented at the Annual Meeting of the Conference on College Composition and Communication, Atlanta.

Rijlaarsdam, G., & Couzijn, M. (2000a). Stimulating awareness of learning in the writing curriculum. In G. Rijlaarsdam & E. Espéret (Serie Eds.) & A. Camps & M. Milian (Vol. Eds.), *Studies in writing: Vol. 6. Metalinguistic activity in learning to write* (pp. 167-202). Amsterdam: Amsterdam University Press.

Rijlaarsdam, G., & Couzijn, M. (2000b). Writing and learning to write: A double challenge. In R. J. Simons, J. van der Linden & T. Duffy (Eds.), *New Learning* (pp. 157-189). Dordrecht: Kluwer Academic Publishers.

Rijlaarsdam, G., & Van den Bergh, H. (1996). The dynamics of composing – An agenda for research into an interactive compensatory model of writing: many questions, some answers. In C.M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 107-125). Hillsdale, N.J.: Erlbaum.

Rijlaarsdam, G., Van den Bergh, H., & Zwarts, M. (1992). Incidentele transfer bij produktieve taalopdrachten: een aanzet tot een baseline. [Incidental transfer on productive language tasks: An initiation for a baseline.] *Tijdschrift voor Onderwijsresearch, 17*, 55-66.

Rosenthal, T. L., & Zimmerman, B. J. (1978). *Social learning and cognition*. New York: Academic Press.

Scardamaila, M., & Bereiter, C. (1986). Research on written composition. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 778-803). New York: Collier Macmillan.

Schoonen, R., & De Glopper, K. (1996). Writing performance and knowledge about writing. In G. Rijlaarsdam, H. van den Bergh & M. Couzijn (Eds.), *Theories, models & methodology in writing research* (pp. 87-107). Amsterdam: Amsterdam University Press.

Schriver, K. A. (1991). Plain language through protocol-aided revision. In E. R. Steinberg (Ed.), *Plain language: Principles and practice* (pp. 148-172). Detroit, MI: Wayne State University Press.

Schriver, K. A. (1992). Teaching writers to anticipate readers' needs: A classroom-evaluated pedagogy. *Written Communication, 9* (2), 179-208.

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research, 57*, 149-174.

Schunk, D. H. (1991). *Learning theories: An educational perspective*. New York: Merill.

Schunk, D. H. (1995, March, April). *Social origins of self-regulatory competence: The role of observational learning through peer modeling.* Paper presented at the Biennial Meeting of the Society for Research in Child Development. Indianapolis, IN.

Schunk, D. H. (1998). Teaching elementary students to self-regulate practice of mathematical skills with modeling. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning, from teaching to self-reflective practice* (pp. 137-159). New York: The Guilford Press.

Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology, 77*, 313-322.

Schunk, D. H., & Hanson, A. R. (1989a). Influence of peer-model attributes on children's beliefs and learning. *Journal of Educational Psychology, 81*, 431-434.

Schunk, D. H., & Hanson, A. R. (1989b). Self-modeling and children's cognitive skill learning. *Journal of Educational Psychology, 81*, 155-163.

Schunk, D. H., Hanson, A. R., & Cox, P. D. (1987). Peer model attributes and children's achievements behaviors. *Journal of Educational Psychology, 79*, 54-61.

Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist, 32*, 195-208.

Schunk, D. H., & Zimmerman, B. J. (Eds.). (1994). *Self-regulation of learning and performance: Issues and educational applications*. Hillsdale, N.J.: Erlbaum.

Schunk, D. H., & Zimmerman, B. J. (Eds.). (1998). *Self-regulated learning, from teaching to self-reflective practice*. New York: The Guilford Press.

Smagorinsky, P. (1989). The reliability and validity of protocol analysis. *Written Communication, 6* (4), 463-479.

Smagorinsky, P. (1994). Think-aloud protocol analysis: Beyond the black box. In P. Smagorinsky (Ed.), *Speaking about writing. Reflections on research methodology* (pp. 3-19). London: Thousand Oaks.

Sonnenschein, S., & Whitehurst, G. J. (1983). Training referential communication skills: The limits of success. *Journal of Experimental Child Psychology, 35*, 426-436.

Sonnenschein, S., & Whitehurst, G. J. (1984). Developing referential communication: A hierarchy of skills. *Child Development, 55*, 1936-1945.

Steehouder, M., Jansen, C., Maat, K., Van der Staak, J., De Vet, D., Witteveen, M., & Woutstra, E. (1999). *Leren communiceren. Handboek voor mondelinge en schriftelijke communicatie*. [Learning to communicate. Handbook for oral and written communication.] Groningen: Wolters-Noordhoff.

Torrance, M. (1996). Is writing expertise like other kinds of expertise? In G. Rijlaarsdam, H. van den Bergh & M. Couzijn (Eds.), *Theories, models & methodology in writing research* (pp. 3-9). Amsterdam: Amsterdam University Press.

Van den Bergh, H., & Meuffels, B. (2000). Schrijfvaardigheden en schrijfprocessen [Writing skills and writing processes.] In A. Braet (Red.), *Taalbeheersing als communicatiewetenschap: Een overzicht van theorievorming, onderzoek en toepassingen* (pp. 122-153) [Applied linguistics as communication science: An overview of development of a theory, research and applications.]. Bussum: Coutinho.

Van den Bergh, H., & Rijlaarsdam, G. (1996). The dynamics of composing: Modeling writing process data. In C.M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 207-232). Hillsdale, N.J.: Erlbaum.

Van den Bergh, H., & Rijlaarsdam, G. (1999). The dynamics of idea generation during writing: An online study. In G. Rijlaarsdam & E. Espéret (Series Eds.) & M. Torrance & D. Galbraith (Vol. Eds.), *Studies in Writing: Vol. 3. Knowing what to write: Cognitive perspectives on conceptual processes in text production* (pp. 99-120). Amsterdam: Amsterdam University Press.

Van den Bergh, H., & Rijlaarsdam, G. (2001). Changes in cognitive activities during the writing process, and relations with text quality. *Educational Psychology, 21*,(4), 373-385.

Van den Bergh, H., Rijlaarsdam, G., & Breetvelt, I. (1994). Revision process and text quality: An empirical study. In G. Eigler & Th. Jechle. *Writing: Current trends in European research* (pp. 133-148). Freiburg: Hochschul Verlag.

Van den Bergh, H., Rijlaarsdam, G., & Kuhlemeier, H. (1997, August). *Attitudes versus achievements: An empirical study to the relation between writing and speaking apprehension and writing and speaking skills*. Paper presented at the 7th European Conference for Research on Learning and Instruction, Athens, Greece.

Van der Hoeven, J. (1997). *Children's composing. A study into the relationships between writing processes, text quality, and cognitive and linguistic skills*. Amsterdam: Atlanta.

Van Eemeren, F. H., & Grootendorst, R. (1992). *Argumentation, communication, and fallacies. A pragma-dialectical perspective*. Hillsdale, N.J.: Erlbaum.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*, 329-339.

Zimmerman, B. J. (2000). Attaining self-regulation. A social cognitive perspective. In: M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego: Academic Press.

Zimmerman, B. J., & Schunk, D. H. (1989). *Self-regulated learning and academic achievement: Theory, research, and practice*. New York: Springer-Verlag.

AUTHOR INDEX

SAMENVATTING

Dit proefschrift gaat over *leren-door-observeren*. Het bestaat naast een inleiding en een discussie uit vier hoofdstukken die verslag doen van vier studies: een post-hoc studie gebaseerd op de data van Couzijn (1995), twee experimentele studies en een case-studie. Met dit proefschrift wil ik bijdragen aan een dieper inzicht in de theorie en de effecten van deze didactiek voor leerlingen in de basisvorming bij het leren schrijven van argumentatieve teksten.

HOOFDSTUK 1

In het eerste hoofdstuk zet ik uiteen waarom voor de didactiek van leren-door-observeren gekozen is en schets ik achtergronden van leren-door-observeren. Vervolgens wordt kort het theoretisch kader over de effectiviteit van leren-door-observeren gepresenteerd. Hoofdstuk 1 besluit met de beschrijving van de thema's die onderzocht worden in dit proefschrift.

Een probleem bij schrijven is cognitieve overbelasting (zie bijvoorbeeld Alamargot & Chanquoy, 2001; Flower & Hayes, 1980). Dit betekent dat schrijvers te veel processen tegelijkertijd uit moeten voeren of aandacht moeten besteden aan te veel verschillende tekstuele kenmerken. Daardoor raken zij het spoor bijster, en produceren een zwakke tekst. Het probleem van cognitieve overbelasting speelt vooral bij schrijvers die *leren* schrijven, omdat zij in feite twee taken tegelijkertijd uitvoeren: een schrijftaak en een leertaak. Ze proberen een goede tekst te schrijven, een tekst zoals die van hen verwacht wordt. Tegelijkertijd wordt van de leerlingen verwacht dat zij leren van de taakuitvoering. Maar daar komen vooral zwakkere schrijvers niet aan toe, omdat zij alle aandacht nodig hebben voor de tekst zelf. Verschillende instructies (bijvoorbeeld leerlingfeedback) zijn ontwikkeld en beproefd om leerlingen te stimuleren uit de schrijftaak te stappen en aandacht te besteden aan reflectie op de schrijftaak. Ook de didactiek *leren-door-observeren* beoogt die reflectie te stimuleren. Maar het meest kenmerkende van deze didactiek is dat leerlingen niet leren schrijven door zelf teksten te schrijven, maar door schrijfprocessen en schrijfproducten van andere schrijvers te observeren. Doordat zij niet zelf schrijven, is de cognitieve belasting minder groot en kunnen zij zich beter richten op de leertaak (het leren schrijven).

Uit praktisch oogpunt lijkt de didactiek van leren-door-observeren een goede kandidaat om de procesgerichte aanpak van het schrijfonderwijs bij het schoolvak Nederlands vorm te geven die voorgeschreven wordt in de basisvorming van het voortgezet onderwijs (Ministerie van OCenW, 1998). Bij deze aanpak ligt het accent op het schrijfproces en schrijfstrategieën, moet er reflectie plaatsvinden op de schrijfcontext, het schrijfproces en schrijfproduct, en moeten teksten becommentarieerd en besproken worden, met de leraar en met medeleerlingen (Hoogeveen & Bonset, 1998).

In het theoretisch kader van mijn proefschrift spelen verschillende elementen en invloeden een rol. Kort samengevat: leren-door-observeren stimuleert het gebruik van metacognitieve strategieën als observatie, evaluatie en reflectie en richt de aandacht op de leertaak. Daardoor verandert de kennis van leerlingen over schrijven. Die kennis beïnvloedt de organisatie van het schrijfproces, waardoor de kwaliteit van het resulterende schrijfproduct toeneemt.

De studies in dit proefschrift sluiten aan bij het onderzoek van Couzijn (1995, 1999). Hij toonde aan dat leren-door-observeren effectiever was voor leren schrijven dan de traditionele vorm van schrijfonderwijs, het leren-door-doen, waarin leerlingen steeds oefeningen maken. Couzijn stelde de effecten van leren-door-observeren op schrijfproducten vast. Hij richtte zich niet op onderliggende processen. Het belangrijkste doel van dit proefschrift is om meer inzicht te krijgen in deze onderliggende processen. Ik richt me hierbij op drie thema's:

- *Het observatieproces*. Observeren is zelf een complexe activiteit, die bestaat uit verscheidene deelactiviteiten. De vraag is welk van die activiteiten vooral bijdraagt aan leerwinst: welke elementen uit de observatietaak dragen bij aan de leerwinst in de studie van Couzijn? (Studie 1, hoofdstuk 2). En meer in het algemeen: hoe doen leerlingen dat eigenlijk, dat observeren? (Studie 4, hoofdstuk 5).
- *Relatie tussen leerlingkenmerken en de effectiviteit van instructies voor leren-door-observeren*. Uit onderzoek is bekend dat de effectiviteit van leren-door-observeren afhankelijk is van verschillende factoren (Schunk, 1987, 1991, 1998). Couzijn richtte zich op een relatief homogene groep leerlingen (vrij hoog niveau) en onderzocht één instructietype voor leren-door-observeren. Leerlingen observeerden steeds twee modellen en reflecteerden op het zwakkere model van ieder paar. Men kan zich afvragen of alle leerlingen leren van reflectie op het zwakkere model, vooral wanneer het goede leerlingen betreft. Daarom wordt in studie 2 (hoofdstuk 3) onderzocht welk type instructie effectief is voor welk type leerling.
- *Effect van leren-door-observeren op de tussenliggende variabele 'schrijfprocessen'*. Couzijn vond effecten van leren-door-observeren op schrijfproducten. Hij veronderstelde dat deze effecten werden veroorzaakt door veranderingen in schrijfprocessen, maar onderzocht deze veronderstelling niet. In mijn derde studie heb ik deze veronderstelling nader onderzocht (hoofdstuk 4).

HOOFDSTUK 2

In hoofdstuk 2 rapporteer ik een post-hoc studie waarin data van Couzijn (1995) worden geanalyseerd. Om de effectieve elementen uit de observatietaak te identificeren zijn verschillende materialen (voortoets, werkboekjes, natoetsen) van 84 leerlingen onderzocht. Ik richtte me op twee elementen van de observatietaak: evaluatie en elaboratie op deze evaluatie. Van drie verschillende leeractiviteiten konden de sporen in de werkboekjes worden teruggevonden: evaluatie (bepalen welk model de taak het best uitvoerde), product-elaboratie (uitleggen waarom producten van modellen wel of niet goed zijn) en proces-elaboratie (uitleggen waarom processen van modellen wel of niet goed zijn).

De resultaten lieten zien dat de leeractiviteiten vrij stabiel zijn over verschillende lessen heen. Leerlingen die deze activiteiten beter (evaluatie en product-elaboratie) of vaker (proces-elaboratie) uitvoerden in de ene les over het ene onderwerp, voerden deze activiteiten ook beter of vaker uit in de volgende les over een verwant, maar ander onderwerp. Dit betekent dat onafhankelijk van de precieze lesinhoud, de leeractiviteiten stabiel zijn en consistent gemeten: een product-elaboratie in les 1 is hetzelfde als een product-elaboratie in de tweede les. Tevens bleek dat de uitvoering van evaluatie en product-elaboratie samenhing met de beginsituatie van leerlingen: leerlingen die hoger scoorden op de voortoets, scoorden ook hoger op de leeractiviteiten. Dit betekent dat sommige leerlingen de gewenste leeractiviteiten al beheersen of gemakkelijk aanleren. De leeractiviteiten zijn dus geen wezensvreemde activiteiten, maar activiteiten die goede leerlingen (gedeeltelijk) uit zichzelf ontwikkelen. Ten derde bleek dat evaluatie en product-elaboratie bijdragen aan de effectiviteit van leren-door-observeren. Hoe vaker leerlingen de uitvoering van de modellen correct evalueerden en de producten van de modellen juist becommentarieerden, hoe beter zij scoorden op aspecten van het schrijven van argumentatieve teksten.Van proces-elaboratie had ik ook effecten op schrijfvaardigheid verwacht. Maar van proces-elaboratie waren zo zelden sporen te vinden in de werkboekjes van leerlingen, dat het moeilijk werd om effecten (als die er waren) aan te tonen.

Via deze post-hoc studie is meer inzicht in een aspect van het observatieproces verkregen. We weten nu dat evaluatie en product-elaboratie effectieve elementen in observatietaken zijn: door het evalueren van activiteiten van modellen en door uit te leggen waarom producten wel of niet goed zijn, schrijven leerlingen zelf betere teksten.

HOOFDSTUK 3

Dit hoofdstuk beschrijft een experimentele studie. Hierin onderzoek ik welk type instructie effectief is voor welk type leerling. In deze studie volgden 214 leerlingen (tweede klas vbo-vwo) in twee instructiesessies een cursus over argumentatief schrijven. Iedere klas werd toegewezen aan een van drie condities: twee observatiecondities of een controleconditie. In de beide observatiecondities observeerden leerlingen modelparen (medeleerlingen) die schrijftaken uitvoerden. Leerlingen werd in de ene conditie gevraagd te reflecteren op het zwakke model, in de andere conditie op het goede model. In de controleconditie voerden leerlingen zelf de schrijftaken uit.

De resultaten toonden aan dat de ene instructiemethode effectiever is voor het ene type leerling, en een andere voor het andere type leerling. Welk van de instructiemethoden effectiever is, hangt ook af van de plaats in instructiesequentie. Wanneer de taak *nieuw* was (in instructiesessie 1) leerden zwakke leerlingen meer van leren-door-observeren (mits reflecterend op zwakke modellen) dan van het zelf uitvoeren van de schrijftaak. Dit zou veroorzaakt kunnen worden doordat deze zwakke leerlingen nu wel metacognitieve strategieën als observeren, evalueren en reflecteren toepassen en doordat hun cognitieve inspanningen nu verschoven zijn van de uitvoering van schrijftaken naar leren. Goede leerlingen daarentegen profiteerden bij nieuwe taken niet alleen van leren-door-observeren (mits reflecterend op goede modellen) maar ook van het zelf uitvoeren van schrijftaken. Zij zijn blijkbaar wel in staat om hun aandacht te verdelen tussen schrijven en leren; zij kunnen kennelijk leren van het evalueren van en reflecteren op hun eigen uitvoering. Wanneer de taak meer vertrouwd was (instructiesessie 2), leerden zwakke leerlingen evenveel van het uitvoeren van schrijftaken als van het leren-door-observeren met reflectie op het zwakke model. Omdat de leerlingen al (wat) successen met de taak hebben geboekt, hebben ze waarschijnlijk kennis over de taak opgebouwd en waren ze dus beter uitgerust om te leren van het zelf uitvoeren van schrijftaken. Goede leerlingen profiteerden wanneer de taak bekend is, alleen van leren-door-observeren met reflectie op het goede model. Niet van observeren met reflectie op het zwakke model of van het zelf uitvoeren van schrijftaken. Misschien hebben ze de uitdaging nodig om uit te leggen wat het goede model goed deed. Voor gemiddelde leerlingen bleek het niet uit te maken volgens welke instructie zij geleerd hadden. Dit gold voor beide instructiesessies. Zij leerden evenveel van leren-door-observeren (in beide condities) als van het zelf uitvoeren van schrijftaken.

HOOFDSTUK 4

Het vierde hoofdstuk beschrijft een experimentele studie, waarin de effecten van leren-door-observeren op schrijfprocessen worden onderzocht. 52 leerlingen uit de tweede klas (vbo-vwo) volgden een korte cursus over argumentatief schrijven in drie verschillende condities: twee observatiecondities en een controleconditie. De condities waren dezelfde als de condities in de vorige studie: in de beide observatiecondities observeerden leerlingen modelparen (medeleerlingen) die schrijftaken uitvoerden, en moesten leerlingen reflecteren op het zwakke model, of op het goede model. In de controleconditie voerden leerlingen zelf de schrijftaken uit. Na het volgen van de cursus maakten leerlingen schrijftoetsen waarin zij korte argumentatieve teksten schreven. Deze teksten schreven zij hardop denkend. Op deze manier kreeg ik inzicht in hun schrijfprocessen en kon ik de schrijfprocessen van leerlingen uit verschillende condities vergelijken.

Uit de resultaten bleek dat leerlingen die hebben geleerd door observeren een ander *schrijfpatroon* vertoonden dan leerlingen die leerden door doen. Leerlingen uit de observatiecondities ondernamen meer metacognitieve activiteiten (bijvoorbeeld oriëntatie op het doel van de opdracht en analyse van de tekst) in het begin van het schrijfproces en meer uitvoerende activiteiten (zoals schrijven en herlezen) in het tweede deel van het schrijfproces. Bovendien maakten deze leerlingen in het tweede en laatste deel van het schrijfproces meer gebruik van meta-analyse activiteiten (bijvoorbeeld: ‘Zo, nu eens even kijken of alles in de tekst staat’). Deze activiteiten zijn een indicatie voor regulerende processen. Tijdens het hele schrijfproces lieten de leerlingen uit de observatiecondities meer planningsactiviteiten zien. Daarentegen voerden leerlingen uit de controleconditie meer schrijfactiviteiten uit in het begin van het schrijfproces en meer formuleringsactiviteiten gedurende het hele schrijfproces. Een ander resultaat is dat leerlingen die hebben geleerd door observeren, voor bepaalde activiteiten een verloop in hun uitvoering lieten: zij voerden bijvoorbeeld in het begin van het schrijfproces meer analyse-activiteiten uit dan aan het einde van het schrijfproces. Leerlingen uit de controleconditie daarentegen lieten een monotoon proces zien: zij voerden dezelfde hoeveelheid analyse-activiteiten uit tijdens het hele schrijfproces. Ook bleek dat de processen van leerlingen in de observatiecondities veel minder homogeen waren dan de processen van leerlingen in de controleconditie. Dit wijst er volgens mij op dat binnen de observatiecondities er leerlingen waren die hun routines overboord zetten, en aan het leren waren.

De gerapporteerde effecten van leren-door-observeren op schrijfprocessen zijn in overeenstemming met het theoretisch kader. Door het verplaatsen van cognitieve inspanningen van schrijven naar *leren* schrijven en door het gebruiken van metacognitieve strategieën, ontwikkelen leerlingen die geleerd hebben door observeren rijkere kennis over schrijven. Deze kennis beïnvloedt de uitvoering van schrijfprocessen in de richting die ik vond: meer planning en meer analyse. Daarnaast is het goed mogelijk dat de afwijkende schrijfprocessen van de observerende leerlingen veroorzaakt zijn doordat deze leerlingen verschillende voorbeelden van schrijfprocessen geobserveerd hebben. De medeleerlingen die zij observeerden, dachten immers hardop tijdens het schrijven en boden zo inzicht in effectieve en minder effectieve schrijfprocessen.

HOOFDSTUK 5

In hoofdstuk 5 beschrijf ik een case-studie die uitgevoerd werd om inzicht te krijgen in hoe leerlingen observatietaken uitvoeren. In deze studie leerden zes leerlingen (derde klas havo/vwo) in een korte cursus hoe ze argumentatieve teksten moesten schrijven. In deze cursus pasten ze theorie toe in observatietaken. Om te kunnen volgen hoe observeren in zijn werk gaat, vroeg ik de leerlingen om tijdens de cursus hardop te denken. Na het volgen van de cursus werden de leerlingen geïnterviewd over hun observatiegedrag. Omdat ik het observatiegedrag – ter verdieping van de tweede en de derde studie – in twee verschillende condities wilde onderzoeken, volgden de leerlingen de cursus in twee observatiecondities. In beide condities observeerden de leerlingen twee modellen (een goed en een zwakker model) die schrijftaken uitvoerden, en in beide condities moesten de leerlingen bepalen welk model de taak het beste uitvoerde. Het verschil tussen de condities bestond uit de focus op de modellen bij de elaboratie-vraag. Leerlingen in de zwak/focus conditie moesten daar uitleggen wat het zwakkere model minder goed deed, leerlingen in de goed/focus conditie wat het goede model goed deed.

De resultaten lieten zien dat observerende leerlingen veel verschillende (meta)cognitieve activiteiten uitvoerden, zoals het lezen en analyseren van de argumentatiestructuur waarvan de modellen een tekst moesten schrijven, observeren, vergelijken, evalueren en het beoordelen van de modellen en commentaar geven op het schrijven van de modellen (processen en producten). De uitvoering van deze activiteiten laat zien dat de observanten criteria voor goed schrijven hebben ontwikkeld, toegepast en geïnternaliseerd. Het bleek ook dat de observerende leerlingen (mentale) uitvoerende activiteiten uitvoerden en dat zij tijdens het observeren deze eigen (interne) schrijfuitvoering vergeleken met de uitvoering van de modellen. Deze mentale schrijfuitvoering van de observanten kost wel wat cognitieve energie maar veel minder dan de energieslurpende ‘echte’ schrijftaak. Deze interne schrijftaak is ‘genest’ in de observatietaak: het mentale schrijven van de observanten is een middel om gericht te zijn op de leertaak, het is geen schrijftaak die voltooid moet worden. Bovendien – en overeenkomstig met de resultaten uit studie 1 (hoofdstuk 2) – werd duidelijk dat evaluatie- en elaboratie-activiteiten belangrijke activiteiten zijn. De observanten startten al met evalueren en elaboreren (refereren aan de schrijfprocessen en/of producten van de modellen) tijdens het observeren van de modellen. Ze gaven daarbij vooral commentaar op de argumentatie in de teksten van de modellen (bijvoorbeeld: ‘Model 1 verwisselt het standpunt met het argument’) en op de gebruikte signaalwoorden/ verbindingswoorden (bijvoorbeeld: ‘Model 2 had “dus” moeten gebruiken in plaats van ‘want”’). Het bleek dat de observanten deze evaluaties en elaboraties ook weer (letterlijk) gebruikten bij het beantwoorden van de evaluatie- en elaboratievraag nadat ze de teksten van de modellen gelezen hadden. Hieruit kan geconcludeerd worden dat de procesobservaties een belangrijke rol spelen in observatietaak, de geobserveerde schrijfprocessen verschaffen de observanten genoeg informatie om het gedrag van de modellen te kunnen evalueren en uit te leggen wat er goed of fout is aan de geobserveerde schrijfuitvoering.

Ten slotte liet de case-studie een verschil in gedrag zien tussen de beide observatiecondities. Leerlingen die op de schrijfuitvoering van het goede model moesten reflecteren (conditie goed/focus), reflecteerden vaak op beide modellen, dus ook op de uitvoering van het zwakkere model. Leerlingen uit de zwak/focus conditie reflecteerden vaker alleen op het bedoelde (zwakke) model. Ik veronderstel dat het moeilijker is om *alleen* op het betere model te reflecteren omdat uitleggen waarom een model de taak goed heeft uitgevoerd een behoorlijk ontwikkeld bewuste kennis over ‘goed’ schrijven vereist, en een expliciete set van criteria over ‘goed’ schrijven die de observanten kunnen toepassen op de taakuitvoering van het goede model. Deze veronderstelling werd bevestigd door leerlingen uit goed/focus conditie. Zij zeiden in het interview dat ze het moeilijk vonden om uit te leggen wat het goede model goed deed en dat ze liever wilden uitleggen wat het zwakkere model minder goed deed. Waarschijnlijk kunnen observanten die op het zwakke model reflecteren de uitvoering van het goede model als referentie gebruiken, het goede model levert dan de criteria voor een goede taakuitvoering.

HOOFDSTUK 6

In het slothoofdstuk vat ik de uitkomsten van de vier studies samen, vergelijk ik een deel van deze resultaten met de uitkomsten van het verwante onderzoek van Couzijn (1995, 1999), reflecteer ik op de uitgevoerde studies, geef ik suggesties voor vervolgonderzoek en noem ik enkele implicaties voor de onderwijspraktijk.

In tegenstelling tot Couzijn heb ik geen hoofdeffect gevonden van leren-door-observeren. ‘Gemiddelde’ leerlingen uit mijn onderzoek leerden evenveel van het observeren van schrijftaken als van het zelf uitvoeren van schrijftaken. Hoewel mijn onderzoek geïnspireerd is door en gebaseerd is op Couzijn's onderzoek, verhinderen verschillende factoren echter een precieze vergelijking van mijn onderzoek met dat van Couzijn. Zo verschilden het niveau van de proefpersonen, de setting, lesinhoud, aantal lessen en procedures. Vervolgens bespreek ik in hoofdstuk 6 een mogelijke alternatieve verklaring voor de gevonden effecten van leren-door-observeren, reflecteer ik op de vrij kleine – maar consistente – effectgroottes, en overdenk ik het experimentele design. Ten slotte ga ik dieper in op de implementatie van de condities. In studie 4 (hoofdstuk 5) vond ik immers dat het werkelijke gedrag van de leerlingen soms afwijkt van de instructie. Dit gold vooral voor de conditie waarin leerlingen op het goede model moesten letten. Deze conditie nodigt leerlingen kennelijk uit om van de instructie af te wijken en op beide modellen te letten. Gesteld wordt dat deze afwijking ook de interpretatie van de interactie-effecten tussen leerderskenmerken en observatiecondities (studie 2, hoofdstuk 3) heeft kunnen beïnvloeden. De gevonden effecten zouden gedeeltelijk veroorzaakt kunnen zijn door reflectie op het 'verkeerde' model. Deze gedachte zou onderzocht moeten worden in een vervolgonderzoek waarin leerlingen *alleen* goede of *alleen* zwakke modellen observeren.

Een suggestie voor vervolgonderzoek is de verificatie van de veronderstelde invloeden in het theoretisch kader. In dat theoretisch kader neemt 'kennis over schrijven' een vrij grote rol in. Ik veronderstel dat leren-door-observeren de kennis over schrijven vergroot, en dat deze rijkere kennis de organisatie van het schrijfproces beïnvloedt, maar ik heb dit in mijn studies niet onderzocht. Daarom zou in vervolgonderzoek (naast de andere invloeden) allereerst de effecten van leren-door-observeren op kennis over schrijven onderzocht moeten worden. Andere suggesties voor vervolgonderzoek zijn gericht op nieuwe typen instructies voor leren-door-observeren, onderzoek naar type observatietaken (deductief versus inductief), onderzoek naar de effecten van leren-door-observeren op grotere en meer communicatieve schrijftaken, transfereffecten van leren-door-observeren, en de rol van leren-door-observeren in reguliere lessen op school. Een ander interessant punt voor vervolgonderzoek is verder onderzoek naar de sequentie voor instructies. Omdat uit de tweede studie (hoofdstuk 3) bleek dat er interacties zijn tussen bekendheid met de taak en type instructie voor verschillende type leerlingen (niveau), is het zinvol om te onderzoeken welke sequentie nu ideaal is. Gebaseerd op de resultaten uit studie 2 zou zo'n sequentie er als volgt uit kunnen zien: (1) wanneer een leerling zwak is en de taak is nieuw, geef dan observatietaken met de focus op het zwakke model. (2) Wanneer de taak bekend is voor deze (nog steeds zwakke) leerling, dan leert hij naast leren-door-observeren met reflectie op het zwakke model, ook van het zelf uitvoeren van schrijftaken. (3) De laatste stap is wanneer de taak nog steeds bekend is, maar wanneer de leerling een goede leerling geworden is. Dan is de meest effectieve instructievorm voor deze leerling het observeren en reflecteren op het goede model. Onderzoek moet uitwijzen of deze sequentie inderdaad effectief is. Overigens is het – in het kader van zelfstandig leren – ook interessant om te onderzoeken wat de effecten zijn als leerlingen zelf de instructievorm kiezen.

Voor de onderwijspraktijk pleit ik voor een afwisseling van de 'gebruikelijke' schrijftaken met observatietaken waarin leerlingen de docent of medeleerlingen observeren terwijl deze laatstgenoemden hardop denkend aan het schrijven zijn. Vooral wanneer leerlingen alleen (en later plenair) of meteen al in tweetallen of groepjes, de taakuitvoering van de modellen evalueren en verwoorden wat erg goed en minder goed ging en suggesties doen voor een betere uitvoering, zijn leerlingen relatief gemakkelijk bezig met evaluatieve en reflectieve activiteiten. Bovendien is de aandacht op een natuurlijke manier gericht op het schrijfproces. Wanneer de docent een script gebruikt voor de schrijfuitvoering dat gebaseerd is op echte problemen van leerlingen kan hij heel effectief de problemen behandelen waarmee leerlingen worstelen. Voor leerlingen kan het bovendien ook erg inzichtelijk en leerzaam zijn om scripts voor een goede en minder goede taakuitvoering te construeren. Ten slotte suggereer ik dat leren-door-observeren ook voor andere vakken in de basisvorming een goede methode zou kunnen zijn voor actief leren.

CURRICULUM VITAE

Martine Anne Henriëtte Braaksma was born on June 27th 1971 in Choma (Zambia) and completed her secondary schooling in 1990 in Hardenberg (the Netherlands). She studied Dutch Language and Culture at the University of Groningen and graduated in 1995 specializing in Speech Communication and Discourse Analysis. Subsequently she attended teacher training at the University of Groningen and graduated cum laude in 1996.

Martine took her PhD at the Graduate School of Teaching and Learning of the University of Amsterdam during the period 1996-2002 in the field of writing education. From 1997 she combined research with teaching various writing courses and courses in methodology of Dutch Language and Culture at the Graduate School of Teaching and Learning. Furthermore, she participated in the organization of the conferences of the International Association for the Improvement of Mother Tongue Education (IAIMTE) in 1997, 1999, and 2001.

Next to her teaching work, Martine currently works at the Graduate School of Teaching and Learning as a researcher in projects on 'Observational learning in language education' and 'Literary reading processes'.