Observation of reading and writing activities: Effects on learning and transfer

Michel Couzijn

University of Amsterdam
Graduate School of Teaching and Learning

dissertation

Promotor: B.H.A.M. van Hout-Wolters
Co-promotor: G.C.W. Rijlaarsdam

september 1995

= CONTENTS =

CHAPTER 4: Learning to write by reader observation and written feedback: Instructive texts

CHAPTER 5: Observing readers and writers of argumentative text: Learning to write – transfer to reading

CHAPTER 6: Observing readers and writers of argumentative text: Learning to read – transfer to writing

CHAPTER 7: Observations: Looking back and looking ahead

CHAPTER 1: Introduction

The aim of this study is to enhance the effectivity of writing and reading instruction. Two ways are chosen for this purpose. First, by detecting and overcoming weaknesses in the traditional learning–by–doing approach to writing and reading instruction. Second, by investigating possibilities to promote transfer between the writing and reading modes. Two alternative methods of instruction are sketched which will be central in the empirical investigations. Finally, the aims and scope of this study are discussed.

1.1 Reading and writing in the language skills curriculum

When people are asked what their children should learn at school – and one does not necessarily have to add: in the language class – almost all the answers contain these two words: 'reading' and 'writing'. Ever since writing came into fashion, about 5000 years ago among the Egyptians, there have been people who wanted to learn how to read those funny little signs, carved in wood or scratched in stone. Others would say, however, that ever since Man had learned to read, about 5000 years ago among the Egyptians, there have been people who wanted to know how to carve these funny little signs in wood or scratch them in stone so the others would see them. Since then, literacy has become the greatest round game ever invented in this world. Literacy is a true game, with rules, prizes, winners, losers, strategies – and it takes a lot of experience to play the game well. In fact, in societies all over the world, the game has become so important in people's lives, the prizes have got so high, the losses so grave, and the rules so complex, that it was decided that everyone should devote a considerable part of their early life to learning to read and write.

That is still the case in the world today. The use of written language is an indispensable tool in many sectors of professional and social life. To be fully prepared, young people spend the best years of their lives in the school benches learning the rules of a rather serious reading and writing game. If measured by learning time, reading and writing are the most important parts of the language curriculum: together they take up almost half of the instruction time (Otter, 1987). Thus, since so many people are involved in learning to read and write, and so much time is invested by each of them, and the results of this learning are so important to their lives, we may conclude that the development of effective teaching and learning methods for reading and writing should have our concern. The present study wants to contribute to such development.

The problem addressed in this study concerns the relationship between writing and reading. More specifically, it addresses how this relationship can be used to enhance the effectivity of writing and reading instruction in secondary education. Skills and processes of (non–elementary) reading and writing have been studied much more often in isolation than in combination (for reading e.g. Van Dijk & Kintsch, 1983; Just & Carpenter, 1987; Rayner & Pollatsek, 1989; for writing e.g. Hillocks, 1986; Bereiter & Scardamalia, 1987). Besides, more studies are aimed at reading and writing processes than at learning–to–read and learning–to–write processes. As a consequence, there are comparably few studies that directly address the interplay of learning–to–read and learning–to–write processes.

In this study, I will elaborate on two potential weak spots in current reading and writing instruction. One of them has to do with requirements for *learning,* the other with requirements for *transfer–of–learning*. With 'learning' I refer to the change in a learner's behavior potencies regarding a certain domain, as a result of experience or practice *within that domain*. With 'transfer', I refer to the change regarding a certain domain as caused by experience or practice within *another domain.* Reading and writing are considered in this study as different domains, distinguishable though not distinct.

The distinction between within–mode learning and between–mode transfer is maintained here for the sake of clarity. It is understood that transfer can in fact also take place *within* each of these domains (e.g. when someone's learning to write fiction influences his writing behavior for non-fiction) but such within–mode transfer is not addressed in this study.

The learning 'weak spot' of current reading and writing education has to do with the way in which students acquire complex cognitive skills. The instructional method underlying much of current language skill education is based on the credo 'you learn to do something by doing it'. In this study, several arguments – theoretical as well as empirical – will be advanced against this credo. It will be advocated that the acquisition of complex skills in a *Learning–by–doing* situation often asks too much of the learner's self–monitoring and self–evaluation capabilities.

The transfer 'weak spot' of current reading and writing education lies in the strong neglect of interrelationships between reading and writing. While there is more than only intuitive evidence for the existence of such interrelationships (e.g. Shanahan & Lomax, 1986; Simonsen, 1988; Pool & Van Wijk, 1995), they are hardly ever identified or used to design curricula, or to develop instruction that enhances (and profits from) cross–fertilization. Reading and writing instruction is hardly ever more than the sum of its parts: reading instruction and writing instruction. If the instruction reflects such unrealistic separation between the skills, it is likely to affect the acquisition of learners' cognition. It may be the most important cause why so many students do not apply in their writing (e.g. paragraph construction) what they have learned in a reading context (e.g. paragraph analysis).

In this study, the enhancement of instructional effectiveness is pursued by examining additional or alternative activities for Learning–by–doing, and by using the transfer potential between the reading and writing modes. Promoting such transfer may have two benefits. First, more learning goals may be achieved within the same time (or the same goals in shorter time). Second, positive effects on the integration of acquired knowledge may be realized, a benefit which I would call even more important. The motivation for this study can now be formulated: 1) to identify additional or alternative learning activities which enhance the effectivity of Learning–by–doing within the writing or the reading mode, and 2) to investigate possibilities for transfer between the writing and reading domains. As will be explained in chapter 2, possibilities for transfer will be sought in selecting domain elements that are susceptible to transfer, and by identifying factors in the instruction that promote the actual transfer.

As a starting point for the investigation, I will present in this chapter several points of criticism on a traditional type of Learning–by–doing. Then I will move on to a prospection of other types of instruction that may be a good alternative for, or supplement to, Learning–by–doing. Finally the scope of this study and the parts of this book will be explained.

1.2 Criticism on Learning–by–doing in language skill instruction

A traditional method of instruction for reading and writing in secondary school is based on the effect of doing exercises. This method, which I will call 'Learning–by–doing', can be represented as the following sequence:

1. Presentation of the subject–matter: in a text book, and/or by a teacher's explanations.

For instance, 'In an explanatory text, three parts can be distinguished: introduction, body, and ending'.

2. Practice: applying the subject matter in one or more exercises.

For instance, the students must write an essay in which they should use these three parts, or they must analyze the three–part structure of existing texts.

3. Evaluation: The teacher evaluates the product (the essay or the analysis), determines a mark, and returns the product to the student. A classical discussion of the results may follow.

Notice that it is easy to recognize such a method of instruction in other secondary school subjects. It is a so–called 'teacher–centered' method (Boekaerts & Simons, 1993), in which the teacher initiates and controls the learning activities of the students. A similar, rather directive method is described by Gagné's nine 'instructional events' (Gagné et al.,1988).

Such a method of instruction can be criticized in several ways. I will give some points of criticism here with respect to a learning–to–write situation; a comparable criticism can be given with respect to learning–to–read situations.

a) Usually, <u>the exercise is very complex</u> and does not offer sufficient possibilities for the students to get well–trained in applying the subject–matter. Writing a text with the three–part structure (or analyzing such texts) takes a long time; only a small part of this time is actually spent on applying the subject matter. So during most of the time for practice, students have their attention at other affairs. With complex writing (and reading) tasks, students have to divide their attention across so many activities that they will invest comparatively little cognitive effort in what had to be learned, or even neglect it at all. Teachers often find that a subject that had recently been 'learned', appears to be 'forgotten' in a new task. It is not forgotten, of course – it's just not being used because the students' attention goes out to something else.

b) There is <u>no self–evaluation task</u> for the learner. The teacher is often the only one occupied with, and responsible for the evaluation of student performances. Thus, students are not stimulated to think about what they could have done right or wrong. This is all the more important, because many students tend to do their homework as <u>soon</u> as possible (not necessarily as <u>good</u> as possible) so they will not be over–critical towards their own work. This is a pity, because learning is supposed to take place on the edge of what a learner is still capable of, and of what he is not yet capable of. In this respect it is advisable to have the learner participate in the evaluations. They may discover the boundaries of what they know and do not know, or what they can do and cannot do (metacognition;

see Flavell, 1985) and may more consciously work on shifting these boundaries.

c) No self–correction takes place. Usually the subject matter is practised only once or twice (because the tasks take up so much time), and the mistakes that are made don't have to be corrected. The text is returned to the student, errors are marked, and a grade is given; however, making revisions, trying to do it better that you did the first time, is still quite rare. This is a missed opportunity, since self–corrections and renewed trials may well contribute to learning: finding out what went wrong ànd how it can be avoided or corrected. In learning mathematics, it is very common to re–do a sum that failed. Why should this be very different in language skills education?

d) There is quite a distance, in time and in space, between the performance of the student, the evaluation made by the teacher, and the feedback supplied to the student. For this reason, it is very hard for students to connect their writing activities, the positive or negative effects of these activities, and suggestions on how to do it better. These things are simply too far apart for the learner to recognize any connection between them.

e) Many teachers think that they, as professionals, should comment on ALL aspects of the text. In his revision, they will therefore comment on much more than only the correct or incorrect application of the subject–matter. This looks like a nice service, but such extensive commentary is not more effective than limited commentary (Hillocks, 1986). Extensive commentary can even trouble the student's focus of attention to the extent that he forgets the original subject–matter.

f) The millstone of every mother–tongue teacher: this instructional method results in a load of work for the teacher. If teachers make themselves the only ones responsible for the evaluation of the student's performance, it is their job to correct all the texts and other assignments. The essays of one group of 30 students result in almost a day's work for the teacher. Such a burden may lead to handing out fewer writing tasks than would be desirable.

g) This instructional method is evidently product–oriented and not process–oriented. The subject–matter states some quality of the final product, but no information is given on *how* one should arrive at such a product. Should writers start with the introduction? Should they first make a global outline? When and how should they revise? By involving the readers in the evaluation of their products, the recursive character of reading and writing tasks may become clear to them: they may learn that they should not pay attention to all aspects at once, but to do the task step–by–step, and insert some moments of reflection in between to see if they are still on–track.

It is not my intention to disqualify Learning–by–doing as if it were an unuseful method of instruction. On the contrary, I think that doing exercises is an indispensable tool in language skill education – as in other domains of skill education. However, an exclusive reliance on this method may hinder optimal learning due to the abovementioned weaknesses. It is my aim to investigate ways in which the learning–by–doing method can be enhanced or supplemented within the domains of reading and writing.

1.3 Two alternative methods of instruction

In chapter 3 of this study, two alternative methods of instruction will be discussed which may - theoretically - compensate for some of the abovementioned weaknesses. The hypotheses hereabout are experimentally put to the test in the empirical chapters 4, 5 and 6. In each of these chapters, a certain text type is chosen (instructive text or argumentative text) and applied in one or more tasks in either the writing or the reading mode. The concurring instructional methods are then used to construct the experimental interventions. In this section I will give a prospection by sketching these two alternative methods.

The alternatives are variants of a method called Learning–by–observation. They make essential changes to the 'Practising' stadium of Learning–by–doing. In both Learning–by–observation methods, the experimental students - instead of doing exercises themselves - are instructed to observe and comment on peer students doing exercises.

The first alternative, 'Learning by observation of models', is a learning activity in which the observer is looking at task performances of peer or age–group students, which serve as – good and bad – examples (for instance, observing how other students add structure markers to their essays). The main activity of the observer is evaluation: find the criteria in task behavior that distinguish good task behavior from weak task behavior (for instance, choosing the best of several ways of adding structure markers, and motivating the choice). The rationale for this instructional method is that it directs the observer's attention to processes or, better, qualitative differences between processes. The observer is supposed to identify with the observed student, evaluate his/her task behavior, and then to do some self–initiated reflection: 'would I do it the same way?', 'can I do it better?", 'what should I remember from what I've seen?'.

'Learning by observation as feedback', the second alternative, combines performance and observation activities. First the observer has to perform a communicative task (for instance, clarify the text structure of an essay by adding textual markers, such as 'in the first.... in the second place', 'on the other hand...', 'in sum....' etc.). Next, the observer sees a peer performing the 'communicatively complementary task' (for instance, analyzing this text structure and identifying the markers that are helpful in the analysis). The observer collects realistic responses to his/her performance, with a view to self–evaluation and revision. In observing the authentic reading and writing activities, the observer may become more aware of the needs, problems, and strategies of the communicative partner, with a view to future task performances.

Expectations about the effectivity of these two alternative instructional methods are derived from the extent to which they may solve the Learning–by–doing problems. Since the alternative methods are supposedly more process– and evaluation–oriented, contribute more to self–reflection, and are stronger focused on criteria for 'what works' in communication, these methods are expected to be advantageous in comparison with Learning–by–doing.

1.4 Aims and scope of this study

It is the starting–point of this study that there are weaknesses in the traditional Learning–by–doing approach to language skills instruction that set bounds to its effectiveness. With 'effectiveness', I mean within–domain learning effects, as well as between–domain transfer effects.

The aims of this study are to develop alternatives to, or variations on the Learning–by–doing approach, and to investigate the effectiveness of these alternatives. To this end, several tasks need to be undertaken. They are described in the following chapters:

(chapter 2) In order to promote transfer between the writing and reading domains, the nature of transfer and the factors that influence its occurrence must be understood. Therefore, a literature study about transfer–of–learning will be reported, with a view to specification of transferable elements in the domains, and to the development of transfer–promoting instruction.

(chapter 3) Chapter 2 will yield conclusions that need application or adaptation to specific domains, in this case the reading and writing domains. A learning theory is developed in which stimulation of learning and stimulation of transfer can be integrated. This learning theory, which favors 'learning by observation' over 'learning by–doing', will be exemplified. Finally, the main research questions for the empirical part of the study are presented.

(chapter 4) Empirical evaluation of Learning–by–doing and Learning–by–observation (and some variants), applied to writing instructive texts.

(chapter 5) Empirical evaluation of Learning–by–doing and Learning–by–observation (and some variants), applied to writing argumentative texts.

(chapter 6) Empirical evaluation of Learning–by–doing and Learning–by–observation (and somer variants), applied to reading argumentative texts.

(chapter 7) Conclusions of the empirical chapters are reported. Implications for theory and possible practical implications for education are discussed.

The study is interdisciplinary in that domain–independent learning and transfer theories are applied to domains, while within these domains, (pragma–)linguistic theory is used to describe better and worse performance. The theoretical scope therefore extends to the psychology of learning & instruction and to pragmalinguistics. Considering the domain–oriented background of potential readers, I have chosen to report more extensively on the transfer literature than a strictly learning psychological study would require.

Parts of this study have been previously published, either as journal articles, book chapters, or conference papers.

CHAPTER 2: Theoretical approaches to transfer of learning

At the start of this chapter, the concept 'transfer' is defined and the most important distinctions in transfer theory are discussed. An overview of theoretical approaches, historical as well as modern, is given, with special attention to two recent exponents of, respectively, a task–oriented approach and an approach oriented on learning and instruction. Suggestions to promote transfer are categorized as relating to task or learning content, to instructional factors, or to subject characteristics. Finally, it is shown how the concept of transfer depends on the way it is measured by the researcher.

2.1 Introduction

During the last decade, the phenomenon of transfer of learning has received renewed attention within the fields of cognitive psychology and educational research (Cormier & Hagman, 1987; Odlin, 1987; Salomon & Globerson, 1987; Jelsma, 1989; Larkin, 1989; Salomon & Perkins, 1989; Singley & Anderson, 1989; Simons, 1990; Simons & Verschaffel, 1992; Detterman & Sternberg, 1993). Although theory and research on this subject date from more than a century ago, the main questions have remained the same: how does the learning of a certain task influence subsequent performance on another task, or the learning of that other task? Which structural similarities and differences between the prior and the subsequent task account for this influential effect?

To recognize the importance of these questions, and thus to understand why they have never been out of attention, it may be helpful to mention a few instances, such as: does learning Latin improve general intellectual abilities? Is a mathematical background facilitating for learning to work with computers? Does literature education influence the way we read literary works? How can young learners use their knowledge about speech (phonemes, words, sentences) in learning how to write? How is the learning of a second or foreign language influenced by a student's mastery of his or her native language?

Unsurprisingly, the progress of psychological inquiry in the cognitive domain is reflected in the nature of transfer research. Modern transfer research is not only concerned with measuring *the extent* to which previously learned material affects the performance of new tasks; it is considered more important to gain insight in *how* these transfer processes operate in the human mind, and finally how positive transfer can be enhanced in educational settings. After the behaviorist approach of human learning, which dominated the 1950's and 1960's and which was less oriented on the specification of cognitive processes, the focus of psychological interest is nowadays on the mental processes which play a role in transferring previously acquired knowledge and skills to new cognitive tasks.

Some important studies on transfer in education have been published in the last decade. In connection with a symposium on the 'Acquisition and transfer of knowledge and cognitive skills' (University of Jerusalem, 1986), De Corte edited a special issue of the International Journal of Educational Research (De Corte, 1987). An extensive inventory of past and ongoing research and of practical applications of transfer is offered in Cormier & Hagman (1987); In 1989, the cognitive psychologist Anderson published a textbook with the title 'Transfer of Cognitive Skill' in which he and Singley elaborate the influential ACT*–theory with the purpose of accounting for transfer phenomena. In 1990, Simons set the tone for a Dutch transfer research program by means of an inaugurate speech 'Transfervermogen' (Transfer Capacity). Also in 1990, a symposium 'Transfer on Trial' was held at the Annual Meeting of the American Educational Research Association, which was reported in an integrative state–of–affairs publication (Detterman & Sternberg, 1993).

The purpose of this chapter is to offer an overview of recent theoretical literature on the topic of transfer of learning. In relation to this purpose, the transfer concept itself will be discussed and a short overview over the main historical approaches will be given. The chapter has its motivation in a theoretical funding of the 'transfer' concept, which is central to this study. Since 'transfer' receives increasing attention from language educators and researchers abroad and in The Netherlands (Levende Talen, 1992 (special issue), Bonset et al., 1992, 21–24; Verbeek, 1993, 26; Oostdam & Rijlaarsdam, 1995), and since this study also aims at contributing to the domains of writing and reading education, I thought it to be helpful to give a more extensive overview than a general learning–psychological study would require.

This overview is mainly based on American literature. This is caused by the cognitive perspective on transfer with which I started to collect it, and the recent popularity of transfer studies in American Journals. I would like to take away the suggestion that other research communities have not contributed substantially to the study of transfer. For instance, here in The Netherlands, the Psychonomics Department of the University of Amsterdam has done much work in this field (research projects 'Thinking' and 'Memory'). Under supervision of A.D. De Groot and N. Frijda a tradition was continued of researchers like Selz (1935) and Duncker (1935): introspection and thinking–aloud were used systematically as supplementary evidence to quantitative data on the effects of prior learning or education on new learning and new performances.

Although theory and research findings until now do not give a complete, coherent view on transfer processes, I will try to present some connection between research findings and guidelines for educational practice. Several researchers have attempted to propose instructional interventions which are expected to promote transfer, i.e. which would increase the likeliness of the learned material to be used in new task settings. These proposals, partly intuitive, partly based on the results of empirical research, will be summarized and discussed in the fourth section. The last section is on measuring the extent to which acquired knowledge and skills are used in new situations, in other words, on the measurement of transfer effects.

2.2 The nature of transfer

First, it should be noted that *transfer*, like learning, is a *explanatory construct*. Transfer of knowledge or skills is not a kind of directly observable learning behavior; it is a construct that learning theorists induced in order to account for certain behavioral phenomena resulting from learning. Because of the theoretical nature of the concept, its definition is depending on the learning theory that is chosen as a perspective when detecting or describing the phenomena. Probably this is why, during the last century, researchers' approaches to transfer have changed as often as learning theories have.

The various historical and recent approaches to transfer will be discussed in the next section. In this section, I attempt to sketch a global framework for the transfer phenomenon. I will pay special attention to the theoretical distinctions that have survived the different learning paradigms.

I take the definition of transfer given by Cormier & Hagman (1987, 1) as a starting point:

> *Transfer of learning occurs whenever prior–learned knowledges and skills affect the way in which new knowledges and skills are learned and performed.*

The following aspects of this definition should be noticed. By concentrating on transfer of learning, the authors stress the idea that transfer and learning phenomena should be distinguished, and that the former are possible consequences of the latter. We may observe learning without transfer, but not transfer without learning. A practical consequence is that when one tries to detect transfer, they ought to make sure that any learning has first taken place.

Since learning is usually defined as a 'a relatively permanent change in behavior or behavior capacities resulting from experience' (Mayer, 1987, 87; Schunk, 1991, 2), we must apparently distinguish between changes in behavior (capacities) due to learning, and consequential changes in behavior (capacities) due to transfer. This asks for a theoretical specification of which behaviors are considered learning effects, and which transfer effects. Failure to do so would lead to the consequence that all behavioral change resulting from experience can be called either learning or transfer (cf. Butterfield, Slocum & Nelson, 1993).

Cormier & Hagman (1987) use knowledges and skills as the object of prior learning. These are explanatory concepts for the change in behavior (capacities) resulting from learning: learning results in a cognitive change (new knowledges and skills), and this cognitive change results in (the potential for) new behavior. The same cognitive change will affect the acquisition of 'new knowledge and skills'; this affection is called transfer.

According to the definition, transfer concerns the way in which new knowledges and skills are learned and performed. This leaves room for qualitative differences as well as quantitative differences. Consecutive learning can be accelerated, or it can yield qualitatively different results; the same can be said about consecutive performance. In both cases the influence of previous learning is called transfer.

A last point to which I want to draw attention is the type of cognitive activities in the transfer situation. According to the definition of transfer, prior learning influences learning as well as performance activities in the subsequent situation. This refers to the acquisition of 'new knowledges or skills' along with their application in new tasks. However, one could raise the question if the application activities are really necessary for transfer to occur; one could also call it a case of transfer if prior learning affects acquisition of new knowledge or skills, without any direct application (for instance, when a student's understanding of a new physics subject is influenced by a mathematical background). The only way, however, for a researcher to tell whether transfer has occurred, is by having the learners perform a task that will bring the influence of prior learning to light.

On the basis of this definition, several recurring themes in transfer theories will now be discussed.

Transfer–of–training or transfer–of–learning

In every transfer paradigm an initial learning situation A is discerned from a later transfer situation B. In this transfer situation, new knowledge/skills are *learned* and *performed* (according to Cormier & Hagman's definition). If B is a learning situation, for instance in a formal schooling situation, the transfer effect is called *transfer–of–learning*. If B is not a formal learning situation but asks for application of previously acquired knowledge, the effect is often called *transfer–of–training* (Clark & Voogel, 1985; Baldwin & Ford, 1988; Broad & Newstrom, 1992). It should be noted that transfer–of–training comes close to normal learning: the application of knowledge acquired in schooling. The difference is that the acquired knowledge must to some extent be *adapted* in order to be useful in the new situation. Thus, we can distinguish two types of transfer depending on the nature of the task to be executed in the transfer situation.

In fact, every school learning is basically aimed at transfer–of–learning and/or transfer–of–training. If students are taught a certain maths problem or the topography of a country, we want them to acquire and use the new knowledge not only in the particular lesson (the initial learning situation A), but also to build on it in subsequent lessons, in later schooling and, perhaps most of all, to use it in real–world situations. These are all instances of later transfer situations B.

Therefore one can say that factors in the instructional design that promote transfer–of–training are factors that enhance *instructional effectiveness*, if we want to derive the criteria for effectiveness from the performance outside the instructional situation. It is easy to recognize the importance of this type of transfer for a 'learning–for–life' perspective on education.

Even in early transfer research, the focus was on the other type of transfer, transfer–of–learning. It was expected – and sometimes found – that the acquisition of one skill, for instance calculus, could have a positive influence on the way in which another skill, for instance geometrical problem–solving, was learned. The reach of learning appears to extend beyond the goals at which the instruction is directly aimed. People have the capability to learn more from, for instance, memorizing a French word list than just the translations of the words; they may develop their ability in vocabulary learning, or they may derive some regularities in word morphology which they can apply later to their comprehension of other words.

Throughout the century, learning psychologists have carried out experiments in order to examine the systematics of such transfer effects. The main question was: which tasks facilitate or inhibit the learning of another task? Which tasks urge people to exploit their transfer–of–learning capacities? The emphasis was on the investigation of task characteristics, which were considered to be similarly perceived by the experimental subjects.

The importance of understanding transfer–of–learning lies in its possible educational applications. Insight into the influential effects of learning tasks upon each other helps us to put the elements of a curriculum in tune; transfer–of–learning can thus play a role in enhancing *curricular efficiency*. For example, if a group of students knows how to do sums in adding and subtracting, the teacher must decide whether it is apt to continue the calculus programme with division or with multiplication. Similarly, designers of a writing curriculum would like to know if a course in formal logic can support students to acquire a skill in writing persuasive texts, or if such a course is just a waste of time.

Positive and negative transfer

In the given definition of transfer, previously acquired knowledge or skills are said to *affect* mental activities when transfer to novel tasks takes place. The results of this transfer can be desirable or not. If the transfer effect on task B is desirable, the transfer is called positive; if it is undesirable, it is called negative (Singley & Anderson, 1989, Ch. 4). A synonym for negative transfer is *interference*. The distinction between positive and negative transfer is a matter of transfer *direction*.

Note that it is not the transfer itself that is invalid in the case of negative transfer. Task A may be learned perfectly well, and the learning outcomes of this task may be perfectly well retrieved in transfer situation B. The problem is that the use of this information is unsuitable in the transfer situation B, because it hinders the learning of B.

An example of negative transfer–of–learning is a native speaker of German who learns Dutch, and applies his knowledge about German sentence construction in this foreign language. This transfer can be called generally positive, because the rules that govern sentence structure are much the same in both languages. But if the same person would try to apply the same knowledge just as successfully to his learning of Polish, the results would show the signs of negative transfer. The Polish language has a very different sentence structure, so many sentences that this person would construe would be ingrammatical. His knowledge about German sentence structure hinders his learning of Polish, but it is very hard to not invoke it (Odlin, 1987).

General versus specific transfer

Perhaps the oldest issue in transfer research has been the question of specificity (Mayer, 1987, 211; Singley & Anderson, 1989, 2). Is transfer specific and limited to a small amount of tasks or is it broad and ranges across diverse tasks and disciplines? Which characteristics of the task and of the learner influence the breadth of transfer? The educational value of transfer lies in the combination of tasks which are learned and performed with less effort together, than they would be learned separately. But how far does this facilitation go? Are there any 'key skills' or 'basic knowledges' with large positive influence on a variety of other intellectual skills?

For a long time hopes have been high for a general transfer phenomenon. A famous example of such a key skill is learning Latin, which was long thought of as a training for all kinds of other intellectual skills. However, since Thorndike (see next section) reported on his experiments, the hope for very general transfer effects has been abandoned. No evidence has been found for the existence of domain–independent knowledges or skills which transfer across tasks or across groups of similar tasks (domains). By most researchers today, transfer of learning is considered specific (Detterman 1993).

As in many areas of psychology, the case for situation–specific theories has strengthened during the last decades. Most investigators have ceased to look or hope for global theories of transfer. The trend to investigate the specificity of transfer, especially the cognitive activities that cause this specificity, has survived until today.

Vertical or lateral (horizontal) transfer

Much of the transfer research that has been carried out was aimed at assessing transfer between tasks that are considered to be related. Gagné (1973) has addressed the issue of intertask relationships by distinguishing between vertical and lateral transfer. *Vertical transfer* occurs when an existing knowledge element or skill contributes directly to the subsequent learning or performance of a superordinate task or skill. In this perspective, cognitive skills form a hierarchy of higher–order skills and lower–order subskills. For example, someone who can multiply and subtract numbers should acquire the skill of long division more rapidly than someone who has not learned multiplication and subtraction. This expectation is based on a descriptive theory about calculus in which more elementary processes of multiplication and subtraction are subordinated to the more complex process of long division (this example is further explained in the following section).

Lateral or horizontal transfer is defined in less exact terms, because it involves positive and negative transfer effects to tasks which share – again in theory – elements without having part–whole relationships. Gagné refers to lateral transfer as 'a kind of generalization that spreads over a broad set of situations at roughly the same level of complexity' (1973, 231). This means the kind of transfer which occurs when children recognize that the fractions they are learning about in school pertain to real–world situations, such as how to divide an apple pie up into equal shares.

The problem in defining lateral transfer is obviously the question of how to define in an objective and meaningful way the task elements and the tasks to which they can transfer. This is not easy, since the important relationships between tasks often exist in their underlying structure rather than in their surface similarities or dissimilarities. Although there is no standard routine for the cognitive analysis of task structure, the methodologies being currently developed have made substantial progress in identifying principles of cognitive organization (e.g. Sternberg, 1985; Singley and Anderson, 1989; Detterman & Sternberg 1993).

Near–transfer and far–transfer

Closely related to the vertical–lateral distinction is the distinction between near–transfer versus far–transfer. This is also concerned with the extent to which the initial learning task and the transfer task are (dis)similar, but the terms are used in transfer–of–training research rather than transfer–of–learning.

The distinction between near– and far–transfer is not sharp; it is rather a continuum that is meant. Mayer and Greeno (1972) use the term *self–transfer* for the effect that follows from the plain repetition of a task in a constant training situation: the acquired knowledge can be re–applied with minimal adaptation. If the transfer task or situation is somewhat different from the initial learning task, *near–transfer* can be obtained. *Far–transfer* may only be obtained when the two tasks have considerable differences.

The distinction between near– and far–transfer can also be found in the recommendations for transfer–promoting instruction (Simons & Verschaffel, 1992). Specially in learning settings with heterogenous groups, when it is uncertain in which ways the learning outcomes are going to be used by each of the participants, it is desirable to train for far–transfer. This makes the students able to adapt their knowledge or skills to their own situation, or to future changes in their situation.

Forward–transfer and backward–transfer

A last distinction is depending on the moment at which the transfer of previously acquired knowledge or skills takes place. It is mostly during the transfer situation B that an individual activates and adapts some existing knowledge or skill in order to perform a new task. Such retrospective transfer is called *backward* (Salomon & Perkins, 1989). Opposite to this, it is also possible to anticipate on the usefulness of learning content already during the initial task A. Such strategic learning or encoding of a learning task can lead to easier retrieval in the transfer situation B, in which case we speak of *forward transfer*.

2.3 Historical approaches

So far, we have discussed the most important distinctions in transfer theory which are still in vogue. Now this framework has been set out, it is the moment to take a closer look at the ways in which the transfer phenomenon has been considered during this century from different psychological perspectives.

Research into the transfer phenomenon has been going on for one hundred years now. It is not easy to accumulate the research results, neither on a theoretical nor on an empirical level. As said before, this is partly due to the continuing (and necessary) change of prevailing learning and research paradigms in psychology. Some feel that transfer research is nowadays still in its infancy (cf. Detterman 1993): few widely accepted traditions have grown on issues of definition, measurement or utility. But specially in such an infancy stage, it seems apt to learn from former research experiences in order to give direction to the growth to adulthood.

There have been at least two purposes for conducting research on cognitive transfer: a scientific (psychological) and an educational purpose. Of course, these two purposes were not necessarily strictly separated; outcomes of research with a purely scientific aim can be, and often are, considered important for educational applications. But for understanding the realm of transfer approaches during the last century it may be useful to keep the distinction between these purposes in mind.

Transfer research with a mainly scientific aim is directed at questions concerning the structure of the mental tasks, of the human mind and of human abilities. Effects of learning behavior, including transfer effects, are considered possible sources of information for theorizing about learning activities. These activities include perception, encoding and retrieval of information in memory, and processing of information in working memory.

Transfer research with an educational purpose is conducted in order to enhance the effectiveness or efficiency of instructional programs. Positive transfer of learning can be seen as a helpful human ability in designing efficient instructional programs. Positive transfer of training is even an explicit goal of all education: one does not just learn for school, but for human functioning in real life. It is this combination of scientific and educational aspects, and the hope for useful applicable results that have called researchers' attention for a long time and from different perspectives on the transfer phenomenon.

The doctrine of formal discipline; general transfer

One century ago, the prevailing scientific view upon human learning and performance proceeded from the *doctrine of formal discipline* (e.g. Angell, 1908; cited in Singley & Anderson 1989). In this view, the mind was composed of a limited set of separable general abilities, such as perception, attention and reasoning. These abilities or *faculties*, it was thought, could be developed and strengtened like muscles by means of practice. Any kind of practice for which the faculty had to be used would do, although the highest hopes were on learning comparably difficult tasks for which the faculties were driven to their limits, like learning Latin or arithmetic.

In such an approach to the learning of mental abilities, the expected transfer from a learning task is broad and automatic. Broad, because if a learning task would, for instance, strengthen the mental faculty 'reasoning', then all other learning and performance tasks in which this reasoning faculty plays a role would be facilitated. Automatic, because as with physical training, no conscious exercising is necessary to ensure this facilitation of new tasks.

In 1890, William James made an attempt to test for transfer of learning based on the formal discipline doctrine. He tried to improve the memorization of poetry in a subject by training the memory faculty, but failed to obtain transfer (Ripple & Drinkwater 1982).

Thorndike: specific transfer of identical elements

More carefully controlled pioneering studies to transfer were executed by Thorndike (1906). He also failed to find general transfer from distinct learning tasks, and based on these findings his *'identical elements' theory*. According to Thorndike, the mind was composed of many very specific habits and associations. He assumed that the results of a learning task (i.e. acquired knowledge or skill) only transfer to another task if the performance process for this new task shares *essential elements* with the initial task. 'One mental function or activity improves others insofar as and because they are in part identical with it, because it contains elements common to them. Addition improves multiplication because multiplication is largely addition; knowledge of Latin gives increased ability to learn French because many of the facts learned in the one case are needed in the other' (Thorndike, 1906, p. 243; cit. Singley & Anderson, 1989, p. 3).

Thorndike undertook a long–term research program on the generality of transfer. According to him, transfer of learning is much narrower in scope than was predicted by the doctrine of formal discipline. Thorndike succeeded in establishing the non–generality of transfer with the help of correlational studies. For instance, no correlation was found between memory for words and memory for numbers, or between accuracy in spelling and accuracy in arithmetic. Thus, if transfer was specific, teaching for transfer would imply the teaching and learning of exactly that basic knowledge and those elementary skills which will often be invoked later as part of a more complex task.

In order to find proof for his own theory of identical elements, he had subjects executing tasks which were closely related, but which shared no literally *identical* elements. For instance he had subjects doing algebraic exercises using two different notation systems. Unexpectedly to him, he found more transfer than could be accounted for by the identical elements theory alone. There seemed to be evidence for a kind of more general skill transfer between tasks with related elements.

There are at least two important objections against Thorndike's approach. In the first place, the identical elements theory offers no place for adaptation or flexibility of knowledge, which is now widely considered as the main accounting factor for transfer. 'In many transfer situations, what one knows is somehow insufficient; some kind of transformation of adaptation of existing knowledge is required' (Singley & Anderson 1989, p. 5). Transfer by means of abstraction is denied in Thorndike's theory. It is questionable whether constantly making use of the same S–R pairs in different situations is a matter of adequate transfer of knowledge, or whether it is simply doing 'more of the same'.

In the second place, the nature and specificity of Thorndike's identical elements is undetermined. Thorndike lacked the theoretical tools to decompose his tasks into mental components or specified and consistent knowledge bits. The level of specificity/generality of his elements and the notion of 'identity' have remained problematic.

Generalization theory: inducing principles

The first objection to Thorndike's theory was met by theorists of the *generalization* principle (reported in Klausmeier 1975, p. 425). This means that from a variety of experiences, a general principle can be induced. This new principle is a new and separate knowledge element in itself, and can be applied to new situations, tasks and even domains. Such inducing or generalizing capability is not specific for a theory of transfer, but is part of a more global theory of mental processing capabilities (Judd 1908; cit. Singley & Anderson 1989, p.8–9). In connection with the identical elements theory, it can account for transfer between tasks which are not identical, but which can be connected via a higher–order principle (cf. lateral transfer). Please note that such knowledge adaptation or generalization capacity is relevant to transfer between domains, such as the writing–reading transfer that is investigated in this study.

This generalization principle gives the opportunity to comprehensive, long–term educational programs (instead of the 100–task drills advocated by Thorndike). One could, for instance, focus a writing curriculum on a general 'text structuring' principle applicable to many text types, instead of learning to write several text types one by one.

Gestalt theory of transposition

In the 1930's and 1940's, Gestalt psychologists got involved in the transfer debate (Katona, 1940, Wertheimer, 1945). They questioned the 'knowledge element'–oriented way in which transfer had been approached until then, and stressed the importance of instructional factors and the complex nature of understanding. Studies opposing Thorndike's viewpoint showed that the occurrence of transfer was largely dependent on the type of instruction offered. Transfer often depends on whether the learner can find – or is offered – a common representation of the tasks. Tasks may objectively share identical elements, but if the subject does not perceive and understand them as such, transfer is very unlikely to occur.

Gestalt psychologists made a distinction between *rote* learning (studied by associanists like Thorndike) and *meaningful* learning. Meaningful learning involves the comprehension of a piece of knowledge along with the structural relations with other elements and with factors within the field of application of that knowledge. Gestalts claimed that meaningful learning would promote transfer, in contrast to rote learning, because what is meaningfully learned has more relations with other tasks. A necessary condition for transfer would be that a learner tries to understand the initial learning task with at least some of its internal and external structural relations, and tries to *transpose* this entire task structure to a more abstract level which is also stored in memory. When we learn a melody in the key of C, it is no problem to us to recognize or even produce this melody in a different key. It is not the original elements that we have stored but rather the functional relations between them.

The way in which a task is perceived and in which it is transposed into an abstract task representation is highly decisive for the occurrence of transfer. It is therefore necessary to study the strategies and representations of subjects when they perform the learning tasks. The main contribution of the Gestalt psychologists lies in their attention for the instructional process in the initial learning task. Their theoretical claims have not received analytical elaboration neither extensive empirical investigation. They never rejected the idea of identical elements between tasks as basic factors relevant for transfer; 'they merely stressed the pre-eminence of structural identity over the kind of piecemeal identity that Thorndike advocated' (Singley & Anderson 1989, p. 12).

Transfer of memory skill: The analytical Verbal Learning approach

In the early transfer studies, it was enough to observe transfer in terms of decrease of learning time and/or increase in mastery of the transfer task, as a consequence of initial task acquisition. However, when transfer turned out to be quite enigmatic, questions for the precise activities of knowledge encoding and retrieval became important. An empirical research paradigm, referred to as the Verbal Learning approach, was followed during the 1940's (and later) in order to gain insights into these activities (Osgood, 1949). The Verbal Learners attempted to subject the transfer phenomenon to a thorough analysis in laboratory circumstances.

Many experiments were conducted to establish the effect of stimulus and response (dis)similarity on the recall of responses. In many of these studies, subjects had to learn two word lists consecutively, each list consisting of paired stimuli and responses (words or syllabes, sometimes images). The similarity of the material to be learned was systematically manipulated, and the effect of this manipulation on the learning time or quality of the second list was measured. Osgood (1949) systematically integrated these experiments. Unsurprisingly, it was found that positive transfer only occurs between identical or similar stimulus words connected with identical or similar response words (identical words evoking the largest amount of transfer). If the similarity decreases, to neutrality or opposition, transfer decreases and becomes negative. Positive as well as negative transfer effects were found to be stable.

It is not easy to draw conclusions from these experiments with respect to education. The skill trained here is merely the storage and retrieval of concrete information elements in long–term memory. The importance of the Verbal Learning paradigma is that it is the first move to an analytic, cognition–oriented approach to the study of transfer phenomena in which attempts are made to specify transfer content and processing activities. This property is shared with more recent transfer approaches.

Gagné and the analysis of task performance

As said, the relevance of the tasks used by Verbal Learners for educational programs is limited. Gagné (1973), an educational psychologist, felt attracted to the idea of applying the analytic cognitive approach to more complex classroom tasks. His interest was in 'designing well–formed curricula that maximized learning and transfer and that could be used in conjunction with the new *programmed instruction* technology' (Singley & Anderson, 1989, 15). By means of this type of instruction, Gagné hoped to approximate the strict conditions of laboratory research which the Verbal Learners imposed to their work. Besides, in developing a programmed instruction course, the complex learning task needs to be decomposed in subtasks or *frames*; an activity that matched Gagné's intentions for the study of transfer.

In specifying a skill for which a curriculum needed to be designed, Gagné used some kind of rational decomposition. When describing a certain cognitive skill, he gave a hierarchical representation of subskills called *learning sets*. It was assumed that a superordinate skill could only be acquired if one or more of the necessary subordinate skills is mastered. For instance, if multiplication is in part adding, then mastery of adding is a necessary prerequisite for learning to multiply. In a scheme:

[insert figure 2.1 here]

Figure 2.1*: Division two prerequisites: adding and subtracting.*

Gagné made the distinction between *vertical* and *lateral* transfer. Lateral transfer was defined as the kind of transfer that spreads over a broad set of situations at roughly the same level of complexity (Gagné 1973). In this case, transfer could be expected from counting apples to counting numbers one–by–one, or from addition to subtraction. Vertical transfer takes place between lower–level and higher–level skills which exist in a part–whole relationship to one another. The skill of multiplication *includes* the skill of addition, and is thus superordinated.

The ratio for task analysis was not specified by Gagné; not a singular consistent analytic instrumentarium was used to specify cognitive operations in each subtask. Decomposition of a skill was done by some rational specification of thinking steps and the required knowledge in the execution of that skill. The questions for these thinking steps and knowledge are put recursively until one arrives at thef so–called basic skills.

In constructing a curriculum on the basis of such skill analysis, more than one way could be followed as long as two conditions are fulfilled: a) the direction order of the skills offered in a learning program must be upwards and b) the mastery of a subordinate skill must be established before a superordinate skill is taught. But Resnick (1976; cit. Singley & Anderson, 1989) showed that very good students are able to span a gap, i.e. that they are able to acquire a skill even if the direct subordinate skill is not yet mastered.

Gagné tested his transfer theory (and his decompositions of classroom skills) by presenting students with programmed instruction based on a skill hierarchy, and monitoring their transitions between subskill levels (e.g. Gagné & Paradise, 1961). As confirmative evidence for the transfer theory he considered those performances which included mastery on both subordinate and superordinate skill (someone who can add and multiply), and performances which included failure on both these skills (someone who can neither add nor multiply). Someone who masters a superordinate skill, but fails on the necessary subordinate is in direct contradiction to the theory; it means that either the theory or the task analysis is wrong. And, at least, someone who masters the subordinate skill and does not master (yet) the superordinate skill is irrelevant to the testing of the theory; such a person would probably have to receive more instruction or practice on the superordinate skill.

Empirical research conducted by Gagné offered no strong support for the effectiveness of the program. His assumption was that such principled curriculum design would be helpful in enhancing transfer. However, even when all subordinate skills were mastered, the superordinate proved often to be still too difficult for students in 13 to 50 % of the cases. Gagné put the blame to the discutability of his task analysis. Many decompositions of cognitive skill are possible, and in fact, Gagné had to cope with the same sort of representational problems as Thorndike. A good and consistent theory is needed for the determination of cognitive elements shared by tasks, and for a detailed specification of mental activities.

Conclusion: What information pertains?

Many researchers know the transfer phenomenon from their own experience. It is easier to learn something new if you have some experience in learning. Some knowledge is more useful than other knowledge. There are more and less successful ways in which scholastic knowledge is applied to real–world problems. The problem is that this transfer phenomenon has shown to be rather unruly in scientific research.

However, important lessons can be learned from the different historical approaches to the study of transfer. Three main research questions regarding transfer can be distinguished:

1) a question concerning <u>description</u>: <u>Under which conditions can transfer be observed?</u> This question does not only call for a definition of transfer phenomena, but also for a description of - learning and performance situations in which they occur, of learning contents (knowledge or skills) susceptible to transfer, and of the kind of learners capable of transferring what they have learned.

2) a question concerning <u>explanation</u>: <u>How does transfer work within the learning and task–performing mind</u>? This is the question for explanatory theories about cognition, encoding of learned material, retrieval and adaptation of knowledge and use of skills in performance situations, and influence of educational circumstances.

3) a question concerning <u>prediction</u>: <u>How can we design learning situations in order to</u>

promote transfer? This is the question for the applicability of the knowledge gained in answering the first two questions. This applicability does not only depend on the possibilities of cognitive transfer, but also on other, external factors in education such as limitations to teaching time, individual attention and priority of educational goals.

Conclusions with respect to the nature of transfer can be drawn with respect to its specificity. The hope for overall generality of transfer has held off. Klausmeier (1975) identified a trend away form the all–embracing theoretical claims that attempt to account for transfer phenomena under all conditions. This general trend has continued. As in many other areas of psychology, the case for situation–specific transfer has strengtened during the last decades. Most researchers have ceased to look or to hope for global theories of transfer (cf. Detterman, 1993).

Conclusions with respect to research methods are not unidirectional, but it is important to notice the need for a comprehensive, validated theory of cognitive task analysis. The research of Thorndike as well as Gagné, who both tried to combine scientific control, educational relevance of the tasks chosen and detailed accountability was impeded by the lack of such a theory of analysis.

As we have seen, at least two conditions must be fulfilled before transfer between two mental tasks will occur. One condition is concerned with the relevance of the *content* of these *tasks*, i.e. some kind of perceived similarity or analogy between the performance processes of these tasks. Transfer appears to be specific, but not in a too strict way: learners may transfer similar as well as identical elements. We need to know more about the relation between task (dis)similarity and the likeliness of transfer. The second condition is concerned with the *learning or performing circumstances*. Factors in these circumstances (instruction, motivation, activation prompts) seem to have a strong effect on the appearance of transfer, or at least on the amount of transfer obtained. These factors affect the way in which the prior knowledge or skills are stored and/or retrieved, thus affecting the likeliness of transfer.

These two conditions reflect two transfer approaches that can be discerned in present psychological research. Both approaches are analytical with a cognitive background. In the one approach, Thorndikes theory of identical elements is revived with the help of an elaborated theory about the acquisition of cognitive skill. A modern version of this theory has been represented by Singley and Anderson (1989). In the other approach, study is made of the way in which different instructional methods or learner characteristics affect transfer. These approaches will be discussed in the next section.

2.4 Contemporary approaches

Two recent approaches to the study of transfer are sketched in the present section. It should be noted that their mutual relation is complementary rather than concurrent. They both inherit their interest and demands from the varying results of earlier research. On the one hand, they are concerned with the structure of mental functioning in transfer situations; on the other, with the implications of this functioning for effective education.

Both approaches have to a certain extent inspired the experiments that will be reported in the second part of this study. Therefore they are described here more extensively than the previous historical approaches.

Singley & Anderson: A resurrection of the identical elements theory

Every theory of transfer is in fact based on a theory of learning, just like every learning theory reverts to a theory of performance. It is this idea that got Newell & Simon (1972) to discuss the internal logic of a research agenda for cognitive psychologists. It is also this idea that made Singley and Anderson (1989) decide to extend their theory about the performance and acquisition of cognitive skill, called the ACT* theory (pronounced 'act–star'), with a theory of transfer. They elaborate on Thorndike's identical elements theory, attempting to overcome the handicap that the absence of task analysis theory has imposed on transfer research (see section 2.3).

In any identical elements theory, the critical question is how to define the elements. Singley & Anderson's attempt is aimed at specifying the identical elements in terms of so–called *productions*, a certain type of cognition. Because they substituted this crucial concept of their ACT* learning theory for the identical elements, it seems sensible to discuss this ACT* here in short.

The ACT* theory is a detailed, comprehensive theory of skill acquisition. It lies within a broader class of theories that use *production systems* to model human cognition (Newell & Simon 1972). Every production system consists of at least a long–term memory and a working memory. The long–term memory contains, among other things, *procedural knowledge*: sets of condition–action rules which are called *productions*. A condition–action rule, or production, is a piece of knowledge which is principally aimed at some intentional activity. The general form is:

IF (condition = I, II, III ...), THEN (action = A, B, C ...)

For instance,

IF (conditions): I. the goal is to use a <pronoun>;
II. the referent is a <person>;
III. the grammatical function in the subordinate clause is an <object>;

THEN (action): I will use the word <whom>.

A production such as this 'fires' as soon as conditions of the IF–clause match the state–of–affairs in the working memory. At that moment the action, specified in the THEN clause, is executed. The conditions in the IF–clause include a goal statement and context– and situation–dependent variables. According to Anderson (1983), all cognitive functioning can be theoretically described in terms of productions.

Unlike other production system models, the ACT* theory has an extra device: a so–called declarative memory, containing *declarative knowledge*. This kind of knowledge refers to 'facts' or assertions with truth value, in contrast to *procedural knowledge* (as represented in the procedural memory) which is about the specification of actions in order to achieve goals (procedures).

It must be noted that the nature of procedural knowledge is quite different from declarative knowledge. Certainly not all knowledge about procedures is procedural knowledge. The knowledge used, for instance, to tie your shoe–laces is indeed an example of procedural knowledge: we automatically execute the action if all conditions are fulfilled, even without consiously knowing what we do. But we can also have declarative knowledge about procedures like, for instance, a male gynaecologist knows how to give birth to a baby, or like a student learns rules for writing summaries of scientific articles before putting them into practice. This knowledge can be verbally stated without too much effort, and is not used routinely such as procedural knowledge.

Singley & Anderson (1989, p. 198–200) list the following differences between declarative and procedural knowledge:

Declarative knowledge		Procedural knowledge	
1.	can be stated verbally with comparable ease;	1.	can hardly or not be stated verbally;
2.	sensitive to forgetting; more difficult to relearn;	2.	not so sensitive to forgetting; easy to relearn;
3.	worse candidate for *automatic* transfer;	3.	better candidate for *automatic* transfer;
4.	broadly accessible (easy generalizing across or associating on content);	4.	narrowly accessible (hardly generalizing across or associating on content);
5.	better candidate for non–automatic or *deliberate* transfer.	5.	worse candidate for non–automatic or *deliberate* transfer.

Besides, the two types of knowledge are neurophysiologically separated. Amnesic patients are often able to learn the execution of procedures, but hardly any facts.

The ACT* theory breaks down the acquisition of cognitive skill in two major stages: a *declarative stage*, where a declarative representation of the skill is interpreted by provisional general productions (consciously 'knowing' what the task 'writing introductory paragraphs' encompasses), and a *procedural stage*, where the skill is directly embodied in automatized domain–specific productions. An example of such a production is, for instance:

IF (conditions):	I.	the goal is to write a paragraph
	II.	the focus of the paragraph = main subject of text
	III.	the textual function of the paragraph is a conclusion
THEN (action):		**I will re-read the posited problem and main points in the text**

The transition from the declarative to the procedural stage is achieved by the process of *knowledge compilation*. Knowledge compilation consists of two separate activities. The *composition* activity collapses the provisional general productions into many, highly specific productions (finding and recording the precise steps required in paragraph writing, like the example above). The *proceduralization* activity deposits the resulting domain knowledge from long–term memory directly into productions (declarative knowledge about paragraph writing is transformed into automatic writing behavior).

A part of this theory which has recently been added is *structural analogy,* an activity for translating related declarative knowledge into action (Anderson & Thompson, 1990). During this translation process, a generalized production is produced by abstracting common features of the source and target of the analogy:

– 'in an essay, the main viewpoint must be expressed either at the beginning or at the end'	&	– 'in a paragraph, the main viewpoint must be expressed either at the beginning or at the end'.

→ may lead to generalization:
'in any text, the main message ...'

These are the essential concepts and relations within the ACT* theory of skill acquisition. In short, all skills are initially acquired as pieces of declarative knowledge concerning a specific domain; they are transformed into condition–action rules which are practised, adjusted and elaborated until conscious interpretation of task elements is no longer necessary. At this moment the skill is automatized; the declarative knowledge is transformed into procedural knowledge.

How does this learning theory concerning cognitive skills relate to the transfer topic? According to Singley & Anderson, it is the productions that can take the place of the identical elements that Thorndike was searching for. 'A first approximation to an understanding of transfer involves comparing two sets of productions for different tasks. To the extent that the production sets overlap, transfer would be positive from one task to the other. To get a slightly more quantitative prediction, weights might be assigned to the productions according to their relative frequency of use in transfer tasks' (Singley & Anderson 1989, p. 31).

Compared with Thorndike's static 'elements', the concept of production has several advantages, according to the authors. A production is 'versatile and powerful' in that it contains variables instead of fixed values. It can contain more than one variable and thus be elementary as well as complex. It is 'abstract and can be used to represent many different yet functionally equivalent methods at various levels of generality' (Singley & Anderson, 1989, p. 32). Productions can be used to represent internal cognitive behavior (reasoning steps, monitoring activities) as well as visible behavior.

Singley & Anderson distinguish four types of transfer, depending on the kind of knowledge (declarative or procedural) acquired or applied in the initial and the transfer task:

TARGET KNOWLEDGE (Transfer task)

		to procedural knowledge	to declarative knowledge
SOURCE KNOWLEDGE (Initial task)	from procedural knowledge	*1. Productions in the learning task apply directly to the transfer task.* E.g. driving in two different, but similar cars; sentence construction during writing and during typing. This kind of transfer is automatic as long as the transfer task is represented in such a way that the production is *directly* applicable (e.g. not if the steering wheel is at the opposite side).	*2. Acquired productions are helpful in later acquisiton of declarative knowledge.* This is comparably indirect transfer, e.g. basic skills such as reading, writing and studying skills, made up of procedural knowledge, which are helpful in acquiring new knowledge in school subjects. These advanced skills are strong candidates for general transfer across domains (Larkin, 1989).
	from declarative knowledge	*3. Declarative knowledge acquired in the learning task is helpful in the acquisition of productions in the transfer task, by means of the 'structural analogy' activity.* Declarative knowledge about a solution to an 'old' problem is activated (e.g. read the newspaper's headlines to select important news) and modified to fit a novel problem (e.g. scan a book's index to select important parts). As a by–product, a production rule may be generated which captures the essence of the solution and generalizes across specific features in the source and target (e.g. check for summary information when having to process a large piece of text). The difficulty in creating the analogy lies in recognition and retrieval of existing knowledge, and in modifying the solution (Gick & Holyoak, 1987).	*4. Existing declarative structures have influence on the acquisition of new declarative structures.* E.g. the type of transfer studied by the Verbal Learning psychologists (see 2.3). But also the adaptation of an 'old' problem solution to a new task (like with declarative–procedural transfer) which does not directly result in new, more general productions (e.g. the advice to scan a book's index before reading it, without making any further generalization). In general, early stages of skill acquisition involve the construction of declarative knowledge on the basis of existing knowledge, both declarative and procedural.

Figure 2.2: *A taxonomy of transfer types (adapted from Singley & Anderson, 1989, 33).*

The main objective of the ACT* approach to transfer is to test the ACT* theory as a learning theory in a transfer context. Singley & Anderson restrict themselves to transfer from declarative and procedural knowledge to procedural knowledge (types 1 & 3). In their research, Singley & Anderson

(1989) use novices' task representations to allow a large role for declarative knowledge (i.e. not yet proceduralized knowledge) and task control (i.e. of not yet automatized tasks). These representations are obtained by means of rational task analysis. If possible, protocol analyses and quantitative modeling yielded complementary information. The level of analysis is comparably detailed, since exact understanding of the way in which knowledge elements transfer calls for more than just general assessments.

Their study is an attempt to validate the usefulness of productions (or their declarative precursors) as candidate elements for transfer, and to establish the activities of lateral, negative, specific and analogical transfer. The most important tasks under study are programming, text editing and calculus. Singley & Anderson report to be quite successful in predicting the amount of transfer between tasks, the prediction being made on the basis of task analysis with productions.

Neither the exact ins and outs of their experiments nor the theoretical working–out of the task analyses can be discussed at this place. Here it must be sufficient to notice that a promising approach to transfer has risen with the attempt to theoretically specify the cognitive activities of skill acquisition and transfer, and to explain transfer by means of 'identical elements' made up of productions. It should be noted that Singley and Andersons approach is not as new as it may seem: about 20 years ago, Elshout found transferrable 'identical elements' among several thinking skills (Elshout, 1976). Much work must still be done in order to test the theory, by trying to analyze, for instance, divergent tasks in addition to the convergent and automatized (problem–solving) tasks that Singley and Anderson studied.

Educational psychology: Promoting transfer by instruction

In the second half of the 1980's, several educational psychologists have tried to account for transfer effects, or rather for the absence of transfer effects, in terms of factors regarding the learner, the instructional situation or the performance situation (Brooks & Dansereau 1987; Voss 1987; Pea 1987; Salomon & Globerson 1987; Salomon & Perkins 1989; De Corte e.a. 1990; Snow 1990; Simons 1990; Simons & Verschaffel, 1992). Some influential contributions were offered by Gavriel Salomon, who published two articles about the concept of *mindfulness*. I will pay some special attention to this concept, since it may explain the type of transfer (lateral, forward transfer between the writing and reading modes) that this study is aimed at.

Salomon & Globerson (1987) and Salomon & Perkins (1989) give an account for the fact that people, including expert learners, do not always make the best possible use of their knowledge; a fact all too easily recognizable in transfer experiments. Besides, the authors want to explain the often contradictory findings in transfer research by means of a new learning categorization: the categories of *'high road learning'* and *'low road learning'*, and determining the conditions under which these two types of learning can yield transfer.

The gap between how people usually perform new tasks and how we could optimally perform, is due to a transfer problem, according to Salomon, because previously learned knowledge is not or not optimally used in executing novel tasks. Possible reasons for unnecessarily poor task performance include cognition (e.g. problems with memory usage, misconceptions, non–effective strategies), motivation (e.g. success expectation, self–esteem, wants) and personality (e.g. a tendency to rely on certain habits or methods). Salomon et al. add a fourth reason called *mindfulness*, defined as:

> *'the volitional, metacognitively guided employment of non–automatic, usually effort demanding processes'* ((Salomon & ..., 19.., ..).

Mindfulness is a factor in the working method that a person applies in his task execution. Mindfulness is applicable to consciously controlled, thus non–automatic processes, and allows an individual to withold initial responses to problems, examine situational cues and meanings, generate alternative solution strategies, generate information, draw connections and make abstractions. In less mindful task performances, the individual relies too heavily on already known schema's and strategies. Mindful performance is always more effort–taking than non–mindful performance.

Educational circumstances which provoke the learner to be mindful have usually large positive effects on learning outcomes. They stimulate learners to carry out the above mentioned non–automatic activities which contribute to the quality of the performance. But what has such a feature to do with transfer? This depends on the type of learning that precedes the transfer.

According to Salomon, there are two types of learning activities: learning by way of a 'low road' and of a 'high road'. All knowledge and skills that we acquired have come to our minds by way of one of these roads. Low–road learning takes place by means of extensive and varied practice (learning word processor functions by writing various texts), automatizating stimulus–response pairs (knowing other people's phone numbers) or motor skills (riding a bicycle). Execution of skills, retrieval of knowledge learned in such a way is fast, but inflexible. The learner is, for reasons of efficiency, trained to react directly on the situation. Wrong responses can hardly be witheld or adapted and must be corrected afterwards. Transfer of low–road skills is limited in scope to those applications in which the (inflexible) knowledge can work without having to be modified.

'High road learning' is concerned with mentally demanding reflection on tasks or abstract problems which cannot automatically be solved (e.g. solving complex math problems, writing an essay). It requires analysis of the goal structure and weighing alternative solutions. Explicit instruction is often required, and for a longer time.Learning results of this type can be abstracted from the learning situation or example and transferred to comparably new fields.

Each of these two types of learning activities supports the other; neither of them is superior to the other as long as they are executed in their own suitable fields of learning content. E.g. for student learning, high–road idea selection and low–road note–taking work well together in the complex task of processing information from lectures.

With these concepts, the authors try to account for the varying degrees of success in transfer research. They state that in many studies where transfer failed to occur, the conditions for the transfer of neither high–road learning nor low–road learning were met. Which are these conditions that mediate transfer of low–road, respectively high–road learning?

Low–road transfer. By practising in varying contexts, the learning material gets decontex-tualized. The decontextualized material will be efficiently invoked in different task settings. (Abstracted features of) the task situation will initiate performance processes (cf. Anderson's condition–action rules). It is usually unintentional, implicit, model–based knowledge or skills that are transferred. Reflective decomposition of the behavior is difficult, if not impossible. We can think of habits, socialization, personal beliefs, motor skills and cognitive styles as good candidates for low–road transfer to new tasks.

High–road transfer. Also decontextualization takes place, but now in a mindful manner. The learner must attempt to deliberately make an abstraction by dropping details from the original task setting, in order to widen the category of tasks to which the solution is applicable. This is called the strategy of *mindful abstraction*: 'the deliberate, usually metacognitively guided and effortful decontextualization of a principle, main idea, strategy or procedure, which then becomes a candidate for transfer; or – in rarer cases – the learning of such a principle in abstract form in the first place' (Salomon & Perkins, p. 126). The abstraction forms the bridge between contexts of application, since they lead to more generally applicable productions. Mindfulness is a condition for understanding and connecting the abstraction with usage contexts, or for explicitly deducing the abstraction. The effects of metacognitive abstraction of a single problem–solution pair, by means of writing summaries, on transfer task performance are impressive. It should be noticed that the act of mindfully creating the abstraction does not only yield new prodctions, but can have a learning effect in itself: the student may become better in the skill of 'abstraction'. In this respect, mindfulness is a personal characteristic with broad applicability.

Sometimes, high–road learning has to come into play if automatic processes fail (an American driving in an English car; native speaker of English applying his skill in sentence–building to the French language). Abstractions can be encoded with possible applications in mind (forward transfer) or they might be made 'on the spot' on the basis of existing information (backward transfer), although this is usually a difficult task which bears heavily on the memory and processing capacity.

Salomon & Perkins (1989) discuss examples of transfer failure in older research, and try to put their finger on the spot. For instance, instruction in the children's computer programming language LOGO has hardly shown any positive effect on planning skills (Verschaffel, De Corte & Schrooten, 1992). According to Salomon & Perkins, the initial task was poorly learned, and no opportunity for automatization/decontextualization is given, neither any stimulation to abstract their knowledge.

Salomon & Perkins conclude in stressing the need to explicitly teach for transfer. This means that either automaticity or a mindful abstraction process must be provoked if we do not want knowledge to be inert for most students. Besides, learners should be able to cue task situations for solution–relevant features, in order to facilitate backward–reaching high–road transfer ('in which tasks have I seen these features before?'). Thus, structural *relevance* of a learned topic is a necessary condition for transfer, but it may not be sufficient; fulfilment of one or more transfer conditions is needed as well. Without this, transfer effects will remain unstable or simply absent.

In sum, a major shift in theoretical approaches to transfer can be observed from task–orientation (formal discipline, Thorndike) to more cognition– and instruction–oriented approaches (generalization, Gestalt, Verbal learning and Gagné's task hierarchy). Both foci are still present in modern approaches, as witnessed by Anderson & Singley's and by Salomon's work. It is important to note that the two approaches do not conflict nor compete. The two approaches seek to explain transfer phenomena by means of different determinants of the learning situation. It is most probable that more than one factor will influence the occurrence of transfer, so both task, learner and learning–environmental variables may well be investigated in their own right.

2.5. Transfer-promoting instructional methods

The practical pretension of transfer research has been mentioned before. Many researchers have tried to translate their findings into recommendations for instruction. In this section I will pay attention to these recommendations. It must be stated that they are provisional, rather based on intuition and theoretical suppositions than on research. This is specially the case because empirical transfer research has for a long time focused on the content and structure of tasks rather than on instructional factors. Thus, I cannot give a certified list of recommendations that can be followed in today's classroom instruction in order to facilitate transfer of learning. The discussion which will follow here may function as material for a 'context of discovery' in determining which instructional interventions deserve theoretical elaboration and empirical inquiry.

I will categorize the proposed educational methods into three: those related to tasks, related to the learning environment (including the instruction) and related to learner characteristics (cf. Simons & Verschaffel, 1992).

Transfer-promoting methods related to tasks

Teaching basic skills. There is a striking analogy between Thorndike's opinion concerning teaching for transfer on the one hand and the actual quest for the teaching of basic skills on the other. According to Thorndike, precisely those knowledge domains or elements and those cognitive skills have to be identified which occur often as part of a variety of (also complex) real–world tasks (Thorndike 1906; cit. Singley & Anderson 1989, p. 24). Thorndike's pedagogy contains that this learning content must be repeated over and over again, in order to ensure optimal transfer.

We observe similar viewpoints in the 'basic skills' discussion (Nijhoff, 1990): the wish to identify those skills that everyone needs in order to fulfill their present or future functions, although these functions may be quite unspecified. One can also recognize this idea in the determination of 'eindtermen' or 'kerndoelen' (basic goals) in The Netherlands. These are educational goals that must be attained by all secondary school students in the Netherlands at the age of 15 (for mother–tongue language, see Rijlaarsdam, 1989). Imposing these goals to the curriculum of the first three years of secondary education can be seen as an attempt to guarantee transfer–of–training from education to the basic requirements that the average student must meet after this part of his schooling.

Teaching realistic tasks. To a large extent, the potential for knowledge acquired in schooling depends on its usability in real–life. Therefore, teachers should teach realistic tasks and use realistic examples (Simons 1990; Detterman 1993; for mother–tongue language, see Ten Brinke, 1976).

A problem lies in the determination of future task requirements. An analysis of real–life tasks for each school subject in general education is not easy to make. Another problem lies in the large expected variability of task requirements that all students will meet in the future. The basic skills for John, destined to be a journalist, will most likely not be the same as those for Mary, a future professor in genetics. It gets more troublesome to establish realistic goals that are valid for both of them as their age increases.

Transfer-promoting methods related to the learning environment

Several factors in the learning environment influence the occurrence of transfer. Most of them have to do with the type of instruction, and the learning activities following from this instruction.

Teaching for meaning–making: Gestalt psychologists have stressed the relative unimportance of rote learning if the goal of teaching is to obtain transfer. If we want students of German to use prepositions with the correct casus, we should not only teach them a list of these prepositions, but also each of these prepositions in connection with verbs that invoke a certain casus. Even better would be to teach students the prepositions and verbs in a variety of sentences. As a result, functional relations between prepositions and verbs (and more abstract: a usage context) will be stored and more readily used. It is rather the resulting *understanding* of how knowledge can potentially be used, and the acquaintance with usage contexts that lead to actual use (cf. Detterman 1993).

Klausmeier summarizes in his review of the transfer phenomenon (Klausmeier 1975) the 'principles for teacher behavior' in relation to transfer. They are most of all based on *explicit* teaching–for–transfer, which means directly involving the students in the possibility of knowledge use.
Explicitizing the goals. As in any intentional human activity, setting a goal initiates executing activities. Therefore it is wise to explicitly teach students for transfer if transfer–of–training is the aim. This means: teachers should make clear why they want students to know what they teach. They should sketch ways in which the knowledge will be used. Students must be given clues on how to remember the material in relevant circumstances, and the most important parts must be stressed.

Explicitizing relations between tasks. Related to this is the advice to practice the knowledge in more than one task, along with making the functional relations between these tasks explicit. If two goniometric problems can be solved in a similar way, for instance by making a auxiliary line, the teacher must show the students this, and *why* this is the case, and not only how they can be solved.

Supplying feedback. Experiencing success in the execution of a learning task facilitates retention of this execution. That would mean that correct responses must be rewarded in order to strengthen the retention.

Explicit and varied representation of relations within the content. For vertical transfer, the mutual relations between knowledge and skills must be clarified, or sometimes abstractions must be made from discrete knowledge elements. Make these relations or abstractions explicit and emphasize them, and find proof for your students understanding it. Some researchers make a plea for various representations of the concepts and their relations: in words as well as in graphical /schematized forms (cf. Simons, 1990).

Varied practice. Retention of problem–solving methods will be enhanced if the solution or its underlying principle is applied to novel problems with, for instance, a different appearance. So it is advisable to provide *various application tasks* (cf. Salomon & Perkins, 1989; Simons 1990).

Cumulative learning. Learning over a period of time will certainly give more stable abilities and knowledge. That is why learning should be *sequential* and *cumulative,* so that students' knowledge will not become inert after they have done a test.

It is remarkable how often the advice to *explicitly* teach for transfer is given. Voss (1987) asserts that a strong 'perception–aiding component' in the initial learning task should be present, with respect to the perception of new tasks. It means that the learning material is encoded together with explicit abstract knowledge of the concept categories and problem types of the transfer domain of application.

Transfer-promoting methods related to learner characteristics

So far I have only discussed the tasks and factors in the learning environment that may evoke transfer. However, learners are not passively subject to these influences. Even if we would offer the same tasks under the same conditions to a group of learners, we would observe various amounts of transfer within this group. The accounting factor for this must be found in the learner's cognitive or motivational characteristics.

Memory–access skills: Pea (1987), Simons (1990) and Sternberg & French (1993) stress the importance of memory–access skills. According to them, transfer is an interpretative problem. Students fail to transfer most often because they are not aware at all that they have the cognitive tools to solve a problem. Task features for a species of problem types must be taught explicitly, together with tools for retrieving relevant knowledge. Visual mapping of learning content is advisable. In the end, the student must be autonomous in discovering possibilities of using his knowledge, without a scholastic instruction being necessary.

Task–monitoring skills: McKeachie (1987) adds to this the explicit teaching of task–monitoring skills which guard the execution of a task. Many tasks that students must master are complex, and call for attentive monitoring and regulation. Task–monitoring skills can therefore be applied to a variety of school and real–life tasks. Possible transfer of such skills would therefore be very general.

Learning attitude: Prawat (1989) mentions an essential difference in goals that students can aim for while executing a task: some will want to *finish the task* in casu, while others want to *learn how to finish such tasks.* The first are oriented on performance, the last on mastery. Both groups define the means–end–relation reversely. For the first, learning is a means to the end of task completion. For the last, task completion is the means of becoming able to master such tasks. I will pay special attention to this distinction in the following chapter.

Simons's transfer conditions

Finally, Simons (1990) discusses in some detail a set of ten 'transfer conditions'. The fulfilment of these conditions to learning may account for the occurrence or absence of transfer. I discuss these conditions in connection to the category 'learner characteristics' because, according to Simons, the learner's influence on their fulfilment is crucial. Simons ascribes a property 'Transfer Capacity' to each learner, indicating the extent to which the learner is capable of making memory representations that are adequate for (far) transfer, by fulfilling each condition (Figure 2.3).

Category	*Transfer condition*	*Description*
quality of memory representations	1. memory trace	– the initial task must be mastered in order for transfer to occur
	2. number & strength of memory connections	– the retrievablility of the acquired cognition determines the likeliness of activation in the transfer situation
	3. compilation	– declarative knowledge must be proceduralized to the extent that it can be used in real problems
general knowledge and cognitive strategies	4. adequate strategy use	– learners must have strategies at their disposal for problem decomposition, self–regulation and transfer
	5. metacognitive knowledge	– learners must acquire and use knowledge about their own knowledge and their task performance
	6. usefulness	– learners must determine the usefulness of the acquired knowledge
relation between learning context and transfer	7. context	– learners must perceive similarities between learning and transfer contexts (e.g. realistic contexts)
	8. decontextualization	– learners must distinguish specific details of the learning context from abstracted, transferable ideas
	9. affective aspects	– motivational and affective factors in the learner co–regulate memory access

Figure 2.3: *Simons's transfer conditions*

According to Simons, transfer conditions 1, 2 and 3 can be fulfilled by meaningful learning and practice in various contexts, selecting main points in subject matter, explicit relations and diverse representational systems. Conditions 4, 5 and 6 are usually first fulfilled by the teacher, who explicitly explains and/or performs the strategies, and monitors/comments on their use by the students. In a further stage, the students must become responsible for monitoring themselves. Usage and usefulness can be explained by the teacher, but also discovered by the student. Conditions 7 and 8 make up a plea for realistic/simulated learning contexts. Decontextualization is stimulated by intentional reflection and construction of metacognition.

The learner is to a large extent co–responsible for being prepared to learn–for–transfer, states Simons. It is not sufficient to find whether transfer between tasks, skills or domains is feasible at all, but under which conditions it is feasible and with which learners (Simons & Verschaffel, 1992).

This concludes our inventory of transfer–promoting features in the instruction. In order to build a transfer theory that could account for transfer effects in various domains, the effects of these methods ought to be studied in certain relations to each other. Some of them are more suitable to evoke near transfer, others to evoke far transfer. A curriculum that is developed specially to promote transfer should, like any other curriculum, take more than one of the categories 'task', 'environment' and 'learner' into consideration.

2.6 The measurement of transfer effects

Not only the theoretical approaches to transfer have changed over the years, development can also be observed in the technology of transfer measurement. Particularly during the decennia in which the Verbal Learning approach dominated transfer research, psychologists have engaged themselves with developing adequate experimental designs and quantitative measures which could be used in their type of empirical transfer research (Gagné, Forster & Crowley, 1948; Woodworth & Schlosberg, 1955; Ellis, 1965; Singley & Anderson, 1989). In this section, the most important of these designs and measures will be discussed. The relevance of this discussion lies not only in the arguments for selecting a particular design for our experiment, but also in the clarification of the concept 'transfer': which learning of performance behaviors should be taken into account when measuring 'transfer'?

A *transfer design* is a plan concerning the organizational structure of a transfer experiment. This experiment must enable the researcher to measure the effect of learning some task on the execution of a subsequent task, which may include learning as well as performance activities.

In discussing the varying transfer designs, I will use a standard example: a researcher who wants to determine the transfer of learning grammar (task A) to writing (task B). How will he or she conduct the experiment? That is, on what grounds will a design be chosen and how can transfer effects be assessed? It is assumed that, according to Cormier & Hagman's definition (section 2.2), the 'prior learned knowledges and skills' will be acquired by practice in grammar exercises, that the 'new knowledges and skills' will consist of the ability to write grammatically correct sentences in essay composition (so the 'new knowledge' calls for the application of grammar knowledge to sentence construction'), and that 'the way in which' this writing task is affected is assessed by quantitative measures: the number of syntactical errors and the number of (non–syntactical) spelling errors.

While I will use the term 'task' in this discussion to indicate the learning content of the initial learning situation, it is understood that 'skills' or 'knowledge' can be acquired in this situation as well. Thus the discussion is also applicable to the acquisition of skills and knowledge.

Gagné, Forster & Crowley (1948) mention four transfer designs and their matching formulas. One of the earliest and most insightful experimental designs is as follows:

Figure 2.4: *Prototypical transfer design*

Group	Learning Task	Transfer Task
Experimental	A	B
Control	—	B

In such a design, an experimental group is instructed in a (learning) task A, followed by executing a (transfer) task B. Their learning and performance of this transfer task is directly compared with the learning and performance of a (comparable) control group, which received no instruction in task A. If the experimental group outperforms the controls on task B, there is positive transfer from A to B; if the controls outperform the experimentals, there is negative transfer from A to B.

Two formulas are used with such a design that yield a quantitative measure of the transfer amount. The first is directly derived from the raw task scores:

(1) $$T_{raw} = E_B - C_B$$

In this formula (and in this section), T stands for Transfer effect, E for Experimental score and C for Control group score. It is assumed that higher scores mean better performance, as is the case with with most test scores. With tasks for which the reverse is true (such as time–on–task and amounts of spelling mistakes), the E_B and C_B terms should change places in order to have a correspondence between positive transfer and a positive value of the formula result.

Singley & Anderson (1989) mention the simplicity and precision of this formula as its strongest points: the amount of transfer is expressed directly in the units in which performance is measured, like number of correct items or time span of task execution. Its major weakness of this formula is that comparisons between different transfer tasks are usually impossible.

For instance: our researcher who is interested in the transfer from grammar skill to writing skill may use a double operationalisation of writing skill. A 100–word dictation and a 20–sentence essay are written by two groups of students. The number of correctly spelled words is calculated for the dictation, as well as the number of syntactical correct sentences for each essay. The experiment is conducted according to the prototypical design: only one group is previously instructed in grammar. Suppose results are that the control group reaches an average score of 50 (correctly spelled words) on the transfer task while the average experimental group score is 60. Besides, on the syntactical dimension, the average control score may be 10 (sentences), while the average experimental score is 15. Although the increase in absolute numbers is twice as large for spelling as for syntax, the researcher cannot conclude anything about the relative quantity of these transfer effects: it is not allowed to interprete the two absolute differences scores in the same way. Specifically, it would *not* be allowed to conclude that 'grammar instruction transfers twice as much to spelling as to syntax'. Thus, a conceptualization of transfer as 'difference between experimental and control scores' suffers from lack of comparability due to the different scales of measurement.

An adaptation of this formula can be found which takes away part of this objection. In the new formula, the difference scores of the experimental and control group are normalized by expressing them as a percentage of the control group level of assessment:

$$(2) \qquad T_{\%\ \text{improvement}} = \frac{E_B - C_B}{C_B} \times 100\ \%$$

This adaptation is quite often used (Hayes & Simon, 1977; Smith, 1986) because it offers a more comparable interpretation of amounts of transfer to different tasks.

For instance, if the difference in 'spelling score' between control and experimental group is 10, and the mean control group score was 50, the percentage of improvement due to the task A (i.e. the improvement due to transfer) would be 20 %. In other words, the experimental subjects would score 20 % better than the controls, due to having learned grammar task A. If the mean test scores for the syntactical test would differ 5 points between the two groups, and the mean control score was 10, then positive transfer from learning task A to syntactical skill would be 50 %. The researcher would conclude that grammar skill transfers more than twice as much to syntactical skill as to spelling.

It was still felt (Katona, 1940) that such measure depends too much on the nature of the transfer task, instead of possible transfer virtues of the learning task. The following example may make this clear: if the grammar instruction A) concerns learning the difference between 'defining relative clauses' and 'non–defining relative clauses', what will be the effect on B) writing skill, as measured by the syntactical quality of a 20–sentence essay? There are comparably very few problems with defining and non–defining clauses in most essays; therefore the gain will anyway be very small (e.g. 1 or 2 sentences on a total of 20). The transfer amount calculated with formula 2 will then be comparably small (10 or 20 % if the average control score is 10), *even though* the acquired knowledge (about the difference between defining and non–defining relative clauses) *transfers completely* to the transfer task: the experimental students make no errors at all in the relevant sentences. Theory tells us here that the shared learning component in the two tasks is likely to have transferred, but this is hardly reflected in the result of formula 2. The conceptualization of transfer as 'the gain of an experimental group, as a proportion of the control group score' does not satisfyingly represent such transfer phenomena.

For this reason, the concept of *'learning possible'* received a place in transfer formulas. The rationale is that a meaningful formula expresses the gain resulting from learning an initial task A as a percentage of *the amount of learning possible in task B*. If task B exists of components of which some are already mastered or capable of only small improvement, and others still need to be learned, then transfer effects must be aimed at these latter components.

Katona (1940) offers an approach to the assessment of this amount of learning possible that is still in use nowadays (e.g. Schmidt & Young, 1987; Singley & Anderson, 1989). It requires a design in which the control group performs task B during the learning situation as well as in the transfer situation. In this way, the 'learning possible' is determined by the difference of both control group scores, and 'transfer effects' by comparing experimental and control group scores:

Figure 2.5: *Transfer design in which 'learning possible' is assessed*

Group	Learning Task	Transfer Task
Experimental	A	B_1
Control	B_1	B_2

(B_1 means: score on first trial/set of trials on task B; B_2 mean score on later trials)

The formula that matches this transfer design is:

$$(3) \qquad T_{\%\ improvement} = \frac{E_{B1} - C_{B1}}{C_{B2} - C_{B1}} \times 100\ \%$$

The learning possible is in this formula represented in the denominator: the difference between the initial control performance in the learning (B1) and in the transfer (B2) situation. This learning possible is obviously determined by the duration of learning. It is the researcher's responsibility that the learning tasks for experimental and control groups are comparable in duration and other relevant factors, so that the 'learning possible' and the associated transfer effect can be interpreted as 'possible within the time span of the learning situation'.

The gain resulting from learning task A is represented in the numerator: the difference in experimental group score and initial control group score. The transfer amount is calculated by the ratio of these two, and expressed in a percentage.

Suppose that, in the example, C_{B1} would be 10 syntactically correct sentences, and C_{B2} would be 15 on the transfer task; suppose further that the experimental group score on E_{B1} would be 12 sentences (2 sentences more than C_{B1} due to their mastery of task A). Then the transfer effect would be 2/5 or 40 %.

Under the assumptions that the denominator cannot be negative (i.e. training task B does not make performance worse), and that higher scores stand for better performance, we can discern three situations with respect to the scores of Experimentals and Controls on their initial performance on the task B:

1. $E_{B1} < C_{B1}$ –> Exp. performance worse than Controls –> transfer is negative

2. $E_{B1} = C_{B1}$ –> Exp. performance same as Controls –> transfer is absent

3. $E_{B1} > C_{B1}$ –> Exp. performance better than Controls –> transfer is positive.

In the last situation, a special phenomenon may occur:

3a. $E_{B1} > C_{B2}$ –> Exp. performance better than Controls on the Controls' transfer trials –> transfer is positive; a case of 'supertransfer'.

In the last situation, learning task A results in even better performance of task B than directly learning task B itself. Such situations occur when mastering task A is a necessary precondition to learning a more complex task B, the direct mastering of which would be out of reach of the students under study. Such situations are rare, so the measure will vary usually between 0 and 100, with negative values for cases of negative transfer.

Formula (3) also enables comparisons between transfer effects on different tasks. The researcher may have organized a spelling training for the control subjects, which takes place on the same moment that the experimental group is instructed in grammar (defining vs. non–defining clauses). If we add to the previously given research data (syntax: $C_{B1} = 10$; $E_{B1} = 12$; $C_{B2} = 15$; and spelling: $C_{B1} = 50$; $E_{B1} = 60$) that the controls *after* their training had an average spelling test score of 65 (C_{B2}), then we can calculate the transfer effects (in percentages) of learning task A (grammar instruction) for both transfer tasks:

$$T_{\%\ \text{improvement}}\ (\text{spelling}) = \frac{60 - 50}{65 - 50} \times 100\ \% = 67\ \%$$

$$T_{\%\ \text{improvement}}\ (\text{syntaxis}) = \frac{12 - 10}{15 - 10} \times 100\ \% = 40\ \%$$

The researcher would conclude that learning grammar is *almost twice as beneficial* for improving spelling as for improving syntactical skill, *with respect to the possible improvement* of these two latter abilities in the given time span, on their respective levels of mastery.

Thus, formula (3) turns out to be superior, with the concession that the amount of transfer measured is related to the (duration of the) instructional program used in the experiment and (represented by $C_{B2} - C_{B1}$) and to the initial skill level (represented by C_{B1}). It is the researcher's responsibility to guard the external validity of the transfer measure by choosing a programme and a sample of subjects that will enable him to generalize the results to some intended population..

More transfer designs have been developed and used with one or more sometimes helpful modifications for special purposes (Gagne, Forster & Crowley, 1948; Woodworth & Schlosberg, 1955; Ellis, 1965). Since our study is about mutual transfer between writing and reading skill, we can make use of a special design. The following design is in fact a combination of two standard designs (as in Figure 2.4). It is a symmetrical design by which mutual transfer effects between two tasks or skills, A and B, can be assessed. The two experimental groups are each other's control group, and set each other's standard for 'learning possible':

Figure 2.6: *Transfer design enabling measurement of two tasks' mutual transfer effects.*

Group	Learning Task	Transfer Tasks
Experimental 1	A1	A2 & B1
Experimental 2	B1	B2 & A1

It is necessary to systematically vary the order of transfer tasks for each group (e.g. half of the experimental group 1 performs task A2 first, and then B1; the other half performs B1 first, and then A2). This will avoid structural bias by transfer tasks influencing each other.

Formula (3) can be easily modified accordingly to be used with this design:

(3A): transfer of skill A to skill B:

$$T_{\%\ \text{improvement}}\,(A \rightarrow B) \quad = \quad \frac{E1_{B1} \; - \; E2_{B1}}{E2_{B2} \; - \; E2_{B1}} \quad \times 100\ \%$$

The effect of learning A (for instance: writing) on transfer situation B (for instance: a new reading task) is determined by considering E1 as an experimental group and E2 as a control group. Vice versa, the effect of learning B (for instance: reading) on transfer situation A (for instance: a new writing task) can be assessed by considering E2 as experimental and E1 as control group:

(3B): transfer of skill B to skill A:

$$T_{\%\ \text{improvement}}\ (B \rightarrow A) \ = \ \frac{E2_{A1} \ - \ E1_{A1}}{E1_{A2} \ - \ E1_{A1}} \times 100\ \%$$

We will come back to the design issue in each of the reported experiments (ch. 4, 5, 6).

In sum, several methods were discussed for determining the amount of transfer from a certain task A to a consecutive task B. Two designs were presented of which the last enables the investigation of reciprocal transfer effects between two tasks within one experiment. The merits were discussed of a transfer measure that can be interpreted as the experimental gain, resulting from prior learning, as a proportion of the total learning possible on the transfer task.

2.7 Conclusion

Having come at the end of this chapter, I will summarize the findings. Our study is aimed at investigating possibilities for transfer between writing and reading instruction. Such transfer can be expected for theoretical reasons, which will be discussed in the following chapter. In short, both reading and writing activities rely partly on the same knowledge (a related set of coding/decoding procedures) and elements of this knowledge may be susceptible to transfer.

The type of transfer between writing and reading that we would like to investigate can be described, using the dimensions in section 2.2, as:

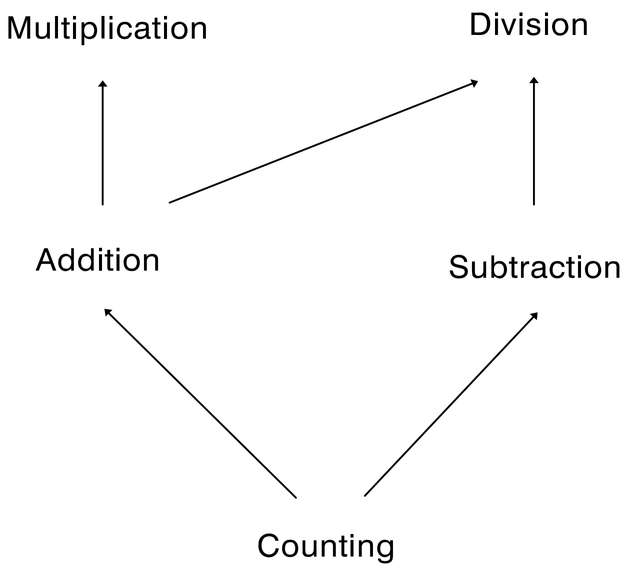
positive,	– prior learning enhancing transfer performance
specific,	– within the same text types and on the same text–structural level
lateral	– the skills having no part–whole relation
comparably far	– to the complementary mode
forward	– stimulated by an intervention in the initial learning situation
transfer–of–learning	– to a new learning situation (new reading task following writing, and vice versa).

I have distinguished two lines of research in historical as well as more contemporary approaches to transfer. The oldest line of research is mainly occupied with the specification of task or skill elements, leading to a better understanding of how people perform these tasks, and to possibilities for curriculum enhancement by ordering and fine–tuning the tasks on the basis of their structural relations. This line of research is continued by cognitive scientists, who attempt to give more detailed, but flexible and quantifiable task analyses by means of production systems.

A more recent line of research puts the accent on factors regarding learning and instruction that influence knowledge or skill acquisition and retrieval. Although the educational value of many supposedly transfer–influencing factors has not yet been determined, two things are evident. First, there is no guarantee for automatic transfer between two, even closely related, tasks; therefore instruction must be specially aimed at transfer. Second, transfer activities usually concern abstraction (for comparably far transfer) or automatization (for comparably near transfer). Most learners must be prompted by instruction to encode or adapt their knowledge in such a way that it will be useful in potential transfer situations.

In this study, elements of both lines of research will be followed. In the next chapter, I will venture to give a theoretical specification of the 'identical elements' that may transfer between related writing and reading tasks. Consequently I will discuss a type of instruction that may stimulate learners in acquiring the 'identical elements' in such a way that transfer is more likely to occur. In short, the next chapter contains an elaboration of more general learning and transfer theories within the domains of reading and writing.

Figure 2.1: *Two prerequisites for division:*



CHAPTER 3: Learning and Transfer within the Domains of Reading and Writing

In this chapter possibilities are considered for effective acquisition of, and transfer between, reading and writing skill. The specification of 'identical elements' is done from a communication–analytic perspective on writing and reading, more specially on their complementary character. In a search for effective instruction, a learning–by–doing approach to instruction is criticized for its relative weak support of self–observation and self–evaluation activities, resulting in non–optimal learning. A social–cognitive theory, advocated by Bandura and Schunk ('learning by observation') stresses observation and evaluation of task execution activities; this theory is used for developing types of instruction that may promote learning as well as transfer of writing and reading skills. Two recent empirical studies are discussed in which the observation of communication processes has been an effective intervention. The chapter is concluded with the research questions that are central to the empirical part of this study.

3.1 Introduction

The previous chapter offered a rather general, domain–independent perspective on transfer–of–learning. In this chapter, I will attempt to apply elements of this general perspective to the domains of writing and reading. Promoting transfer within or between these domains is desirable, since it may contribute to the effectivity of instruction as well to the quality of learning (integration of the acquired knowledge). Obtaining such transfer is not easy; a study by Rijlaarsdam, Van den Bergh and Zwarts gives some evidence that incidental or 'automatic' transfer is not likely to occur even within the same mode (Rijlaarsdam, Van den Bergh & Zwarts, 1992). Thus, there is a need for investigating and developing transfer–effective instruction concerning writing and reading.

In section 2.5, three groups of factors in the learning situation were considered essential for explaining or promoting transfer within or across domains (cf. Simons & Verschaffel, 1992, 8–13):

a) LEARNING CONTENT
A theoretical analysis of learning goals (task, behavior, knowledge, skill) in the initial learning situation and in the transfer situation, in which it is specified on which grounds transfer may be expected (e.g. 'identical elements').

b) LEARNING ENVIRONMENT
A theoretical analysis of factors in the learning environment (instruction, materials) that influence the chance that prior learning will be usefully applied in a transfer situation. This concerns factors in the initial learning situation (promoting forward transfer) and/or the transfer situation (promoting backward transfer).

c) LEARNER CHARACTERISTICS
A theoretical analysis of subject characteristics (intelligence, prior knowledge, motivation) that are supposed to influence the occurrence the way in which, or the extent to which transfer occurs.

For our goals, the development and testing of effective and transfer–promoting instruction for writing and reading education, this general perspective must be applied to the domains of writing and reading. The type of transfer that is aimed at is *lateral* (either intramodal: between similar tasks within one mode, or intermodal: from writing to reading or vice versa) and *forward* (the influence of instructional factors within the initial learning situation on the application of learning results).

This application leads to the following goals of this chapter. Concerning 'learning content': an analysis must be given of writing and reading in which sources for possible transfer are presented and explained (section 3.2). Concerning 'learning environment': a well–argued choice must be made from the transfer–promoting instructional factors in section 2.5: which type of learning activities may contribute to both learning and transfer in the domains of writing and reading? To answer this question, the merits of two instructional methods are discussed and contrasted: learning–by–doing (section 3.3) and learning–by–observation (section 3.4) The choice is made also on the basis of empirical intervention studies that have been successful in obtaining transfer (discussed in section 3.5). Next, the theory must be integrated and adapted to form a basis for the present study (section 3.6) and research questions must be derived which will be central to the experiments in the empirical part of this study (section 3.7).

I want to make explicit that this study is not occupied with 'learner characteristics' and the way in which they may influence writing–reading transfer. Although I consider these characteristics as certainly relevant to the study of transfer and to the validation and actual implementation of research results, I have given priority to the investigation of task– and instructional factors. The interactions of these effects with subject characteristics is, at least for the present, outside the scope of this study.

3.2 Relationships between reading and writing skill

The supposed relationship between writing and reading skill has been since long the object of study. Although most studies into written language skills are aimed at either writing or reading, there are also many that try to shed light on the connection between these skills (for overviews see Stotsky, 1983; Kucer, 1985; Shanahan & Lomax, 1986).

A simple indication for the idea that writing and reading are connected is given by correlational studies of written language skills (e.g. Van Gelderen, 1987; Van den Bergh, 1989; Rijlaarsdam, Van den Bergh & Zwarts, 1992). Although the results of such studies vary, moderate to strong correlations between reading and writing tests have repeatedly been found. However, an important objection to such correlational approaches is that they are usually poor on theory and do not offer a description or explanation of the way in which the skills are psychologically connected. Moreover, alternative explanations for the suggested causal relations (e.g. so called 'third variables' such as intelligence) are insufficiently ruled out. For our aim, the description of cognitive elements that are shared by writing and reading, we can therefore not rely on the correlational approach wich takes the empiry as its starting point.

Studies with a theoretical starting point usually result in a model–like representation of reading and writing processes as complementary activities. Models of the writing and the reading process have been integrated in various ways in overall–processes of text production and processing (e.g.

Pearson & Tierney, 1984; Kucer, 1985; Pool & Van Wijk 1995). Like in models that only represent the reading process (Just & Carpenter, 1987; Rayner & Pollatsek, 1989) or only the writing process (Hayes & Flower, 1980; Hayes 1992; Bereiter & Scardamalia, 1987), also in integrated models are functional categories of activities discerned, such as 'generating', 'monitoring' and 'evaluating'. The categories are connected by arrows that indicate temporal sequences, or streams of information, or regulation activities. The status of these connections is not always clear. Nevertheless, these modelled descriptions of the writing–reading–connections offer more detail than the correlational descriptions, and thus more starting–points for the explicitation of transferable elements.

Models describing the mutual connection between reading and writing can be placed into three categories, depending on the direction of transfer they predict:

Reading–to–writing models were based on the audio–lingual theory of language acquisition, which describes a hierarchy of language skills in which receptive skills are acquired earlier than productive skills, and oral skills earlier than written skills ((Ney 1966, Raub 1967, Miller & Ney 1968, Hall, Moretz & Statom 1976; all cited in Simonsen, 1988). The explanation that learning psychologists gave for this order was that young children first hear the voice of their parents (listening), and imitate it afterwards (speaking), and that this sequence is repeated when mastering written language, which is learned by reading and understanding before children kan produce it.

The problem with these models is that they are hardly informative with respect to how the transfer from reading to writing operates. No specific skills or other elements are mentioned that could transfer in some way. Only the direction of transfer is stipulated, but without explanation or detail within each of the modi.

The production of writing–to–reading–models was stimulated by the viewpoint that reading, like writing, is a constructive process as well. A mental text must be construed during reading on the basis of textual and contextual information and background knowledge ((Rayner & Pollatsek, 1989). The aim of reading as well as writing is the production of a complex message. Transfer from writing to reading is based on the idea that readers can understand verbal messages by using the same knowledge as with which the message was construed. Empirical studies (Combs, 1975; Straw & Steiner, 1982) were based on the hypothesis that instruction in sentence–combination, aimed at the production of complex verbal messages, would contribute to the processing of such messages during reading. Such transfer effects were in fact found in these studies.

The rather limited operationalization of writing as sentence–combining is one of the major weaknesses of this approach. As a consequence, there is little to conclude with respect to transfer from writing to reading; only that such transfer is possible. There are no data on precise psychological mechanisms, the knowledge elements that are transferred and the generality of this type of transfer. Just like in the reading–writing category, the informativity of these models is low.

It can be concluded that these two families of reading–writing–models, which both predict one–sided transfer, are both plausible with respect to the described direction, but that the research falls short in the one–sided orientation on *confirmation* of the model instead of *falsification*. If one assumes that transfer can occur from reading to writing, but not from writing to reading, it is adviceable ot put the complementary hypothesis (transfer in the opposite direction) experimentally to the test with a view to rejection (Popper, 1963). This strategy was not explicitly followed in any of the two approaches. However, looking back we can conclude that confirmative results for each of the families of models yields falsification of the one–sidedness of the other family. If transfer from writing to reading cna be obtained, this is an apparent falsification of a one–sided reading–to–writing model. The fact that transfer can apparently be induced in both directions motivates the creation of a two–sided or <u>interactive model</u> instead of a model predicting one–sided transfer only.

Several researchers have tried to describe reading and writing as related, mutually influencing skills that feed partly on the same mental resources (Shanahan & Lomax 1986; Pool & Van Wijk 1995). Shanahan & Lomax (1986) propose a combined reading–writing model that describes the <u>development</u> of these skills in terms of verbal units. *Readers* learn how the letters of the alphabet and syllables sound, than they learn the meaning of words, then word combinations and propositions, and finally text comprehension is acquired. *Writers* follow a similar road: they learn to spell (splitting up words in sounds, transforming sounds into letters), they learn to write words, to use syntactical knowledge in formulating propositions, and knowledge about text grammar (e.g. story structures). Their model is represented in figure 3.1.

Figure 3.1: *Interactive Reading–Writing model (Shanahan & Lomax, 1986)*

[hier figuur 3.1]

Each of the seven constituting variables was measured with a group of over 500 second– and fifth–graders. The fit of three regression models was compared: the regression of reading variables on related writing variables (representing a one–sided transfer from writing to reading), the regression of writing variables to related reading variables (representing one–sided transfer from reading to writing) and the regression of related reading and writing variables on each other (representing the interactive model). The last regression appeared to fit best. This result indicates that, at least for these age groups, the development of writing and reading is synchronous, and the possibility of mutual transfer is still open.

Congeniality or relationships between the variables, which are indicated in the model by arrows, can be interpreted as relying on common knowledge. We will discuss a few examples. An example of 'receptive' knowledge, that is knowledge that is used in understanding or decoding verbal information, is:

'The sign <letter> *sounds like <sound>'*

or, in a developmental stadium in which complete words are recognized:

'The word <word> *means <concept>.'*

For learners who asssume a 1–to–1 relationship between sound and verbal symbol, the first knowledge element is equal to its 'productive' transformation:

'The sound <sound> looks like <letter>*"*

or, on a word level:

'The concept <concept> is written like <word>*'.*

Accordingly, more to the right side of the model, learners will equal the syntactical or textual knowledge:

'An adjective relates to the noun in from of which it is placed'

will be equalled by this learner with the text production rule:

'If you want to relate an adjective to a noun, place it in front of the noun'

Also in this latter case, the learner assumes a 1–to–1 relation between the semantic aspect (the adjective's referring aspect) and the verbal aspect (its order within the sentence). It is plausible that learners assume such 1–to–1 relationships, as a kind of general learning strategy, because it simplifies their understanding of the domain and enables their active functioning on the domain. From the moment learners are confronted with counter examples, they will construct more nuanced extensions of the rule. For instance, many beginning Dutch spellers learn to write the sound <*t*> with the letter <t>, also in cases where this is incorrect, such as 'root' ('ret' instead of 'red') or 'hooft' ('heat' instead of 'head'). Only if they have been confronted for some time with the fact that the sound <*t*> is in some cases written as <d> (namely if a <*d*> is heard in the the conjugation or in the plural form: rood – rode; hoofd – hoofden), they will specialize their productive rule.

A comparable reasoning can be given for homonymous words. A Dutch child who has learned a meaning of the word <*tas*>, (standard meaning: bag; non–standard meaning: cup) will have no problem making its knowledge productive, and ask 'Where is your <tas>, grandma?' In making this knowledge productive, *the child assumes* that other people will understand the intended, unequivocal meaning. Only when this assumption is violated (for instance when grandma replies 'Which <tas> do you mean, honey?') the learning child will abandon the 1–to–1 assumption and specialize its knowledge about the word <tas>.

The language learner acquires much knowledge about relationships between verbal signs and their meaning. They may acquire this knowledge either in a receptive form or situation ('what does this word mean, mama?'; 'where can I find the topic of this paragraph, sir?') or in a productive form or situation ('how do I write <*baby*>, mama?'; 'how should I order my arguments, sir?'). The assumption that receptive and productive rules can be deduced from each other is necessary for the development of communicative skills: writing is not only transforming thoughts into verbal code, but transforming *in such a way that readers may decode it* by using the same set of rules. This is why, in the communicative system that language is, 1–to–1 relations between meaning and verbal sign or code must be more rule than exception (Van der Horst, 1986, ch.6). The verbal signs can be of many levels: from letters through words to sentences and texts.

In an attempt to explicitize the types of knowledge that are at the basis of the related variables, the model of Shanahan & Lomax's can be extended (figure 3.2).

Figure 3.2: *Reading–Writing model (Shanahan & Lomax, 1986) extended with types of grammatical knowledge*

[hier figuur 3.2]

More detailed models, such as this interactive model, offer more possibilities for the development of theory about transferrable elements and transfer mechanisms.

Also in the more recent model of Pool & Van Wijk (1995), writing and reading are sketched as related and parallel processes. This allows again for the possibility of bilateral transfer. The – somewhat simplified – model can be represented as in figure 3.3.

Figure 3.3: *Writing–reading model (Van der Pool & Van Wijk, 1995)*

[hier figuur 3.3]

As a common resource for the reading and writing process they mention 'knowledge about the world', which would form the basis for the conceptual components Inventing and Comprehending. Language users call for this type of knowledge when giving meaning to the texts they read (semantization) as for generating the information that must be transferred[1]. A second resource is the 'lexicon', which must be understood as the source for linguistic knowledge that is related to lexemes: syntactical and conceptual information, phonological and morphological information. The lexical knowledge is used during writing, when the writer transforms his 'invented' meaning into verbal code ('coding'), and it is used during reading, when the reader decodes verbal units and their relations to a conceptual level.

We may assume, even though it is not explicitly mentioned in the presentation of this model, that text–structural and pragma–linguistic knowledge are also part of the 'lexicon'. Writers use such knowledge when structuring verbal units that surpass the sentence level, and readers use it when comprehending such larger texts. For example, textual and pragma–linguistic knowledge must be used to comprehend or produce argumentative, enumerative, and explicative relations between sentences. Also, indirect or implicit language ('Well, John, how come you're so funny today?') is interpreted by using pragma–linguistic knowledge.

In sum, these two interactive models are more informative, since they stipulate which parts of reading skill and writing skill have mutual relations, and – in the case of Pool & Van Wijk – the

---

[1]) Pool & Van Wijk give a meaning to declarative and procedural world knowledge that differs from the Singley and Anderson's (1989) meaning that was presented in section 2.4. Pool & Van Wijk refer to Levelt (1989).

nature of the cognitive elements which account for these relations. This contributes to a specification of reading and writing processes in which 'identical', or at least transferrable elements can be identified.

In order to get more grip on the identical or related elements of writing and reading, we can undertake an analysis of writing and reading tasks using Singley & Anderson's ((1989) production rules, or goal–action–descriptions. We start from a complete communicative transfer between writer and reader, and try to describe, in general terms, the essential cognitive activities that constitute the writer's and reader's task.

In some communication–analytic approaches (e.g. Schultz von Thun, 1977) the communicative information transfer is – in a simplified way – represented as in figure 3.4.

[hier figuur 3.4]

Figure 3.4: *A simple model for communication (Schultz von Thun, 1977)*

The complementary nature of writing (coding) and reading (decoding) processes is evident, and also the instrumental character of the verbal message. It is the writer's task to formulate the message in such a way that it is safe to assume that the reader will succeed in reconstructing the intended meaning. Reversely, it is the reader's task to reconstruct the meaning in such a way, that it is safe to assume that it reflects the writer's intentions (Van Eemeren 1980).

As an example for the analysis of a communicative information transfer we will take a writer who wants to produce an argumentative text with the goal to convince a reader. Thus, the 'message' is in this case an argumentative text. This writer's thinking activities for the performance of the task depend on his understanding of 'argumentative text'. Let us assume that the writer has the textual knowledge:

> *'a standpoint (S) and argumentation ((A) are the essential parts/properties of an argumentative text (AT)'*

for instance because he/she has learned this in school. Thus, the writer will know these three concepts, the terminology to describe them, and the mutual relation: AT = S + A.

We can now describe the writing task as a sequence of the following thinking activities.

1. Goal of the task is to produce a text of the type AT.
2. IF (goal = produce text of type X) THEN (actions = realise the essential parts or properties of type X)
3. *'a standpoint (S) and argumentation (A) are the essential parts/properties of an argumentative text (AT)'*
4. IF (goal = produce text type B) THEN (action = realise parts/property S and realise part/property A).
5. So for this task I must realize an S and an A.

In this case, thinking step 1 is the part of the task orientation: determining the goal. Thinking step 2 is the actualization, based on the goal set in step 1, of a general production rule for text composition. Thinking step 3 is the actualization of textual knowledge relevant for this particular task. Combining steps 2 and 3 yield the specific production rule 4. It must be noticed that rule 2 is more flexible than the rule in step 4: variables (X) are part of step 2, while substitution of the information from step 3 yields a production rule made of relative 'constants' (AT, S, A). In step 5, the result of combining steps 1 and 4, a conclusion is drawn about how to execute the task.

Executing a complementary reading task can be described as:

1. Goal of this task is to typify the text.
2. IF (goal = typify text as type X) THEN (actions = determine essential parts/properties of type X).
3.. In this text, I identify a part/property S and a part/property A
4.. *'a standpoint (S) and argumentation (A) are the essential parts/properties of an argumentative text (AT)'*
5.. So in this text I identify both essential parts of a type B text.
6.. So I typify this text as a B.

Thinking step 1 is again part of the reader's orientation on the task. For an effective text interpretation, readers may want to determine the type of 'message' first: is it for instance an informative text, an argumentative text, an excuse or a request? On the basis of this goal, a reading strategy is actualized (step 2). The result of this strategy (step 3) is connected to the actualized textual knowledge (step 4), so that the reader can make the correct inferences as to this type of text (steps 5 and 6).

The text may contain many more parts than only S and A, of course. The inference made in step 6 may be premature, or only partially correct (the text can contain other parts than argumentative, such as informative parts or requests). However, it is important to note that the reader, when assessing the nature of the text, relies on the cognition activitated in step 4.

The corresponding element that is used in the writing as well as in the reading task is printed in italics. It is the verbal, in this case textual, knowledge that resembles the subject–matter as offered to students in language skill instruction. We can consider this knowledge element as a possible basis for transfer between both tasks: someone who has learned to write texts on the basis of such knowledge, will be more likely to recognize this type of texts from a reader's perspective. Reversely, someone who has learned to recognize the essential parts in such texts, is more likely to construct his own argumentative texts accordingly.

Like in the examples of the letter <t> and the word <tas>, a learner will assume that the relation between type of message and constituting elements is reciprocal: 'AT ⇔ (S + A)'. That is, a text of the type AT should always contain the parts S and A (and no other parts *instead of* these parts), and a text containing the parts S and A always belongs to the type AT. If learners use the strategy to consider the relation between verbal code and meaning, as specified in the receptive or productive rules, as reciprocal, then they can adapt a rule to the complementary mode.

In fact, this is a process of *knowledge adaptation*, which is characterized by a *reversion* of the coding or decoding rule. We can schematize the distinction between the linguistic 'identical element' and the (de)coding rules for both modi:

linguistic knowledge:	textual property (B ⇔ S+A)	
	⇙ ⇗	⇖ ⇘
coding/decoding rules:	writing (coding): IF (goal = write text type B) THEN (actions = write S and A) (because B ⇒ S + A)	reading (decoding): IF (goal = typify text as B) THEN (actions = identify S and A) (because S + A ⇒ B)

In my opinion, the probability of transfer is possibly co–determined by the way the learner understood and stored the relationship between AT on the one hand and S and A on the other hand: as a reciprocal relationship (indicated as 'linguistic knowledge') or as a non–reciprocal relationship (either the coding or the decoding rule). If the rule is stored directly in the reciprocal form, abstracted from the mode in which it is practised, it requires less adaptation in a transfer situation. The learner is, as it were, already on Salomon's 'high road' (Salomon e.a., 1987; 1990): a certain forward transfer has taken place. This is the kind of decontextualization that teacher can stimulate by explicitly mentioning other usage contexts for subject matter. If, however, the presentation of subject matter and the type of exercises are aimed at one mode only, leading tot a non–reciprocal form of the stored information, the learner will have to put some extra effort in adapting his knowledge in the transfer situation (backward transfer).

In sum, a class of linguistic knowledge can be invoked for reading as well as for writing activities. Reading and writing are related in the application of such knowledge in (productive) coding rules or (receptive) decoding rules. The linguistic knowledge is considered as an 'identical element' that may transfer between reading and writing tasks. Coding and decoding rules are if–then–productions in which the linguistic knowledge is connected with either a productive or a receptive goal. The coding rules in which the linguistic knowledge is incorporated indicate how conceptual information should be transformed into verbal code. This transformation can operate on all verbal levels, from letters through words to sentences and texts. Decoding rules specify, reversely, the transformation of verbal information into meaning (conceptual information). Since coding and decoding rules are strongly related by the linguistic knowledge on which both are based, it may be assumed that these rules as well as the linguistic knowledge elements are potential candidates for transfer. The transfer may be obtained by a process of *reversion* of the relation expressed in the coding/decoding rules.

3.3 A problem analysis of learning–by–doing

In the previous section, it was shown how linguistic knowledge can be integrated in coding or decoding rules, in which conceptual information is linked to their verbal expression. I suggested that this linguistic knowledge can function as 'shared element' of writing tasks and their complementary reading tasks, and that it is as such a potential candidate for intermodal transfer. The coding/decoding rules can be represented as productions in the way described by Singley & Anderson (1989). Applied to one of the two modes during learning, the piece of knowledge can be adapted or 'mirrored' to become useful for communication within the other mode. These activities are in line with recent learning theories that stress the importance of active knowledge construction, rather than the reproduction of knowledge that is handed over during instruction (Parys & Byrne, 1989; Smith, 1993).

If the theoretical possibility of transfer between reading and writing is granted, the next question is how to promote this transfer through instruction. To answer this question, I will make a problem analysis of a traditional, deductive kind of language skills instruction, referred to as learning–by–doing. It consists of presenting a prescriptive rule for language use, which the learner must apply in one or more (writing or reading) exercises. 'Application' and 'practice' are the key words for the learning–by–doing method of instruction, which is based on the idea that skills can only be acquired by repeated and varied exercise:

> *'For most academic skills like reading, writing and arithmetic, there simply is no substitute for repeated practice. Only with much practice will these habits become automatic and be performed rapidly and effortlessly.'* (Bower & Hilgard, 1981, 539–540).

In this citation, the authors stress the importance of practice for learning complex skills. Their position is that extensive practice is a necessary condition for the acquisition of expertise in the domains of reading and writing. It would be very difficult to find a teacher, educator, or educational psychologist who disagrees with this viewpoint. Even more, looking at current school books and teaching practices for language skills education, one will find that many teachers and educators have adopted this <u>necessary</u> condition as a <u>sufficient</u> condition (Hillocks 1986; cf. De Glopper 1988). <u>learning–by–doing</u> is the dominating instructional method in the everyday practice of language skill education. The effectiveness of practice as a learning tool, so widely relied on, can however be questioned. In this chapter I will make a theoretical comparison of a learning–by–doing approach to reading and writing with an instructional method which I will call 'learning–by–observation'.

Two reservations should be made regarding the problem analysis. First, I will – for the sake of clarity – comment on a very *simple* type of learning–by–doing: presenting subject–matter, doing exercises, and receiving summative feedback (as described in section 1.2). Of course other, more advanced types of learning–by–doing exist, and various forms of practice and feedback. The problem analysis offered here may not be equally valid for these other types. However, I consider it as one of the main problems of secondary reading and writing education that instruction is often based only on this simple type of learning–by–doing. Therefore, and for the sake of clarity, this simple type will serve here as a contrast to alternative types of instruction. Second, in making a problem analysis I do not intend to discredit learning–by–doing in general. I believe that learning by doing exercises is an essential part of skill education. Nevertheless, it makes sense to detect weak spots of this method, and to invent alternative or supplemental learning activities which can enhance the effectivity of language instruction in general.

At least two arguments oppose the idea that learning–by–doing deserves a status of unassailability. In the first place, the learning output of practice is not equal for every student. Some learners, so–called 'good novices', manage to profit more from practice or exercises than others, even within the same IQ subclasses (Elshout & Veenman 1992; Veenman 1993). Apparently effective skill acquisition is induced by more factors than practice alone. The optimalization of any instructional method requires insight into these additional factors which modify the effect of practice on skill acquisition.

In the second place, learning–by–doing may be a very good method to train automatic, 'rapid and effortless' skill execution – to speak with Bower and Hilgard. However, many skills to be acquired in school are made up of more than only of knowledge proceduralized by practice. This is particularly the case in more complex task domains, such as mathematics, literature, text analysis, or essay composition. Automatic activities should be strategically alternated with more conscious mental activities: the systematic error detection of one's own solution to a problem, the deliberate reflection on one's habits or attitudes towards a topic, or conscious self–regulation of one's task behavior, which may otherwise become less systematic and effective. In short, learning in complex domains often calls for the learner's *self–reflecting* abilities, with the aim of enhancing their *self–regulating* activity during task execution. In this section it will be discussed whether learning–by–doing is also an optimal method for developing such self–reflecting abilities.

Educational psychologists have called attention for the crucial role of self–monitoring or self–observation (perceiving one's activities during task execution), self–reflection (processing the output of monitoring by evaluation, abstraction and attribution) or self–regulation (controlling the task execution for the sake of its effectiveness, based on information from self–observation and self–reflection) (e.g. Simons & Beukhof, 1987; Kuhl & Kraska, 1989; Vermunt, 1992; Ng & Bereiter, 1992; Schunk & Zimmerman, 1994). Many of them place their theory in the context of learning processes. However, it is important to note that regulation of *learning* a certain skill requires – in theory – regulation of the *executive processes* of that particular skill. Task regulation conditions the regulation of the learning process for that task (Ng & Bereiter, 1992). By comparing the processes of *executing* a task and *learning to execute* it, I will try to clarify the key mediating function of self–monitoring and self–reflective activities. I will then discuss the possibilities or limitations of learning–by–doing regarding these activities.

Taking writing as an example, a skill that has often been conceived of as a problem–solving activity (e.g. Hayes & Flower 1980a,b, 1986), we can divide the cognitive activities aimed at resolving a writing problem in *executive* activities aimed at text production (orientation, writing and revising activities), *monitoring* activities aimed at on–line knowledge of one's actual task behavior (self–observation, evaluation and reflection) and *regulative* activities aimed at strategic control of the former types of activities, dependent on their evaluation (see figure 3.5).

[insert fig. 3.5 about here]

Figure 3.5

Functional relationships between levels of performance activities

In figure 3.5 executive and monitoring activities are placed on levels I and II respectively. An effective temporal organization of these activities is governed by regulative activities placed on a third level. Straight arrows indicate the flow of information between activity categories and curved arrows indicate activation prompts. In this representation, a central position is given to monitoring and evaluative activities, since they supply the knowledge base for skilful regulation and thus execution of the entire writing process. Being aware of one's writing activities and their consequences is an essential step towards detecting possible flaws in, and enhancement of one's writing behavior. Thus, good writers invest in being aware of their activities during the course of the writing process.

How is this perspective on performance regulation related to learning? Writers in a learning situation, like students at school, will consider a writing task as being part of a learning task. In fact, they must execute two processes at the same time: a writing process (with a material aim: producing a text) simultaneously with a learning process (with a cognitive aim: acquiring skill in producing such texts). This 'parallel' learning process can be represented with the same morphology as the writing process, including executive activities (orientation on the learning task, performing learning activities), monitoring activities (self–observation and evaluation of learning activities) and regulation of learning (e.g. starting over again, or skipping parts of the exercise). The writing task should be instrumental to the learning task, which implies that the way in which these tasks are connected influences the quality of learning.

This connection (between the writing and the learning task) lies in the writing experiences from which students can learn. To be instructive, writing rules, techniques, strategies must not only be executed by the writer; they must be monitored, conceptualized ('given a name'), experienced, along with their positive or negative effects. Writers use their 'writing experiences' or 'writing evaluation' (the output of the monitoring processes on level II) as input for their learning. These experiences consist of conceptualizing writing behavior (what am I doing now? how should I call it? which strategy must I choose? have I done anything like this before?) and evaluative labeling (this strategy has been very time–consuming; the brainstorm was, or was not, successful; such a sentence may conclude the text). Writers who put some effort in realizing and evaluating their working–method during (or after) writing, invest in the meaningfulness and effectiveness of their learning.

In sum, learning–to–write by doing writing exercises appeals strongly to the learner's self–observing and self–regulative capacities. The same can be said about learning–to–read. The regulation concerns three aspects of the task: a) learners should follow a 'double agenda' with some activities aimed at the production or reception of a certain text, and other activities aimed at learning; b) they need to effectively alternate executive and monitoring activities for each of the agendas, and c) they must orderly control a variety of executive activities for the composition or comprehension of text. Even though some processes will take place without conscious attention, still a large part of the learner's cognitive capacity will be occupied with regulation.

Obviously learning–by–doing demands a lot of the learner's self–monitoring and self–regulative abilities. Of course, it is not possible nor productive for learners to be constantly aware of all of their mental activities. This would be cognitively too demanding, and young learners are probably not yet able at all to switch between task execution and self–reflection. I do not advocate such 'permanent awareness'. What I do want to advocate, is that in a learning–by–doing paradigm, the self–monitoring and self–reflective activities are likely to suffer from a student's tendency to focus on (short–term) task performance instead of (long–term) learning performance. This may be a weak spot of the learning–by–doing method, for which compensation ought to be found.

Considering the importance of self–monitoring and evaluation for learning complex skills, I suggest that a simple learning–by–doing type of instruction for reading and writing may not be supportive enough for every student, perhaps not for most of them. While some students are keen on finding instructive aspects in even simple writing or reading assignments, others will just 'go through the motions' and fail to observe, let alone improve, their writing behavior; mostly because a strategic distribution of attention across learning and writing levels is beyond their capacities. Such students need a type of process–oriented instruction that offers more support for their monitoring of writing (or reading) activities, and in evaluating their effectiveness.

If the learner does not fulfil the conditions to effective learning–by–doing, learning tasks may go wrong on each of the three aspects a), b) and c) mentioned above. The task may go wrong on aspect a) when a student does not consider a school assignment as something possibly instructive, but as a 'job' that needs to be finished as soon as possible. The task execution process, as represented in figure 3.5, will then be initiated and will yield some reading or writing product, but the learner will not invest cognitive energy in attaining learning goals.

Problems regarding aspect b) may arise when learners fail to orientate themselves on the task, if they fail to realize which task activities they are undertaking, or if they are uncritical toward their performance. A poor orientation on the writing or reading task may lead to inefficient or even inadequate task behavior: no clear goals are set, existing strategies are not activated, or no planning is done. Poor orientation on the (parallel) learning task may result in incorrect or hampered integration of new cognition into existing cognition, so that inconsistencies and misconceptions arise (Ausubel, 1968; Ali, 1990) or acquired strategies are not connected with their potential usage contexts and remain inert (Simons, 1990). If learners do not realize what they are doing (no monitoring), no cognitive representation may be constructed of the working method (conceptualization) that can sink in memory. If students are uncritical toward their activities – either on the task level or on the learning level – they may maintain inadequate ones, or refrain from acquiring better ones because they do not observe their positive effects. The same may happen if learners do not evaluate their performance because external feedback is lacking, inadequate, or remote.

Finally, students can have problems with regulation on the executive level. Complex tasks are called complex because they call for may activities of a varied nature. Counting two–letter words on a page calls for many mental activities, but it is not a complex task. Writing an essay, on the other hand, is complex because the constituting elements vary (task definition, brainstorming, selecting information, ordering, organizing, formulating, revising etc.). The organization of these subprocesses across time can be more or less successful; a strategic regulation is therefore desirable but also puts a burden on the refulative capacities of the learner.

In sum, one of the weaker spots in a learning–by–doing paradigm may be the lack of support for monitoring and reflective (evaluative) activities; they are left to the initiative of the learner. This lack is in our analysis a bottleneck for effective learning, since learning processes obtain their input from monitoring of and reflection of the task execution process. In this way, the learning process is a derivative of the task execution process: for students to assess if they have mastered a class of problems or tasks, it is necessary to prove themselves that they actually can solve some of these problems. In doing so, they must conceptualize the task behavior (or problem strategy) and label it according to the perceived results or feedback. Students should be able to assess if their task performance is correct of not, or else they will never assess to which extent they master a task – and will finally not become more responsible for their own progress in learning (Vermunt 1992).

For some tasks, it is far from easy for a learner to find out in which aspects the task behavior was adequate or not. It is not only important in skill acquisition that students learn to perform some (mental or other) activities, but also that they acquire the criteria by which the activities and their results can be evaluated. For instance, in the composition class students are instructed that texts of a certain length should start with an introduction. The problem is initially not whether the students write an introduction or not (since they are explicitly instructed to do so) but if they can check if what they have written satisfies the criteria for 'proper introductions'. If these criteria remain unclear to the student, no cognitive elements of that kind can be constructed and positively labeled, and the instruction may well remain ineffectual.

Therefore in this study, instructional factors are stressed which may reinforce the monitoring and reflection components of task execution, in comparison to learning–by–doing. We will now look more closely at such factors.

3.4 An alternative instructional method: Learning-by-observation

In this section, I will focus on *learning–by–observation* of writing and reading activities, which are performed by others than the observer. Thus, for the observers it is no longer necessary to simultaneously perform, observe, evaluate and regulate *task execution* activities (e.g. writing or reading) as well as *learning* activities (e.g. learning to write or learning to read): they can more conveniently concentrate on the observed writing or reading activities.

In order to evaluate and reflect on one's own task execution when learning–by–doing, the learner must perform *self–observation* activities: observation *by* the learner, *of* the learner's activities. Such self–observation is also stimulated in what is known as 'process–oriented instruction'. This kind of instruction does puts a smaller accent on the critical inspection of qualities of the final product, but stimulates students to recognize and understand their own mental activities that resulted in the product. This helps the students in maintaining or adjusting their working method (De Jong, 1992; Vermunt, 1992; De Jong & Van Hout–Wolters 1993).

An obvious method for process–oriented instruction is *direct instruction* (Pressley e.a., 1987). Students are informed, by verbal explanation and/or pictures, which steps (mental and other) they should take in order to perform a certain task. Depending on how the students interprete and follow the instructions, this method can be more or less successful. Another method is 'modeling', when a person (and expert, usually the teacher) shows the students how to perform the task, after which the students imitate him/her. It is the teacher's task to supply 'corrective feedback' where needed. Thus, external feedback remains necessary for this method, which proved successful in comparisonwith other methods of strategy instruction (Pressley, o.c.).

This method of learning by imitation of models has received much attention in the transition from behaviorist to cognitivit learning psychology (Schunk, 1992). The authority of the 'observational learning' paradigm, Albert Bandura, stressed the idea that people learn by imitation of behavior in their social environment. By evaluating the consequences of observed behavior as positive or negative, people are disposed to take over or to avoid such behavior, without any direct instruction (Bandura, 1977; Bandura 1986; Schunk 1991; Schunk & Zimmerman, 1994). I will now go further into learning–by–observation as defined by Bandura.

According to Bandura (1977), learning can take place in two general ways: in an *enactive way*, when learners performs the new behavior themselves (learning–by–doing) and in a *vicarious* way by observing models performing this behavior (learning–by–observation) or by interpreting verbal information (learning–by–understanding). Of course these types of learning can be combined in an instructional program, but the learning activities are considered essentially different. Enactive learning implies learning from the consequences of one's own behavior, and vicarious learning is learning from the consequences of other people's behavior. We can consider this last type of learning as an extension of 'you should learn from your mistakes', namely 'you should learn from other people's mistakes, so you need not make them yourself' or, more positively, 'a good example is worth following'.

Bandura, who was initially educated in the behaviorist tradition, developed a theoretical perspective on knowledge acquisition by observation which allows us to place him among cognitivists: Bandura takes his conceptual frame and explanative factors for the studied phenomena from the mental system:

> *'Learning is largely an information–processing activity in which information about the structure of behavior and about environmental events is transformed into symbolic representations that serve as guides for action'* (Bandura, 1986, p.51).

This definition bears similarity to the construction of schematic productions or strategic plans for action (Anderson 1990; Van Hout–Wolters 1992). The definition refers to both types of learning, enactive and vicarious.

Modeling, learning by imitation of other people's succesful behavior, has been closely studied by Bandura in a number of contexts and tasks: cognitive, motor, social and self–regulative tasks, regarding the acquisition of knowledge and skills.

Bandura distinguishes *live* and *symbolic* models, depending on the physical presence of the model in the social environment of the learner. He defines *observational learning* as a kind of modeling by which comparably new behavior is observed and imitated. The models supply informtation to the observer about how the task or skill can be performed best. Observational learning consists of four essential subprocesses:

1) directing attention toward relevant behavior components, as demonstrated by competent models
2) enhancing retention or consolidation of the perceived behavior by rehearsing and coding in visual or verbal form, relating it to existing knowledge
3) production and self–evaluation of the modeled behavior, possibly followed by corrective feedback
4) enhancing motivation by understanding the (un)desirable consequences of the behavior, which (de) motivate their use.

Compared to a direct verval/pictorial instruction of processes, or to stepwise exercising of subprocesses, observational learning has repeatedly proven more effective (for a review see Schunk, 1991, ch. 4) in experiments in which teachers are models and supply corrective feedback.

The models do not necessarily have to give 'good examples'. 'Cognitive modeling can include other types of statements. Errors are built into the modeled demonstration to show students how to recognize and cope with them.' (Schunk, 1991, p. 109).

For more than twenty years have Bandura and his colleagues investigated the applicability of observational learning to the field of cognitive skills. Meichenbaum & Goodman (1971) based a self–regulation instruction on the principle of cognitive modeling. They found that the complex tasks they asked their students to perform, were carried out more systematically and resulting in a higher quality. Schunk, a former student and one of the most active colleagues of Bandura, writes about other cognitive tasks (1991, p. 110):

> *'Cognitive modeling has been used in training programs with a variety of tasks and types of students (Fish & Pervan, 1985) and is especially useful in teaching students to work on tasks in a strategic fashion. In teaching reading comprehension, the beginning of the preceding instructions can be modified as follows:* 'What is it I have to do? I have to find the topic sentence of the paragraph. The topic sentence is what the paragraph is about. I start by looking for a sentence that sums up the details or tells what the paragraph is about' (McNeil, 1987, p.96). *Statements for coping with difficulties (e.g.* 'I haven't found it yet, but that's all right') *can be built into the modeled demonstration*'.'

Please notice that modeled processes bear some similarity with the IF–THEN–statements used by Singley & Anderson (1989): IF (goal = find the topic sentence) THEN (action = check for sentence type A and sentence type B).

Bandura specified a number of factors that would influence the nature and effect of learning–by–observation. These factors are outlined in figure 3.6. I will give some elaboratation with a view to the research reported in the empirical part of this study.

Figure 3.6: *Factors influencing the effectivity of learning–by–observation*

Factor:	***Influence:*** **Modeling is more effective...**
Developmental status	.... with older students as observators, because they pay more attention, remember more and are more motivated;
Model prestige and competence	.... with high–status and competent models; observators consider it their task to find out who the best models are;
Vicarious consequences	.... if it is clearly indicated which models are successful and which are not, because it enhances observation;
Expected result	.... if students have high expectations regarding the success of the observed strategy in the domain.
Goal setting	.... if a goal is attached to the behavior that is observed.
Self–efficacy	.... if the observed behavior is believed by the observers to be within their reach; ....and if more than one model can be observed. Seeing more models leads the observer to think the he can perform the task as well.

Although in most experiments on observational learning, the teachers served as models, Schunk (1991, p. 123) gives two arguments for using students as models. One can understand her argumentation as a pay–off of self–efficacy and status/competence: two different reasons two attach importance to the observed behavior. The teacher is appreciated for the expert quality of his/her task execution and the clear formulation of thought processes, or because his/her behavior may seem a good criterion for test behavior ('in order to have a high mark, do like the teacher does'). Peer students, on the other hand, are appreciated because they give the observer an impression that the observed behavior is within reach ('if he/she can do it, then so can I'). Besides: observing different peers is more effective than observing only 1 peer (Thelen, Fry, Fehrenbach & Frautschi, 1979). Some observers prefer *mastery models* (peer students who perform the task correctly; such models show effective task behavior and self–confidence) and *coping models* (peer students whose hesitations and errors are recognizable, so that the observers identify very easily) (Thelen et al., 1979).

Schunk and Hanson (1985) compared the effectivity of 'mastery', 'coping' and 'teacher' models with 'no model' in an experiment with elementary school children acquiring the skill of subtraction. They found that observing both peer models resulted in the largest gain in self–efficacy and subtraction skill, followed by the teacher model and finally the 'no model' condition. In a study by Schunk, Hanson en Cox (1987) observing *one* coping model, *several* coping models or *several* mastery models was found to be more effective than observing *one* mastery model; the skill that was acquired was 'calculus with division'.

By observation of models, learners usually acquire at least a rough approximation of a complex skill (Bandura, 1977). They can refine or correct their skills by further observations or practice, which may be helped by corrective feedback of a teacher. Such feedback is only effective if the observer has already such rudimentary idea about how to do the task (Carroll & Bandura, 19982). However, the learners must become able themselves to distinguish faulty behavior and correct behavior.

It is an interesting question whether learning–by–observation of models can offer an effective alternative or supplement to learning–by–doing, regarding the acquisition of complex skills. Before attempting to answer this question from an empirical perspective in the next section, I will summarize the main points of learning–by–observation:

* learning–by–observation is a process–oriented instructional method which may be compared to teaching by demonstration: the behavior to be acquired is shown and not only explained;
* empirical evidence suggests that some cognitive tasks are more effectively acquired in a *vicariouos* way – by observation and imitation – than by *enactively* performing these tasks;
* the student constructs mental plans which specify the activities for task execution; these plans guide the acquired – and subsequently performed – task behavior;

* observation can be instructive for various types of tasks, including cognitive and self–regulative tasks;
* the effectiveness of observing models depends on the type of model: peer models may be more effective than teachers;
* for different reasons, coping models and mastery models are both valuable; maybe for different groups of learners;
* the role of feedback is ambiguous: corrective feedback by teachers may enhance performance, but it is preferrable if students learn to evaluate the imitated behavior themselves.

At the start of this chapter, I suggested that the learning–by–doing method is very common, while there are reasons to doubt if this method is optimal for the acquisition of complex skill. This method does little to stimulate students' self–observation and self–reflecting activities, and therefore the possibilities for learning may not be optimally used. Next, we saw that social–cognitive learning psychologists handicap a comparable kind of *enactive* learning as compared to several kinds of *vicarious* learning, among which learning–by–observation. In several experimental studies learning–by–observation has yielded good learning results, using both expert ('mastery') models and less expert ('coping') models, regarding executive activities and regulative activities.

In sum, the learning–by–observation method may be a potential alternative for learning–by–doing, because it can yield good *learning* results regarding the acquisition of task strategies. However, this study does not only aim at learning effects, but specially at *transfer* effects. There are several reasons why one can expect more transfer from *observational (vicarious) learning* of reading and writing processes, than from *enactive learning*.

First, we expect more transfer *between* the writing and reading modes as a consequence of larger learning effects *within* the modes. In section 3.3, I suggested that learning–by–doing may demand too much of some learners' self–regulative capacities, and that they would probably not pay much attention to attaining learning goals. If the learning–by–observation method succeeds in partially – and temporarily – relieving the self–regulation burden, in favor of conceptualization and evaluation of the observed activities, then we must find an effect on learning results, which will in turn contribute more to transfer effects.

In the second place, the acquisition of coding and decoding rules by observation may yield results that are qualitatively different from the results of an exercise–centered learning–by–doing method. The question to be answered is whether the transfer of the 'identical elements' (the linguistic knowledge at the base of these rules) must be stimulated through the high road or the low road. The high road certainly qualifies, because intermodal transfer is comparably far. Is the acquisition of a coding or decoding rule less domain–specific when learned by observation than when learned by application in exercises? In comparison with learning–by–doing, learning–by–observation induces a larger psychological distance to the activities (one of the aspects of Bandura's distinction between 'vicarious' and 'enactive' learning) which may lead to a more abstract, less context–bound awareness of the observed strategy. In new task situations, it may be easier to identify other people's approach and to evaluate its result, then to identify one's own approach during execution. Possibly the condition that Salomon & Perkins (1990) put to high road transfer (the active and 'mindful' abstraction of the specific task situation) is fulfilled with less effort during observation from an 'outsider' perspective than during a personal involvement in task execution.

Finally, there is empirical evidence for the idea that observation of language processes promotes transfer. In the next section, the results will be discussed of two different lines of study which are related in that they both investigate the effect of observation activities on learning and transfer.

3.5 Effective observation: Two examples

We can compare the learning–by–doing method of instruction with the learning–by–observation method, regarding the acquisition and transfer of the coding/decoding rules described in section 3.2. Sonnenschein & Whitehurst (1983; 1984) have reported on several experiments comparing several types of observation and practice; they have paid attention to learning as well as transfer effects. Since these experiments have (among others) served as starting point for the experiments reported here, I will discuss them in some detail.

A second line of research was conducted by Schriver (1987; 1989; 1991; 1992). Schriver offers a different perspective on observation of language processes in instructional programs. However, this different perspective is an interesting addition to the learning–by–observation paradigm, and has also proven to be effective; therefore it will receive attention in this chapter either.

Sonnenschein & Whitehurst's interest went out to the acquisition of *referential communication*. The communicative task that Sonnenschein & Whitehurst asked their subjects (young pupils, age 6) to do concerned speaking and listening activities. For the speaking task, two objects were shown to a subject which corresponded on two of the three dimensions 'color', 'size' and 'shape' and were different on the third dimension. The subject had to describe one of the objects in such a way that a good listener could identify the object; in other words, that an unambiguous reference was given using the dimension on which the objects differed. For example, if a big yellow triangle and a big blue triangle were on the table, and the subject had to describe the first, than the answer 'the yellow one' was correct and 'the triangle' was incorrect. For the listening task, subjects had to take the correct object that would match a given description, such as 'take the blue one' or 'take the triangle'. When this was not possible - because the description was ambiguous ('take the triangle') - the subject had to say so. The common rule underlying both speaking and listening task, or the 'identical element', was called the 'difference rule', because a correct performance of the both types of tasks depended on the property that was different instead of equal. After the subjects had given an answer, they received feedback on the result: either 'correct' or 'incorrect'.

Direct training (learning–by–doing) appeared to yield good learning effects within each mode (speaking and listening), as measured by a posttest of comparable speaking or listening tasks (see table 3.1). However, hardly any transfer was found between the modes. Obviously the 'common element' does not automatically transfer from speaking to listening tasks, or vice versa. A group of subjects that practiced with speaking as well as listening tasks reached higher scores for speaking than a control group. This can, however, not be qualified as a transfer effect but as a learning effect since the subjects were (also) trained in the speaking mode; a learning effect had apparently held off.

Table 3.1: *Results of the Sonnenschein & Whitehurst (1983) study; posttest scores indicate proportion of good answers on a series of dichotomous items (base line: .50)*

Training Condition:	Posttest Speaking:	Posttest Listening:
Control:	.75	.67
Speaker training:	.95 *	.65
Listener training:	.77	.78 *
Both training:	.91 *	.68

** = different from control group (p<0.05)*

Later Sonnenschein & Whitehurst (1984) extended their experiment with learning–by–observation conditions. Subjects had to observe models performing speaking and/or listening tasks (giving correct and incorrect answers) and did not have to perform these tasks themselves. In the experimental conditions, several instructional factors were systematically varied: the communicative role that had to be observed (only speakers, or only listeners, or both speakers and listeners: a complete communication); the supply of evaluative comments by the observer (some conditions had to motivate their answer to the test assistant: why was the observed behavior correct or incorrect?); and the reception of feedback (the test assistant would tell subjects from these conditions if their answer was correct or not). The observed performances were unfortunately rather artificial. The subject was either shown a doll who performed the 'speaking task' and told the subject which triangle etc. to pick; or this doll would play a roll–game with another doll who performed as the 'listener'; or the subject would have to speak (or observe other children speaking) to a picture that symbolized a listener.

As learning and transfer tasks, not only 'speaking skill' and 'listening skill' were measured, but also a new task focusing on evaluation: 'commenting on others' performance'. This last task consisted of a series of observations of speakers as well as listeners, with the question whether the performance of each of them was correct or not (four alternatives: base level = .25).

The most striking result (see table 3.2) was that observation and evaluation of both speakers and listeners – the complete communication – resulted in very high scores on all posttests (condition 1 in table 3.2): the performance of speaking and writing tasks and the evaluation tasks. Observation of both roles without giving evaluative comments (condition 2), however, yielded much smaller effects. Observation of both roles without receiving feedback (condition 3) nullified the performance effects. Observation of only one role, either speaker (condition 4) or listener (condition 5) yielded large learning effects (comparable to learning–by–doing, Sonnenschein & Whitehurst 1983) but no transfer effects, neither to the complementary mode, nor to the commenting tasks. Condition 6 is a kind of reflection–condition, not an observation condition: subjects first performed the speaking and listening tasks, and then told the test assistant whether they thought they had performed well or not. This 'reinforced self–observation' yielded high scores on the speaking and listening tasks (which they had trained), but low scores on the commenting tasks (which they had not trained) so we can say that transfer–of–learning for this group is also low.

Table 3.2:

Results from Sonnenschein & Whitehurst (1984): Posttest scores for Speaking and Listening indicate proportion of good scores on a series of dichotomous items. (base line: .50). Criticism test questions had 4 alternatives (base line .25).

Observation Condition:	Observed roles:	Give evaluation:	Receive feedback:	Speaking	Listening	Criticism:
1. Observation:	S and L	yes	yes	.98	.98	.91
2. No evaluation:	S and L	no	yes	.73	.63	.64
3. No feedback:	S and L	yes	no	.49	.50	.27
4. Only speakers:	S	yes	yes	.97	.67	.58
5. Only listeners:	L	yes	yes	.60	.73	.52
6. Self–evaluation	self: S and L	yes	yes	.94	.95	.48

The researchers conclude that speaking and listening tasks are 'subordinated' to the commenting or evaluation task, in the sense that who masters the commenting task appears to master the speaking and listening tasks as well, but not vice versa. In this respect their approach to transfer reminds of Gagné's learning hierarchies (section 2.3). Moreover, they could conclude that an effective acquisition of speaking or listening skill can be obtained by observation and evaluation of others performing such tasks.

This study gives some important hints for the effectiveness of learning–by–observation regarding learning and transfer effects. First, we notice that learning–by–observation is as effective as learning–by–doing, *as long as feedback is given* (as was the case in the learning–by–doing groups). Second, intermodal transfer is not obtained by training in, or observation of, one group only, neither to the complementary communicative skill, nor to the commentary skill. Last, learning and transfer effects decrease dramatically if the subject does not evaluate the observations, or is not told how to evaluate what she saw. We may conclude that some kind of information processing or labeling is desirable to make the observations meaningful. This 'making meaningful' may be all the more important since the subjects are not personally involved in the communication: they are outsiders, looking in. In the next part of this section we will see that observation is also possible from an 'insider' viewpoint.

The 'insider' viewpoint can be recognized in the research line followed by Schriver, who reports on textual and learning effects of an instructional program for writers who observe authentic readers

(Schriver, 1987; 1989; 1991; 1992). Schriver focuses particularly on ways for writers to enhance the comprehensibility of their texts. The students she works with are college juniors (19 y.). She found that these adult writers can benefit much from a method called 'Protocol–Aided Revision', which holds that a reader is given a text (composed by the mentioned writer), is asked to read it aloud and externalize all his/her thoughts during reading: every inference, comprehension difficulty, semantic link, or problem that he/she may find. The writer is given a typed protocol and/or an audiotape of this reader and learns to revise his text according to the signaled problems and other information. In this way, the proof–readers take the burden of error detection from the writer, by signaling comprehension problems on–line.

This type of observation is different from the ones discussed before, in that the observations are not directed at models for the writing process, but at collecting feedback on the writing performance. On the other hand, one could assert that writers should be able to take possible reader's responses into account during writing. This is a difficult part of the writing process: to look at text that you have just written through the eyes of a reader to whom the text is new. In this respect one could say that the observed readers' responses can function as a model for the reader whom the writer should imagine during the reviewing part of the writing process.

Schriver does not work with live confrontations or videotapes, but with audiotapes and typoscripts. Thus the word learning–by–observation has a different meaning regarding her research. The audiotapes give a real time account of the reading process, although without on–time visual information. Nevertheless, it is the essence of reader observation that the observer can get insight into the reader's comprehension activities. In this respect, Schriver's typoscript method has a similar aim as observing readers live or on video, and has proven to be effective. In an all the more interesting way, because her students were able to transfer or even translate their acquired knowledge about reading processes to the writing subprocess of error detection.

The Protocol–Aided–Revision method does not only yield textual benefits. In a training study (1992), Schriver experimented with two groups of writers (college juniors and seniors). Control students learned to revise by means of traditional audience–analysis theories and techniques (e.g. role playing, Lannon, 1985; purpose–oriented audience pedagogy, Emig, 1981), and collaborative peer–response methods in which peers give written or oral comments on each others texts after reading (e.g. peer critiquing, Bruffee, 1985). In short, these students learned to revise by doing error detection and revision exercises.

The experimental group learned to revise by means of Protocol–Aided Revision, using on–line reports of readers' comprehension protocols. As a posttest, the control and experimental subjects had to underline parts in a text (words, clauses etc.), which would likely cause comprehension problems for readers. The text problems were actually assessed by realistic readers, and it was determined to which extent each group made correct or false predictions regarding 'reader problems'. Schriver found that, after the training – which took ten lessons for all groups – the experimental group subjects were able to detect about twice as many probable reader's errors than the control group subjects (means: 5.25 errors for the experimentals, 2.46 errors for the controls). So the writers' experience with recorded comprehension processes of real readers resulted in a better 'nose' for the problematic aspects of texts.

The nature of the errors in which the experimental subjects excelled can be described by two categories: errors by 'omission' (a lack of information which would make a part of the text difficult to comprehend), or 'global level problems' (related to a problematic organization of text parts larger than the sentence). Moreover, the ratio between accurate and inaccurate predictions was 32/68 % for the experimentals, and 15/85% for the control subjects, which is clearly in favor of the first group.

Schriver's explanation is that the knowledge accumulated by the experimental subjects on readers' problems and needs can be transferred to new text editing tasks. It still has to be investigated if this increased 'reader awareness' will also transfer to new writing tasks.

It should be noted that Schriver, although her work on Protocol–Aided Revision is original in its use of on–line processes, stands firmly within the tradition of peer–response writing instruction. Research into peer response is abundant (see Rijlaarsdam 1987, 1993), although by far most of it concentrates on evaluating writing products, not on writing processes. However, the social aspect of peers evaluating products is in agreement with the more process–oriented instruction advocated by Schriver. In both cases, the writer has access to information regarding the quality of his text that he or she would not be able to access in another way.

3.6 Diverging perspectives for the observation of reading and writing

Both lines of research discussed in the previous section use observation and evaluation of reading or writing activities as a kind of 'vicarious' learning (although Schriver's research is much further from Bandura's observation of models). In both lines of research the effectivity of learning–by–observation is compared to more 'enactive' learning activities of subjects who do not observe reading or writing, but perform reading or writing tasks themselves. In both cases, observation activities appeared to have some advantages, specially when supported by a commenting task and/or by feedback.

There are also some important differences between Sonnenschein & Whitehurst's approach and Schriver's approach. These differences concern 1) the goal of the observations for the learner, and 2) the role of the learner in the communicative process. I will explain here why these differences are enough significant to speak of diverging perspectives on the observation of communicative processes. This divergence is reflected in the way that we organized the set of research questions for our study (section 3.7).

Concerning the first difference: in the Sonnenschein & Whitehurst experiments, it is the observer's task to evaluate the observed speaking or listening process *as if it is a model or example* for one's own task behavior. The observer focuses on the quality of the process and intends to acquire the knowledge or skill that enables performance of such processes. In Schriver's experiments, however, the observed activities are not example for the observer's task execution; the observation is aimed at acquiring authentic reader's responses as feedback, in order to evaluate and adapt one's own task execution process. This evaluation may have effect on the short term (the revision of the particular text) and an effect on the longer term (the activation of knowledge about authentic reading processes during writing, in particular during revision). In fact, Schriver's method of reader observation–as–feedback demands a special mental activity of the learner: transforming the observed reading activities in to consequences for one's own writing skill.

The second difference is related to the first. As I noted at the end of the previous section, the observers in the Sonnenschein & Whitehurst experiments are outside the observed communication (except condition 6 in table 3.2). They observe the productive part of the communication, or the receptive part, or both, but in none of these cases can the observer influence the communication. This is different in Schriver's experiments, in which the writers/observers initiate the communication, and both observation and evaluation are strongly related to their own performance.

In sum, there are some essential differences between the two types of learning–by–observation (observation of models, and observation as feedback). It is not certain which of the two will prove more effective in a direct comparison. Sonnenschein & Whitehurst's 'outsider' perspective or Schriver's 'insider' perspective. Therefore I will in this thesis investigate both instructional methods in contrast to learning–by–doing methods.

In combination with both written language modes, this yields the following possibilities (see table 3.3). Sonnenschein & Whitehurst have, apart from observations of writers *or* readers, also involved observation *both* writers *and* their readers. We will follow this idea.

Ad table 3.3:

The five perspectives in the two right–hand columns are variants of learning–by–observation, while the other two at the left side are variants on learning–by–doing. These perspectives will be further explained in the commentary on the research questions.

Table 3.3: *Perspectives on learning to read and write categorized according to perspective*

	Learning by doing *(traditional)*	Learning by observation as example *(Sonnenschein & Whitehurst)*	Learning by observation as feedback *(Schriver)*
Writing	Learning to write by doing writing exercises	Learning to write by observing writing processes as examples	Learning to write by observation of reading processes as feedback
Reading	Learning to read by doing reading exercises	Learning to read by observing reading processes as examples	Learning to read by observation of writing processes as feedback
Writing & Reading		Learning to read and write by observation of reading and writing processes as examples	

3.7 Research questions and scope of the study

On the basis of the theoretical considerations in chapters 2 and 3, and particularly the five learning–by–observation perspectives sketched in section 3.5, the main and secondary questions for the empirical part of this thesis can be formulated. The main questions concern the effectivity of the **instruction factor learning–by–observation**, of wich several variants will be compared with (variants of) **learning–by–doing**. Secondary questions concern the modes under investigation (this is a task factor) and the effect aimed at (either learning or transfer).

With 'learning effect' I mean the increase in a skill within the same mode as in which the subject has practised (e.g. the effect of writing practice on writing skill, or the effect of reading practice on reading skill) With 'transfer effect', the increase in skill is meant within the communicatively complementary mode (transfer of reading practice to writing skill, or vice versa).

The starting point of our research are the 5 experimental interventions concerning learning–by–observation. Apart from these, we used two control conditions concerning learning–by–doing: in one control condition the subjects performed reading or writing tasks with normal self–reflection (leaving the extent of self–reflection up to the subjects), in the other they performed these tasks with reinforced self–reflection (provoking self–reflection by one or more extra tasks). The reason for the extra control condition is that the control interventions in Schriver's and in Sonnenschein's studies were also enhanced versions of the simple learning–by–doing type discussed in sections 1.2 and 3.3. In Schriver's research, control subjects did much more than just writing; they took part in role–playing, in group critiquing, and they received feedback. The non–observing students in the Sonnenschein & Whitehurst studies were given direct feedback and explanation. I consider this as a means to support the learners' self–reflection. In order to make the comparisons similar to those made by Sonnenschein & Whitehurst and Schriver, a comparable 'reinforced self–reflection' condition is added to the experimental groups. It enables us to see how much of the 'learning possible' (see section 2.6) from normal to reinforced self–reflection can be won by each of the observation activities.

Although feedback enhanced the effectiveness of observation in the Sonnenschein & Whitehurst (1984) study, I chose to leave external feedback out of the experimental and control interventions. This thesis focuses on the effects of various learning activities which are initiated, performed, and evaluated by the learner – and only by him/her. External influences such as external feedback will probably modify these effects; I think the investigation of such, maybe drastic, modifications has no priority over investigation of the effects themselves. Moreover, adding external feedback would make the learning–by–doing conditions more dissimilar from reality: direct feedback on task performance, other than a mark and perhaps some product–oriented comments, is rare in writing and reading education. Thus, we made all interventions self–instructing so that subjects can work without being helped by research assistants.

All research questions will be answered in relation to the domains of reading and writing. Figure 3.7 gives a structural overview of the research questions central to this study.

[insert figure 3.7]

Fig 3.7: *General structure of research questions*

The main research questions of this thesis are formulated as follows:

Regarding learning effects:

Q 1: Is learning–by–observation of models more effective than learning–by–doing?

Q 2: Is learning–by–observation as feedback more effective than learning–by–doing?

Q 3: Is learning–by–observation of models of both modes more effective than learning–by–doing?.

Q 4: Is learning–by–observation of models of both modes more effective than observationof models of one model?

Q 5: Is learning–by–doing with reinforced self–reflection more effective than learning–by–doing with normal self–reflection?

Q 6: Do the abovementioned questions 1–5 yield different answers for learning to read and learning to write?

Regarding transfer effects:

Q 7: Does learning–by–observation of models promote more transfer than learning–by–doing?

Q 8: Does learning–by–observation as feedback promote more transfer than learning–by–doing?

Q9: Does learning–by–doing with reinforced self–reflection promote more transfer than learning–by–doing with normal self–reflection?

Q10: Do the abovementioned questions 7–10 yield different answers for learning to read and for learning to write?

(ad Q 3 and Q4: Transfer effects of the condition 'Observation (model) of both modes' cannot be distinguished from its learning effects. Therefore no question will be asked as to the relative size of this transfer.)

In short, the overall research hypothesis is based on the idea that learning activities involving observation can promote learning as well as transfer, compared to learning activities confined to individual practice. More specific hypotheses derived from this overal hypothesis will be put to the test in three experiments, which are reported in chapters 4, 5 and 6. The experiment reported in chapter 4 addresses only research questions Q2, Q5, Q8 an Q9 since the interventions are limited to learning–by–observation–as–feedback. The experiments in chapters 5 and 6 address all main research questions (see bottom of figure 3.7).

In every chapter, the general variable structure as given in figure 3.7 will be used as a starting point. The research questions, which are represented at the right side of this figure by comparisons of two conditions, are uniformly presented in the empirical chapters by their Q–number. Additional research questions in these chapters, which include specific variants of the above interventions, will be presented by extending the general variable structure, and by consecutive Q–numbers. In this way, the connections between the experiments may become clearer in spite of their differences. I will now explain some of these differences.

Learning and transfer as assessed in the first experiment is differt from the other two experiments. In chapter 4, the experimental intervention for writers consists of composing a text, observing readers trying to comprehend the text, and revising the text based on the observations. The 'amount of learning' is measured by the quality of the text written at the end of the intervention, i.e. the revised version, in comparison with the quality of first drafts. The 'transfer' effects measured are not intermodal effects, but effects on the knowledge that subjects acquire – or better: construct – about 'good instructive texts'. In a 'generalization task' the knowledge is assessed that students externalize in the form of advice to other students.

'Learning' and 'Transfer' in the other two experiments are more conventionally measured by posttests in which the students must apply their reading or writing skill. The type of transfer is intermodal: what is learned in reading tasks must be applied in writing tasks, and vice versa. Thus, the research reported in chapters 5 and 6 follow more the general outline sketched in this section.

In the last chapter, I will give a revision of the theory on the basis of results reported in the empirical chapters.

Although part of this thesis is based on previous studies, research into observation of task execution processes as a learning activity has not often been systematical or cumulative. I will indicate in which respects the present study adds new elements to the reported research. This study differs from Sonnenschein & Whitehurst and Schriver in the age of the participating students. Sonnenschein & Whitehurst's studies involved only young children in the age from 5 to 7. Schriver's studies were carried out with college juniors and seniors, so the average age will be around 20. We do not know to what extent the effectiveness of the instructional method learning–by–observation is influenced by the age of the observers. Secondary students (8th and 9th grade; about 14–15 y.) will participate in our experiments, since from the age of 15 students receive language skills instruction with a stronger accent on textual skills. Moreover, this is the age–group to which we would like to generalize our results, with a view to the development of curriculum materials.

Also the text types included in our study are comparably new. Sonnenschein & Whitehurst used 'referential communication', that is text containing information that directly refers to the outside world (objects, facts, or events), such as descriptions of rooms or realistic accounts of events. Schriver used educational texts, including lessons on how to operate a university computer system, and short expository science texts from popular magazines such as Time. We have chosen two different text types, namely instructive texts (manuals) for a physics experiment, and argumentative texts including complex rhetorical and argumentation structures. The instructive text type we use is related to Schriver's computer lessons, in that both texts instruct the reader to perform certain physical acts, and that both texts should explain to the reader the purpose of these acts. The more 'cerebral' argumentative text type, to which the reader's responses are more difficult to observe, has not yet been used in research into reader or writer observation. Because such texts are closer to educational reality, we thought it would be interesting to include them in the study.

A third novelty of this study is the way in which reading and writing activities are observed. As we have seen, the observed task performances in the Sonnenschein & Whitehurst studies were rather artificial (with dolls and pictures instead of real speakers and listeners). They do not represent authentic communication. Moreover, the speaking and listening tasks were extremely short. One trial usually consisted of one or two words: 'yellow triangle', 'blue circle', 'red', 'square'. In such situations, there is hardly any difference between observing products and observing processes or activities. Schriver used authentic material in her studies, and recorded much longer reading processes, which were transformed into typed protocols. In this study, we adopt a middle course regarding the length of the reading and writing tasks (between 2 and 10 minutes) and attempt to present it in a comparably natural way by means of videotape. The observer is able to observe practically every aspect of the authentic reading or writing process: the tone, the hesitations, the intonations as well as the actual writing or reading activities. Use of the videoplayer allowed the necessary experimental control, while the authenticity comes close to real–life observations.

Other differences between the present study and previous studies are, that two essentially different variants of learning–by–observation are investigated and to a certain extent compared. Furthermore, not only learning effects are assessed (like in Bandura's research on observational learning or in Schriver's research on protocol–aided revision) but also transfer effects. And finally, the instructional interventions are tested in two complementary language modes, so that the results on learning and transfer may yield information about the relationship between these modes. Are there structural differences between the acquisition of reading and writing, as promoted by learning–by–doing or learning–by–observation?

Figure 3.1: *Reading–Writing model (Shanahan & Lomax, 1986)*

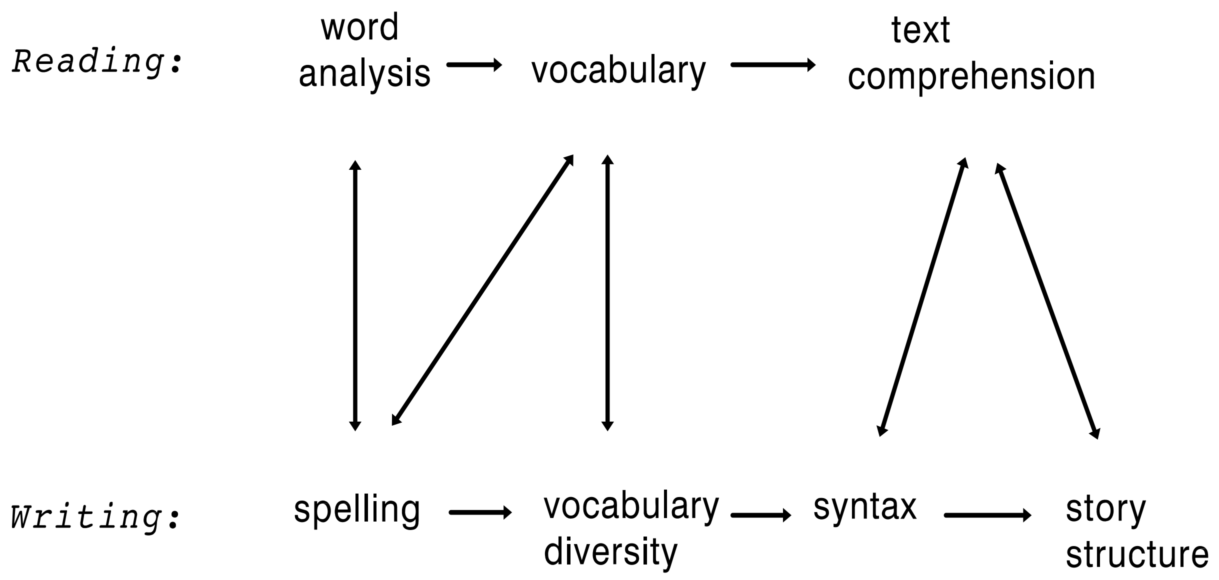


Figure 3.2: *Reading–Writing model (Shanahan & Lomax, 1986) including knowledge types:*

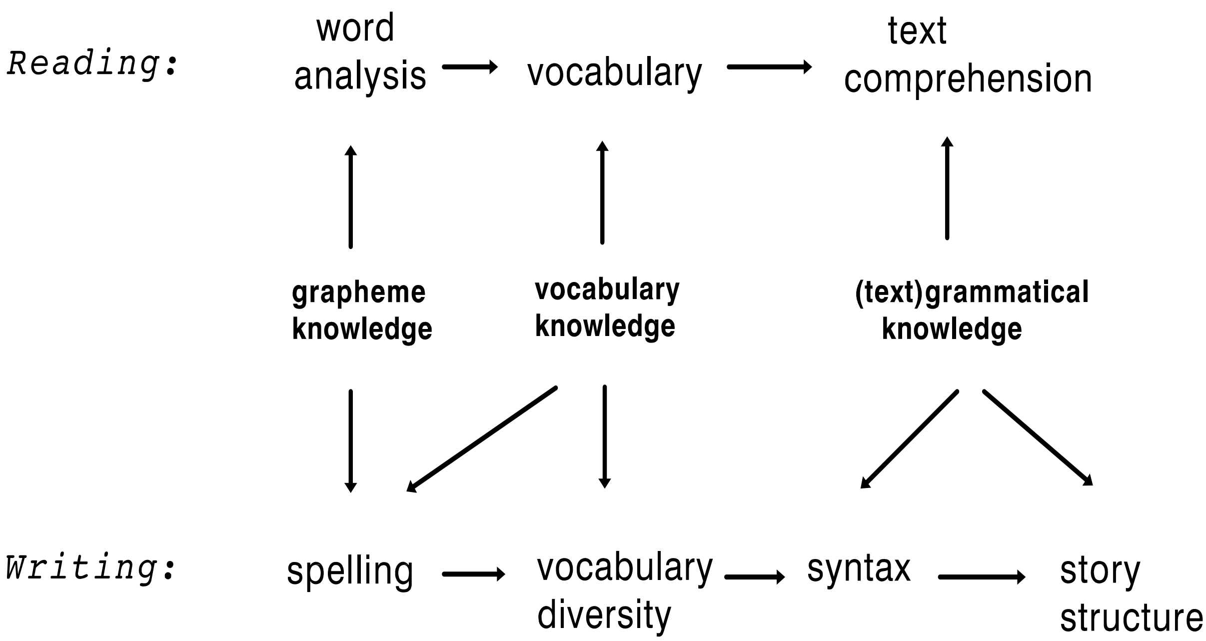


Figure 3.3: *Writing–reading model (Van der Pool & Van Wijk, 1995)*

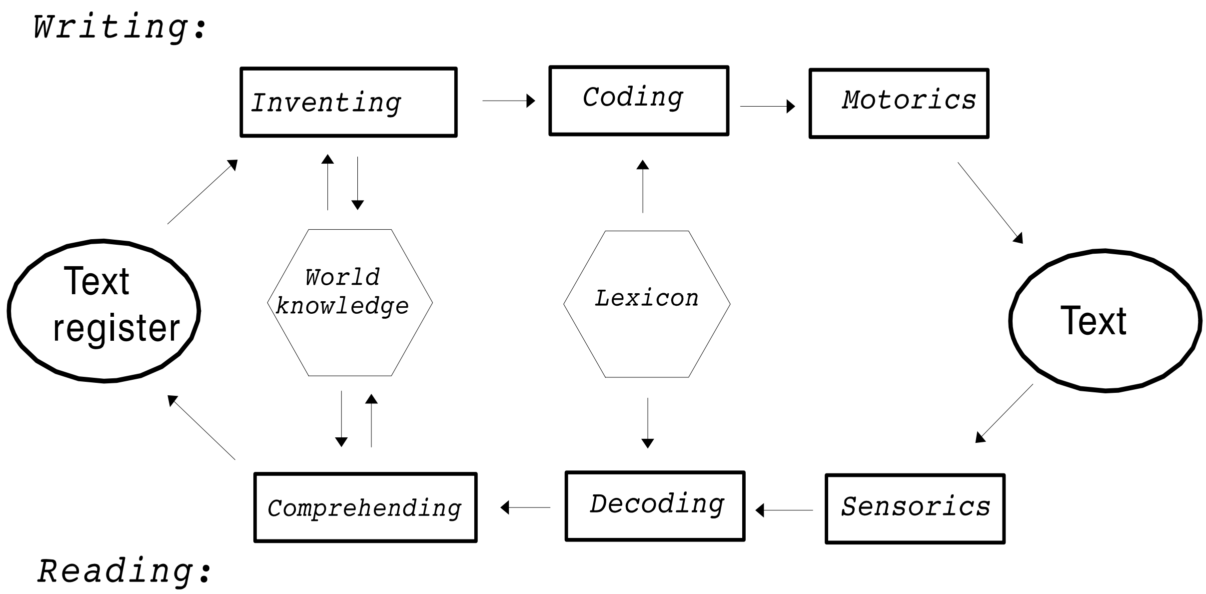


Figure 3.4: *A simple model for communication (Schultz von Thun, 1977)*

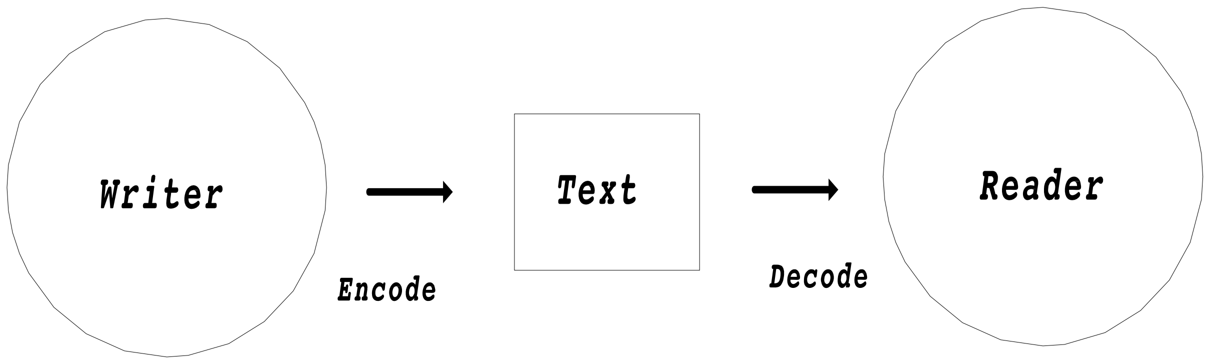


Figure 3.5: *Functional relationships between levels of performance activities*

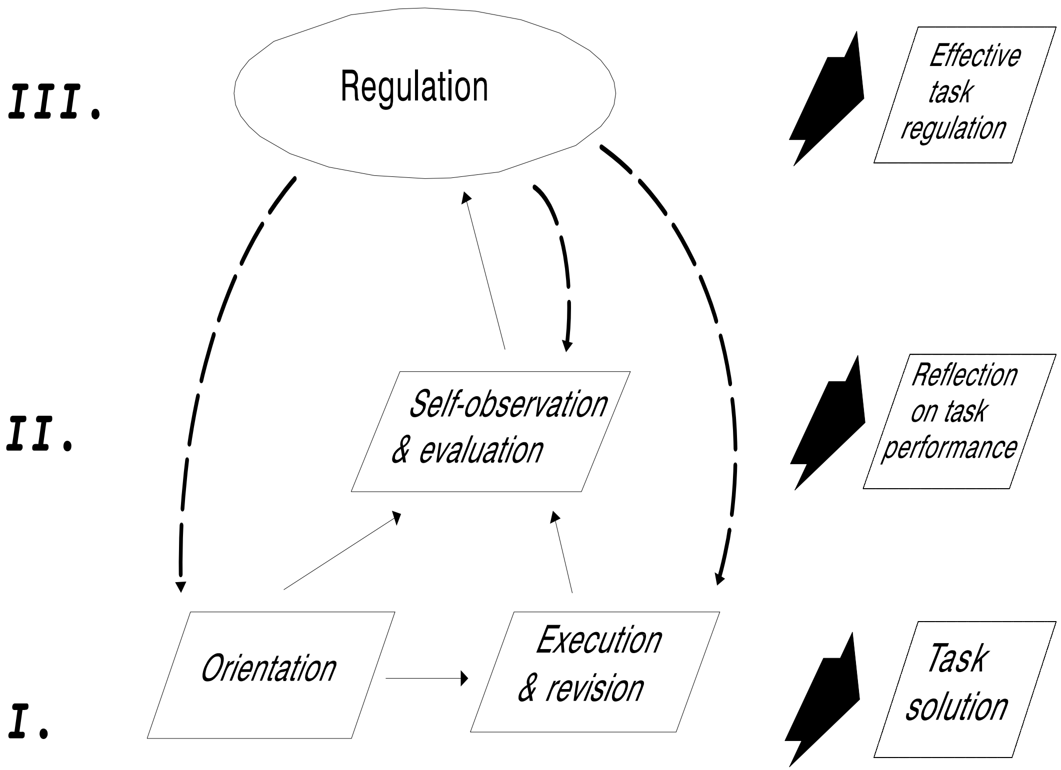
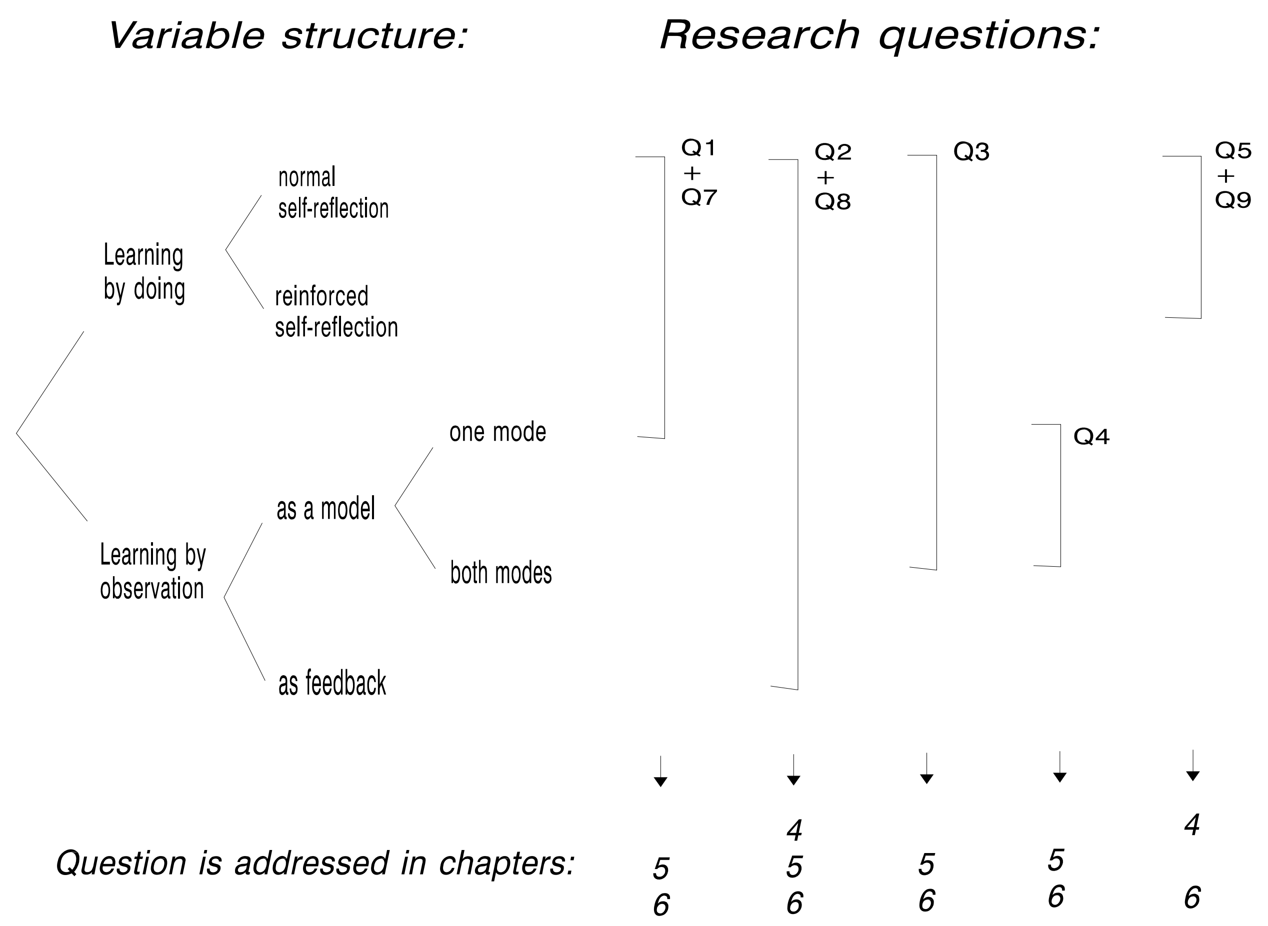


Fig 3.7: *General structure of research questions*



CHAPTER 4: Learning to write by reader observation and written feedback: Instructive texts

Two variants of the instructional method learning–by–doing and three variants of learning–by–observation–as–feedback were compared regarding their effectivity on learning to write instructive text. Generally, the observation methods outscored the learning–by–doing methods. Revising a text by means of self–generated comments did not improve text quality; neither did a orientating reading task preceding writing improve the quality of writing. Writers who evaluated their text by observing authentic readers' comprehension processes succeeded in improving the quality of their writing, an effect that was amplified by supplying readers' written comments on the text. Generalization of the learning experiences to a transfer task was in agreement with the learning effects.

Special attention is given to qualitative differences between the instructive texts (manuals) written before and after revision. The number of instructions remained the same, while the number of object descriptions, explanations, and precautions increased.

4.1 Introduction: Text revision as a learning activity

In the previous chapter, a theory was presented about learning activities for reading and writing, and their effectiveness regarding learning and intermodal transfer. I suggested that learning within each of these domains can be enhanced by stimulation of monitoring and evaluation activities aimed at writing and reading processes, instead of the executing activities so heavily relied on in a learning–by–doing approach to instruction. I further suggested that transfer of the learning results may be enhanced if domain–specific knowledge is acquired in a more abstract way, so that the learner is able to adapt it to other tasks or even other domains. Coding and decoding rules were identified as knowledge elements that may be susceptible to adaptation by a process of generalization or reversion, and are therefore candidates for transfer.

Learning–by–observation was identified as a learning activity that may meet these demands. Compared to learning–by–doing, one can assume that observation focuses the learner's attention more on the monitoring and evaluation of reading and writing activities than on their execution. Learning a communication rule by observing its use, instead of by merely applying it, may result in qualitative differences with respect to the context–specificity of the knowledge, opening the way to transfer.

The present experiment is concerned with learning to write. The instructional methods that are tested are variants of learning–by–observation–as–feedback, and some simple variants of learning–by–doing. The subjects are secondary students (9th grade, ca. 15 y.) of intermediate and high level. The text type to be written and evaluated is an instructive text, more particular a manual text for a simple physics experiment. Research questions Q2, Q5, Q8 and Q9 as presented in section 3.7 are addressed, along with additional questions concerning variants of the learning activities.

It is important to notice the special meaning of the terms *learning effect* and *transfer effect* in the reported experiment. With *learning effects*, I mean the increased writing performance, as measured by the increase of general text quality, after the experimental learning activities of observation and revision have taken place: to which extent can students improve their writing performance after each intervention? Furthermore, genuine *transfer effects* are not assessed in this experiment: we do not ask subjects to perform different writing tasks or tasks from different domains. Instead, in this experiment an important condition to the occurrence of transfer is assessed (cf. Simons (1990): do the learning activities result in new, retrievable and decontextualized cognitive elements regarding writing instructive texts? As a transfer task or 'generalization task', we asked the subjects to write down as much advice as they could for a class–mate who had to write a manual text for another experiment. In doing so, we want to assess if the subject had constructed new cognitive elements ('memory traces') about writing such texts, which can be activated on demand ('connectedness' and 'transfer strategy for memory search') and are to a certain extent abstracted from the learning context ('decontextualization'). The results of this task are considered an indication for possible *transfer* to similar writing tasks or to the reading mode. As said, the actual transfer is not assessed here.

I will first explain why reader–oriented revision may be a powerful activity in writing instruction. The intervention used in this experiment is based on Schriver's (1987; 1992) method of protocol–aided revision, in which writers observe authentic readers' activities in order to collect feedback which they can use in revision. Schriver (1987) reports on a study in which students evaluated the same text twice: first without extra information, and later by using a reading–aloud protocol from a person belonging to the intended audience. She found that confrontation with authentic readers' comprehension activities, problems and comments in the protocol resulted in many new problem detections within the same text, and, consequently, in a higher quality of the revised text. Moreover, writers who became more experienced in processing such realistic feedback developed sensitivity to potential readers' problems in text. Thus the profits of the intervention are twofold: an increased quality of the revised text, and a better knowledge of criteria for comprehensible text. In other words, reader's feedback contributed both to quality of writing and to the quality of the writer's knowledge.

These results were confirmed in Schriver (1992). Writers taught with the reader–protocol teaching method could improve the quality of their texts and learned to predict and diagnose comprehension problems in expository science texts, specially problems caused by textual omissions, and global–text problems. Moreover, writers' knowledge of audience acquired on one text type (instructional text) transferred to another type (expository science text).

Revision, so it seems, can be an effective learning activity for writers, but the textual and learning effects depend on the kind of information at the writer's disposal, and the way this information is processed by the writer. The contribution of revision activities to learning may lie in focusing the writer's attention to that part of his/her performance that is susceptible to improvement. The revision stadium has the educational function of receiving and processing feedback on performance. Revision activities may lead to changes in the performance, i.e. the text (the writing product) but also to changes in the knowledge and behavior of the writer (the writing process). For instance, if a writer finds that his introductory chapter is not in line with the content of the body of his text, he may improve the chapter, and also learn that it is better to delay writing introductory chapters to a later moment in the writing process. Such knowledge–from–experience is the basis for transfer to new writing tasks.

There are usually two sources of information for a writer to revise his text or his working method in writing: self–evaluations, either during or after writing, and evaluations made by external readers, e.g. teachers, in some way communicated to the writer (a note, a written or oral commentary). Each of these two sources has serious limitations.

Self–evaluations suffer from the fact that writers are not good proof–readers of their own texts, since it is hard to place themselves mentally in the position of their readers. When re–reading their own text, the writers have all the necessary prior knowledge about 'the writer's intention' to fill in gaps, solve ambiguities, activate contextual and situational knowledge, anticipate on content, and more. This will disqualify him/her as 'normal reader' and impede appropriate self–evaluative activities.

Evaluations made by external readers, as is usual in writing education, suffer from a distance in place and/or time separating the writing performance from the evaluation. This is a serious problem for writing instruction, since feedback is known to be effective mainly if it follows directly on task execution (Mayer, 1987, 102–112). If a writer has to wait days or even weeks before getting information about the adequacy of his/her task writing, it will be nearly impossible to link contrete writing activities to (evaluations of) their consequences. Moreover, external 'readers' supplying post–hoc feedback will function more like external evaluators who identify flaws rather than perform functional reading tasks. In this respect, the writer is not offered direct insight in the communicative consequences of his writing: the evaluators will necessarily offer a subjective selection, evaluation and verbalization of their reading experiences. This is a weak spot of indirect post–hoc responses.

Looking at the results obtained by Schriver, we infer that observations of on–line readers' responses can be superior to self–evaluation as input for revision and, as a consequence, for learning. In Schriver's study, the writer does not only use the evaluative output of external readers (their comments on text quality) but also information on the reading processes. The writer must connect the observed problems in the reading process to flaws in his text, to the applied writing method and to his knowledge about how to write such texts.

In the present study, the following variables are investigated (see left part of fig. 4.1). We will compare the effect of writing followed by self–evaluation (learning–by–doing) with writing followed by reader observation (learning–by–observation–as–feedback'). In order to evoke many and literally visible responses from the readers, we chose a typically appellative text type: instructive texts, i.e. texts conveying instructions to the reader to perform one or more actions. The writing student must compose a manual for a physics experiment. In the learning–by–doing conditions, the writing students must evaluate the completeness and comprehensibility of the text themselves. In the learning–by–observation conditions, a 'reader' (peer student) is asked to perform this experiment using the writer's manual, while reading aloud and giving oral comments on his or her comprehension. I will explain the variants of learning–by–doing and learning–by–observation that are distinguished in this experiment.

[insert fig. 4.1 about here]

Two variants of learning–by–doing are investigated for writing students, depending on whether self–reflection is stimulated or not. The first variant is a common writing procedure in which only a first version of a text is composed; this condition serves as a base–line (in the figure: condition 0). Subjects are not explicitly instructed how to write instructive texts; they must find out by themselves and imagine which instructions and other information they would need in order to do the physics experiment. 'Learning' is rather limited in this condition: it is a kind of learning by (one–time) experience. Such activity does, however, reflect traditional writing education in which little or no instruction is given before writing, and in which there is only one trial for learning to write a type of texts (Hillocks, 1986). 'Self–reflection' is not stimulated in this condition, but the extent to which subjects evaluate or reflect on their performance is left up to them ('normal' self–reflection, which may strong for some subjects, and absent for others).

More advanced learning–by–doing is performend by subjects who write a first version, followed by *guided revision* (condition I). By prompting the learner to re–read, generate comments, and revise the text, we can assess the extent to which the writer is capable at all of detecting problems in his text. This guided revision activity is supposed to get the most out of students self–evaluation and self–revision capacity, and will serve as a 'reinforced self–reflection' condition (see section 3.7).

Three variants of learning by reader observation will be investigated. In the simplest variant, a *writer observes a reader* who is actually using the text to do the experiment (condition II). While doing so, the reader is thinking aloud and commenting on the text's comprehensibility, thus supplying on–line information about this comprehensibility. Different from Schriver's protocol–teaching method is that the writer is more directly confronted with verbal and visual information from the reader instead of by studying written protocols.

In a second variant, the reader observations are supplemented by *written evaluative comments* from the reader (condition III). Although we have said that post–hoc evaluative comments will be less profitable for the writer than on–line observations, the comments may have a surplus value if added to the observations. They may assist the writer in the abstraction of knowledge from his specific observations and consolidation of this knowledge. For instance, several observations of reader's problems caused by omission can be summarized as 'you have to check for missing information'. For the writer, it is easier to store and access such general or 'condensed' knowledge about flaws in his or her writing than knowledge about each singular observed problem (Schriver, 1992). We expect that more learning will occur if the writer receives a summary of the problems experienced by the reader(s).

The third variant is equal to the first, with the exception that the observed text is not written by the observer (condition IV). It is interesting to know to which extent potential benefits of the reader observation are sanctioned by the observer's authorship. Would the same benefits occur if the writer observed readers using a text written by someone else? It is, however, not likely that texts written by others would lead to the same results. First, the observed feedback would not be tuned to the individual flaws in the observer's performance, and it would be difficult to translate the observed feedback to observer's writing. Second, a motivational counterargument is that the observer would be less personally involved because it is not his own performance that is at stake – he is not part of the communication between the observed reader and an unknown writer.

Subjects from conditions I–IV are generally referred to as 'revisors', since revision is their specific learning activity of which positive influence on performance is expected (in comparison to condition 0). In the context of their initial writing task, however, we will call these subjects simply as 'writers'.

Although much research has been carried out to determine effects of *receiving* feedback on writing, the opposite is true for learners *giving* feedback (Hillocks 1986: 166–168, 219–221). Only in studies concerning peer response groups or collaborative writing, the act of giving critical comments as feedback is part of the learning activities. In these cases, however, the resulting learning effect has rarely been isolated from the effect of receiving feedback, since the students are usually part of write–and–comment rounds (e.g. Rijlaarsdam 1986; Rijlaarsdam, Eiting & Schoonen, 1989).

In our study, we expect beneficial effects for the subjects supplying feedback by using an instructive text, while reading aloud and commenting. These tasks can have positive influence on their writing skill: the reading activity may be considered a strong orientation on a consecutive writing task. By experiencing one text as a reader, the student may learn about the criteria for comprehension that apply to this type of texts. The knowledge about these criteria, acquired during reading and commenting, may be used in their own writing. Sonnenschein and Whitehurst (1984) showed that learning to comment on communication can in fact transfer to performance of the commented activities. Thus, we have some reason to expect potential benefits for the reading and commenting subjects. In our experiment, we will consider their activities (writing preceded by reading & commenting) as another case of learning–by–doing with reinforced self–reflection (condition V).

The orientating quality of the reading task may be specially strong for subjects who reflect on the detected deficiencies and then write down their most important comments (condition VI). In doing so, they arrive at a stadium of 'condensation' of their experiences. Like the revisors who receive written comments as a supplement to their observations, the orienters may enhance or consolidate their learning by generalizing across their reading experiences. The present experiment will be set up in such a way that this effect can be observed in isolation.

I will generally refer to subjects of condition V and VI as 'orienters' (and in the context of their reading task: 'readers'), since they start the intervention by reading, using and commenting on an instructive text. However, their next activity is to write such a manual text themselves, using their reading experiences, and it is this writing performance that is assessed. It is important to keep in mind that subjects referred to as 'orienters' or 'readers' are, in fact, also occupied with learning to write instructive texts like the other subjects.

The resulting learning effect will be assessed as the quality of the students' final writing performance. For the revisors (writers), this is the quality of the text written after revision; for the orienters (who are 'readers' first, but perform the writing task afterwards) this is the quality of the text they write after performing the reading and commenting task. The final versions are compared to the quality of the first versions, written by condition 0 subjects at the start of the instruction. The improvement in quality can be regarded as an indicator for learning.

The transfer of these learning effects to new tasks depend on the extent to which the subjects can generalize across their particular experience with this task, by constructing cognition that is useful for future writing. In this experiment, this generalization is measured as 'transfer task'. It is attempted to assess the students' declarative knowledge about criteria for manual texts by asking them to write down as many pieces of advice as possible for a classmate who also has to write such a text (De Glopper, 1986; Braet & Schoonen 1994; Schoonen & De Glopper, 1995). Differences in numbers of good pieces of advice given indicate difference in the extent to which the various groups are able to construct consciously accessible knowledge.

In this experiment, transfer is not measured by having the students write another manual and assessing its quality. The reason is that a gain in declarative or procedural knowledge does not automatically lead to a gain in performance on transfer tasks. The construction of declarative knowledge is a first stadium of acquiring skilfulness in a new domain (Singley & Anderson 1989; see also section 2.4 of this study). Since students differ in the readiness with which they apply new cognitions in new tasks, we considered it appropriate to first assess the precondition to better writing: better knowledge about how to write, knowledge about the criteria for writing effective manual texts, before assessing the actual application of this knowledge in a new task. If the students can report their knowledge about these criteria, it can be assumed that the first stadium of learning has taken place.

4.2 Research questions and expectations

The structural relations between the *independent variables* were represented in figure 4.1. On the right side of this figure, research questions are connected to the variables and conditions. For most comparisons between conditions, two research questions are given (e.g. Q2 + Q8). The former research question concerns the differential *learning* effect of the compared conditions, and the latter question the differential *transfer* effect.

The main research questions as presented in chapter 3 are aimed at comparing the effectiveness of learning–by–doing and two types of learning–by–observation. The present experiment is only concerned with observation as feedback. Moreover, only reading processes are observed, so we cannot compare the effectiveness of the instructional methods in different modes (Q6, Q10). Thus, in this experiment only main research questions 2, 5, 8, and 9 are addressed, along with several additional questions.

Dependent variables: learning and transfer effects.

There are two *dependent variables*. First, the *general text quality* of the final version written by the subjects after revision (conditions I–IV) or after the reading/commenting task (conditions V–VI) is assessed. This general text quality is indicative for the quality of the final performance of the subject, after performing the experimental learning activities (as explained in section 4.1). Research questions about the 'effectivity' of certain types of instruction refer to this dependent variable. The 'general text quality' of the written manuals is measured by scoring the content elements. The scoring method is explained at the start of the 'results' section (section 4.4.1).

Second, the acquired *declarative knowledge about writing instructive texts*, as measured by the number and quality of pieces of advice given by the subjects three weeks after the intervention, may indicate possible transfer of a certain condition. More precise, it indicates certain conditions to the occurrence of transfer (Simons 1990). This declarative knowledge is both consciously retrievable and abstracted from the learning context, and may therefore be applied to other tasks. Research questions about 'promoting transfer' are aimed at this dependent variable. The scoring method for the pieces of advice given is also explained in the 'results' section of this chapter (section 4.4.2).

Research questions regarding learning effects.

I will first list the main research questions regarding *learning effects*, and describe the operationalization of the independent variables (the levels of which are underlined). Learning–by–doing is operationalized as performing the physics experiment and writing an instructive text that can serve as a manual; for condition 1 this is followed by generating self–evaluations, and revising the text; learning–by–observation–as–feedback is operationalized as writing an instructive text and revising it after observing readers using the text.

Q 2: Is learning–by–observation as feedback more effective than learning by doing?
→ (Expectation: yes, observation enhances monitoring and evaluation, and thus learning)

There are two types of learning–by–doing for revisors: normal self–reflection is operationalized as writing an instructive text without guided revision; reinforced self–reflection is operationalized as writing followed by a guided self–revision session (reflection after the writing task).

Q 5: Is learning–by–doing with reinforced self–reflection by a consecutive revision task more effective than learning–by–doing with normal self–reflection?
→ (Expectation: yes, reinforced self–reflection promotes learning)

Another type of learning–by–doing concerns the orienters: their reinforced self–reflection is operationalized as writing an instructive text, preceded by extensive orientation by reading and commenting on such a text (reflection before the writing task). This condition is also compared to the normal self–reflection condition described above, which has no prior self–reflection task.

Q15: Is learning–by–doing with reinforced self–reflection by a prior orientation task more effective than learning–by–doing with normal self–reflection?
→ (Expectation: yes, since the reading experience may offer a better orientation on the consecutive writing task)

Several research questions are added to this experiment. As explained in the previous section, it will be investigated whether the authorship of the evaluated text influences the effectiveness of the feedback observation. Observing readers as feedback on someone else's writing performance may be less effective, because the feedback may be less appropriate and the personal involvement of the observer may be lower. Consequently, we must distinguish between two conditions of 'observation as feedback', namely observation as feedback on one's own writing performance (operationalization:

observing readers using a text written by the observer), and feedback on someone else's writing performance (observation of readers using a text not written by the observer). The comparison of these two conditions will answer the following research question:

Q11: Is learning–by–observation as feedback on one's own performance more effective than learning–by–observation as feedback on someone else's performance?

→ (Expectation: yes, if feedback is more appropriate and personal involvement in the communication is stronger, learning effects will be larger.)

Another additional research question concerns the effect of giving or receiving written comments, as a means to enhancing the construction or consolidation of knowledge about one's task performance. Summarizing their comments may help the orienters in constructing a meaningful orientation base for a consecutive writing task. Receiving and processing readers' comments may help the revisors in learning how to revise the particular text, and in constructing more general criteria on how to write such texts. If we compare these orienters and revisors with similar groups who do not write or receive summarized comments, we can answer the following questions:

Q13: Is learning–by–observation as feedback (on one's own performance) more effective with receiving summarized comments than without receiving summarized comments?

→ (Expectation: yes, since the summarized information may supply more information about the writer's performance than the writer derived from the observations.)

Q17: Is learning–by–doing (with reinforced self–reflection) with writing summarizing comments on the orientation task more effective than without writing summarizing comments?

→ (Expectation: yes, since the act of summarizing may yield a better orientation on the consecutive writing task)

Research questions regarding transfer effects.

Next, I will list the research questions regarding *transfer effects*. They concern the same comparisons between conditions as the questions regarding *learning effects*, but the comparison is aimed at another dependent variable. In figure 4.1 it is indicated that Q8 corresponds with Q2, Q9 with Q5, Q12 with Q11, Q14 with Q13 and Q18 with Q17.

Q 8: Does learning–by–observation as feedback promote more transfer than learning by doing?

→ (Expectations: yes, observation promotes learning that is more abstracted from the context in which it is acquired, and thus it promotes transfer)

Q 9: Does learning–by–doing with reinforced self–reflection promote more transfer than learning–by–doing with normal self–reflection?

→ (Expectation: yes, reinforced self–reflection will promote abstraction and thus transfer)

Q12: Does learning–by–observation as feedback on one's own performance promote more transfer than learning–by–observation as feedback on someone else's performance?

→ (Expectation: undetermined, because pro–argumentation (learning effects may be larger, and transfer depends on learning) and counter–argumentation (personal distance enhances abstraction) keep each other in balance.)

Q14: Does learning–by–observation as feedback (on one's own performance) with summarized comments promote more transfer than without summarized comments?

→ (Expectation: yes, since summary information will enhance access and consolidation of knowledge)

Q16: Does learning–by–doing with reinforced self–reflection by a prior orientation task promote more transfer than learning–by–doing with normal self–reflection?

→ (Expectation: yes, since the translation of reading experiences into the writing task is a process of abstraction; the acquired knowledge may be more susceptible to transfer)

Q18: Does learning–by–doing (with reinforced self–reflection) with writing summarized comments on the self–reflection task promote more transfer than without summarized comments?

→ (Expectation: yes, since the more consolidated orientation resulting from writing comments enhances the transfer effect expected in Q16).

4.3 Method

4.3.1 Design

An experiment was set up in order to test the hypotheses. Table 1.4 displays the research design:

Table 4.1: *Design*

cond + description:		n:	1st session: **WRITING**		2nd session (3 weeks later): **EVALUATING AND REVISION**				3rd session (3 w. later) **GENERALIZING**
			performing experiment	writing 1st version of manual	observing reader	receiving written comments	generating self–comment	revising	describing criteria for writing instructive texts
0	Learning–by–doing normal self–reflection	14	x	x					x
I	Learning–by–doing reinforced self–reflection	14	x	x			x	x	x
II	Learning–by–observation feedback own text no written comments	20	x	x	x			x	x
III	Learning–by–observation feedback own text plus written comments	7	x	x	x	x		x	x
IV	Learning–by–observation feedback other text	20	x	x	x			x	x

cond + description:		n:	1st session: **READING & WRITING**			2nd session (3 weeks later): **GENERALIZING**
			reading the text while performing the experiment	producing written comments	writing manual	describing criteria for writing instructive texts
V	Learning–by–doing reinforced self–reflection without written comments	14	x (group III texts)		x	x
VI	Learning–by–doing reinforced self–reflection with written comments	20	x (group II texts)	x	x	x

It it shown that five types of experimental activities of writing students are compared and two types of activities for reading students. Students from condition 0 serve as a base line or control group, writing only 'first versions' of manuals and showing their resulting declarative knowledge about manual texts. All other conditions are experimental conditions to be compared with condition 0, on both dependent variables *general text quality* and *declarative knowledge*. Learning activities of groups 0 and I are called 'learning–by–doing' since they write and revise autonomously. Learning in groups II, III and IV is called 'learning–by–observation' because they obtain their information for text revision from observing authentic readers. Groups V and VI learning activities are called 'learning–by–doing' because they practice their writing skill by writing one text; the preceding reading task is supposed to offer a strong orientation on the consecutive writing task.

It is important to note that the manuals written by experimental groups II and III were used and commented on by their peers from the reader groups V and VI. That is, froup VI functioned as 'authentic readers' for group III, and group V as 'authentic readers' for group II. The texts from groups II and III were randomly assigned to the 'readers' from groups V and VI.

Subjects from conditions 0–IV worked in three sessions: they wrote a manual in the first session after having performed the experiment; they revised their text in the second session, on the basis of observations or self–evaluative comments; and in the third session they generalized across their experiences by giving advice on how to write such texts. The experimental reading students took part in two sessions: they used and commented on a manual text in the first, and wrote a new text in the first session, and they performed the generalization task in their second session. A time span of three weeks was between each two consecutive sessions.

The manuals and letters of advice were scored on content elements. The scoring procedure is explained in detail in the results section of this chapter.

4.3.2 Subjects

In all, 109 students from the 9th grade (intermediate and high level) of two schools took part in the experiment. The average age was 15 years. The data for groups 0, I, III, and VI were collected first. A year later, when we wanted to address additional research questions, we collected and added the data for groups II, IV and V. We can check for the comparability of the writing groups, including groups II and IV, because all students had written a first version during the first experimental session. From a comparison of the mean and standard deviations of the general text quality of each group, we conclude that the groups can be considered equal (see Results section).

The students from groups 0, I, III, and VI were randomly selected from six student groups (three groups from each participating school) and randomly assigned to the experimental conditions, after removing from each student group the five lowest achievers and the five highest achievers on writing skill (as indicated by school grades for a recently written essay). The reason for this homogenization is that the study focuses on *typical* student behavior, not on differences between more or less skilled students. It is assumed that the within–group variance on writing skill decreases as a result.

The same homogenization and randomization procedure was followed one year later when the data for groups II, IV and V were collected. We have no direct possibility to check for the comparability of both reader groups (V and VI). However, there is some indirect evidence. Randomization should ensure sufficient comparability between groups II, IV and V with respect to abilities in writing and physics, just as groups 0, I, III and VI will be comparable for this reason (see Results section). Since no difference in general text quality of the first manual version was found between groups 0–IV, we can extrapolate their level of ability to the reading groups, which we consider as comparable.

The group sizes differ considerably. This is caused by our initial plan to have each manual read and commented on by two readers (from the 'orienters' conditions) and to have the writer observe both readers. This explains the 1:2 ratio of the groups III and VI, and the comparably small group size of group III. Once it was found that each reader took eight to twelve minutes to complete the reading task, we decided that observing both readers would take up too much session time, and that the revisors would observe one reader only. The reason why observation of two readers would take too much time is that the time–on–task for groups I–IV should be equal, in order to avoid alternative explanations of experimental effects.

4.3.3 Procedures

All writing and reading assignments were given individually. Each student worked privately in a room with a research assistant giving standardized information and answering questions. The assignments were given orally, and were read from paper.

I will describe the procedures in the same order as they were planned and executed. That is, session 1 for conditions 0–IV, session 1 for conditions V–VI, session 2 for conditions I–IV, and at last the 'generalizing session' for all conditions I–VI.

The initial writing task

The writing task for conditions 0–IV was the composition of a manual for a simple physics experiment (Inspectierapport PABO, 1988). Each writer was told that this manual should serve two goals: it should enable a classmate to *do* the experiment without problems as well as to *understand* it. This formulation of goals should stimulate the writer to take several needs for information into account (see 1.1): it should contain instructions as well as explanations.

Before handing over the writing assignment, each writer received explanations from the research assistant about the physics experiment and did the experiment him/herself until it was understood and performed faultless. Thus it is assumed that between–group differences in comprehension of the physics experiment are minimized. This is a necessary precondition to an equal start of each experimental group. We could check for possible between–group differences in the comprehension of the experiment by comparing the quality of the manuals written in the first session. These differences were not found (see Results section).

In order to fully understand the writers' task, it is necessary to know the physics experiment in detail. The aim of this experiment is to prove that *air takes up space*. Three pictures were shown to the writer in order to explain the experiment. The accompanying text was read aloud by the research assistant. The student did not receive written explanations, because it should be avoided that he/she used them as a 'model' for writing.

"This little experiment will show you that *air takes up space*. This means that boxes or bottles which are commonly called *empty*, in fact contain air – and as long as the air is inside, nothing else can go in. If you fill an empty box with books, the air has to go out at the same time.
So, at first sight, air may seem equal to 'nothing'. But in some cases the existence of air is something that should be seriously taken into account.

[insert figure 4.2 here]

"To see for yourself that air takes up space, we will do the following experiment. In picture 1 you see a construction that you can make with some objects on the table (one bottle, one cork and one funnel). You have to choose the objects well so that the construction fits exactly, i.e. so that no air can escape."

[insert figure 4.3 here]

"If water is poured into the funnel, this will *not* fall into the bottle. The reason for this is that the air is still inside and cannot escape. You may say that the air 'obstructs' the water."

[insert figure 4.4 here]

"Picture 3 shows how this problem is solved. A straw that is stuck through the funnel creates a passage for the air, just like a chimney. The escaping air will free up space within the bottle, which is immediately occupied by the water. Therefore you will see that the water starts flowing into the bottle. If you put your finger on the top of the straw, the air cannot escape any longer and the water will stop flowing. As soon as you lift your finger, air and water will continue to flow: the air out of the bottle and the water into it."

After these explanations, the student had to perform the physics experiment. Several problems had to be overcome in performing. First, the choice of the necessary objects. At the student's disposal were the following objects: *3 bottles* (one too small, 150 ml; one too big, 400 ml; and one exactly right, 250 ml), *4 corks* (one too small with a hole; one too big with a hole; one with the good size but without a hole; one good size with a hole), *3 funnels* (one small but fitting; one size too big; two sizes too big) and *3 straws* (thin and red; too big and red; too big and blue). The student spent a substantial time finding out the only one fitting combination, i.e. the medium size bottle, the medium size cork with a hole, the smallest funnel and the thin red straw. All other combinations do not fit. In their manuals, the students had to refer exactly to the objects they used, or the experiment was doomed to fail.

A second problem consists of the air–tightness of the construction. This could only be obtained by pushing the objects really firmly together. As long as the construction was not hermetic, the experiment was guaranteed to fail. As soon as water was poured into the funnel, the construction would show if it was hermetic: if not, the water would run immediately into the bottle and not stand in the funnel.

A third problem arose when the straw had to be stuck into the bottle through the water in the funnel. This would fill part of the straw with water, which had to be removed (by carefully blowing or sucking the straw) before the air could pass through.

The student had to overcome these and some more problems to bring the experiment to a successful conclusion. Only when necessary, e.g. if the student got stuck and could not resolve the problem him/herself, the assistant gave advice.

Directly after the experiment, which took about ten minutes, the writing assignment was given. The instruction was as follows:

> "Suppose that in the afternoon one of your fellow students has to do the same experiment. He or she enters the room, and all these objects (all bottles, corks etc.) are on the table. No one is there to give instructions or explanations (like I gave to you) and there are no pictures (like I have shown you).
> What this student needs is a good *manual* for this experiment, in which it is exactly stated what he/she should do and know in order to *perform* the experiment without problems and to *understand* what it is about. Now it is your task to write such a manual on the basis of your experience with the experiment.
> You may only use words. You are not allowed to draw any pictures.
> Be aware that your text is meant to be used by one person who has to do the same experiment as you have done, and should understand what he or she is doing. You should try to write a manual that will ensure a quick and unproblematic execution of the experiment and clear understanding. Use as many details as you consider useful. Again: do not draw any pictures."

The writing task was usually completed within twelve to fifteen minutes. After finishing, every student was prompted to read the text aloud and correct any failure that he or she detected.

Reading, commenting & writing procedures for conditions V and VI

The manuals written by experimental groups II and III were used by peers, namely by the 'orienters': group VI functioned as authentic readers for group III, group V as authentic readers for group II. No explanation was given by the research assistant about the content of the experiment and no pictures were shown. The students or 'readers' were only instructed to perform 'a physics experiment' as described in the manual, using the objects on the table before them. They had to think aloud during the experiment and were prompted to continue talking whenever they fell silent. The 'readers' were also instructed to give on–line comments on the problems they experienced, related to the quality of the manual. The reader's performance of the experiment and the oral comments were recorded on videotape. The following instruction was given:

> "You will need some of the objects on the table in the following physics experiment. To execute the experiment, you will have to use the manual. In this manual, someone has tried to describe for you what you should do in order to perform and understand the experiment.
> Your task consists of two activities:
> 1) First you should *precisely* follow the instructions in the manual. Try to do exactly what is written: no more, no less.
> Do not start working according to your intuition, follow the instructions literally.
> 2) Whenever the manual is not clear to you, say it *immediately.* Also make mention of missing or incorrect information. Think aloud continuously while doing the experiment, so I can follow your line of thinking."

In this way, defects in the manuals could become visible and audible in the execution of the experiment. The readers ('orienters') would run into many problems due to missing, incorrect and unclear information, leading to observable confusion and an imperfect task execution. The video recordings serve as 'reader protocols' in which speech and images of the reading process are combined.

After having completed the experiment, or after ten minutes if they did not succeed to complete the experiment within this time span, the 'orienters' were informed about the procedure and goal of the physics experiment, to compensate for differences in quality of the manuals they had used. The same text was used as the revisors had heard when the experiment was explained to them in the first session.

For condition VI, the reading task was followed by an assignment to write general comments about the quality of the manual. The instruction was:

> 'One of your fellow students has tried to write this manual as clearly as possible, so that it would cover all necessary information for you to do the experiment well and to understand it. Now I ask you to write down your comments on this text. Do you think it is a good manual? What are its weak points: what was unclear or incorrect, what was missing? And what are its strong points, which helped you to do the experiment well? Look it over thoroughly and give as many comments as possible."

Writing a short summary of comments took two to four minutes.

For both reading groups, the second session ended with a writing task. After having used and commented on an imperfect manual, the students had to write their own 'first version'. It was not allowed to use the manual they had seen as a model during writing.

> "Now write your own manual for another student who has to do the same experiment. Think of your experiences with the manual and your comments on it."

The average time for writing their own version varied from nine to thirteen minutes (average eleven). Since variability was low and no linear relation between time–on–task and text quality is expected, time–on–task was not included in the analysis of text quality data. When finished, the students were prompted to read their text aloud and correct any mistakes.

Evaluation and revision procedures for conditions I, II, III and IV

This experiment is partly focused on the various observation, evaluation and revision activities that are part of the second session for conditions I to IV. The nature of the information on which the revisors base their revision differs between the groups.

In condition I, the writer is his own proofreader. Three weeks after having written a first version of the manual, the writer receives it again and is asked to redo the experiment. The same collection of objects is at his disposal. The instruction focuses the student's attention on revising the manual text:

> "Many people discover problems in their texts when they read them after some time. While you are using your own manual, you may get useful ideas for improving it. You will probably find that it contains some good points, but it may also be susceptible of improvement.
> So your last task will be: think of as many points as possible that may cause problems for another student, who has to execute and understand the experiment without problems. Revise your manual such that all shortcomings are corrected."

First, the students were asked to read their text aloud as orientation. Secondly, they had to use the text in redoing the experiment. While doing the experiment, and after finishing it, they were asked to generate as many comments as possible on the quality of the text. These comments were written down by the student. The last task was to write a new manual, on the basis of their self–evaluations. They could use their first version as a model.

Students in condition II and III did not have the opportunity to redo the experiment themselves. Instead of acting as their own imaginary reader, they observed authentic readers. After re–reading the first version aloud as orientation, they looked at a video–recording of a reading student using their own manual. Each writer observed another reader, since only one recording could be made of each participating reader. As a consequence, the amount and the nature of the feedback that revisors received or that readers gave varied between subjects. This is not a severe problem, since this experiment focuses on between–group differences.

The video–observation lasted for eight to twelve minutes and contained non–edited recordings of a student using and directly commenting on the manual texts. The instruction was as follows:

> "A fellow student tried to do the experiment using the manual that you wrote. You will see on video what has come of it. What should you look for? While looking at the video, you may get useful ideas for improving your text. You will probably find that it contains some good points, but it may also be susceptible of improvement. Try to concentrate on information that will help you in improving your manual. Does the student you see understand your intentions? Does he/she perform the experiment well? What are his/her comments on your text?
> Take notes while looking at the video. Whenever you want, the tape can be stopped so you can write or ask something."

After having seen the video–recordings, only the students of experimental group III received and read written comments from two readers: one that had been observed on video, and another reader that had not been observed. The comments supplemented the notes they had taken themselves during observation.

Directly after the observation, the students were instructed to rewrite their first version:

> "You have seen one or two students who used your manual. You may have got ideas for the improvement of your text. So your last task will be: revise your manual such that all shortcomings are corrected. Make good use of the information that you received by looking at the video–recordings."

The revision of the manual took fourteen minutes on the average. The students had to write a new text, on the basis of their observations and notes (and for group III: the written comments). They also could use their first version as a model.

The observation and revision procedure for revisors in condition IV was similar, with the difference that the manual used was not written by the observer, but by an unknown writer. As a consequence, the observer was confronted with communicative failures of another text and had to relate them to possible revision of his own manual:

> "Some fellow students tried to do the experiment using a manual that was written by one of your classmates. You will see on video what has come of it.

> What should you look for? While looking at the video, you may get useful ideas for improving your own text. You will probably find that the manual that is used contains some good points, but it may also be susceptible of improvement. Try to concentrate on information that will help you in improving your own text. Does the student you see understand the intentions of the writer? Does he/she perform the experiment well? What are his/her comments on the manual?
> Take notes while looking at the video. Whenever you want, the tape can be stopped so you can write or ask something."

These students also had to revise their text after having seen the video–recordings. They did not receive any written comments from readers.

Generalization task for conditions I-VI

Four weeks after the revision session, all students were given a writing assignment which is supposed to tap their declarative knowledge about manual texts. All students were asked to write a letter to a fictitious fellow student, who needs some advice on how to write a manual text for a physics experiment. It is assumed that the number and quality of the pieces of advice given are an indicator of the student's knowledge about criteria for writing good manual texts. The following instruction was given by the research assistant:

> "Imagine the following situation: one of your classmates comes to ask you for advice. He or she has to write a manual for a physics experiment, and has heard that you have got some experience in writing such texts.
> Although you do not know what sort of experiment your classmate has to write about, still you think that you can explain what points he/she should pay attention to when writing a good manual text. The manual should be used by students of your age.
> Write a friendly note to your classmate and give **as much advice as possible**, clearly stated, that would help him/her writing a high quality text."

The students were given twenty minutes for writing this advice letter. It was explicitly stated by the research assistant that a manual text should not be written, but described.

4.4 Results

First the results will be reported regarding the dependent variable 'general text quality', as indicator for the quality of the final performance after the experimental learning activities. Further, the results on the generalization task are reported, aimed at the knowledge of criteria for 'writing good manuals' as measured by the letter of advice.

4.4.1 Effects on learning-to-write

I will start by explaining the instrumentation for the scoring method. After reporting on general text quality, addressing the research questions regarding learning, I will go more into detail and adress the research questions again with respect to specific text qualities.

Instrumentation and scoring

For the assessment of 'general text quality' a standard manual was constructed that could function as a scoring model. This model consists of an introduction, of three episodes called 'construction', 'water', and 'straw', and a conclusion. For each of these parts, a list of standard information elements is stipulated. This standard manual was constructed by an analysis of the oral explanations given to the writers during instruction, and also inductively on the basis of some written manuals.

Communicative utterances in a 'good manual'

Introduction
1. (T) The experiment is about ...
2. (I) Notice the initial initial situation

Construction–episode
3. (O) Which bottle should be used?
4. (O) Which cork?
5. (I) Put the cork on the bottle.
6. (O) Which funnel?
7. (I) Put the funnel into the cork's hole.
8. (P) Push everything tightly together.
9. (T) because no air may go out.

Water–episode
10. (I) Put water into the funnel.
11. (I) Check if the water stands still.
12. (T) that is because the air in the bottle stops it.
13. (P) If it does not stop, press everything tightly.
14. (T) that is because air is leaking somewhere.

Straw–episode
15. (O) Which straw?
16. (I) Put the straw into the funnel's hole.
17. (P) Make sure the straw's one end is pushed into the bottle.
18. (P) Make sure the straw's lower end is not put into the water.
19. (T) The straw is needed to remove air from the bottle.
20. (I) Check if the water is now running in the bottle.
21. (T) that is because air can now go out.
22. (P) If water does not run: blow, suck or move the straw;
23. (T) that is because water in the straw must be removed.
24. (I) Hold your fingertip on top of the straw.
25. (I) Check if the water stops flowing,
26. (T) that is because air cannot escape any longer.
27. (I) Lift your finger again to let the water go on flowing.

Conclusion
28. (T) This experiment proves that water occupies space,
29. (T) because the water cannot go into the bottle as long as the air cannot go out.

Each of the 29 elements in the four episodes meets a possible need for information that a reader may have. I will first explain these needs. It is supposed that the quality of a manual depends on the number of 'needs for information' that it can satisfy.

At the beginning of the text, readers want to know what they can expect. Therefore, in the 'introduction' information should be given related to the *subject* or *aim* of the experiment, and the reader should be prompted to *orientate* himself on the initial situation with the many objects on the table before him.

With these objects a construction must be made during the 'construction–episode'. Unambiguous descriptions of the necessary objects are of great value to the reader, who would otherwise get lost in construction problems (not knowing what the intended construction looks like). The reader should be warned here to make a firm assembly; if not the construction was guaranteed to leak.

A first observation is done in the 'water episode', when water does - counter–intuitively - not run into the bottle. An ignorant reader would not be alarmed by the water running in the bottle due to a leak in the construction, so a preventive warning is very useful here. Also, the reader want to understand this unusual phenomenon, so an explanation is suitable. A hint must be given on how to correct an undesirable situation.

In the 'straw episode' the relation is observed between escaping air and flowing water. Readers will run into several problems in this episode, which can all be overcome by correct descriptions, precautions and explanations.

Finally, a concluding part should supply the reader with information on what inference to make from the observations, because 15 year old readers cannot be expected to make such inference themselves.

Each text was scored on the occurrence of each of the 29 information elements. That is, for each element it was checked whether the writer had realized it in the text. If so, the element was scored with a full point. If not, the score was zero. If in serious doubt, the element scored a half point.

Scoring was done by independent and trained scorers. One quarter of the texts was scored by three scorers. Because their inter–scorer–reliability (Cohen's kappa) turned out to be very high (0.86 to 0.88), it was decided that the other three quarters would be done by two scorers only. Their inter–scorer–reliability was 0.86.

Two examples of a comparably weak and a comparably good manual will clarify the scoring method:

Manual for the physics–experiment

> You take a bottle. You put a cork in there. You put the funnel in the cork's hole. You put water in the cork and then you see that no water comes in the bottle and that's why you put the straw in and then the air can go away and the water in.

(7 points: 5, 7, 10, 11, 16, 19, 20)

Readers using this text would most probably run into problems with choosing the correct materials. It is very hard for readers to find the correct combination if they are completely unaware of what construction should be made. The semi–causal links 'and then you see that...' and 'and that's why you...' are not really informative. Probably a reader would not understand what the experiment is about, because the aim is not explicitly mentioned.

Manual for the physics–experiment

- You see 3 bottles and 3 corks on the table.
- Take the middle bottle and the cork without marks or spots, with an opening.
- Put the smallest funnel in the cork's opening.
- Push the funnel really well (no air may escape).
- Pour water into the funnel.
- You'll see that the water in the funnel stands still (if you pushed everything really well together, so no air can escape).
- In order to let the water run through, it must be made possible for air to escape.
- Put a straw (the thin one) through the funnel until halfway the bottle.
- You may not hold the straw in the funnel itself, there's water in there which stops the air from flowing out).
- If the water still does not run through, there must be water inside the straw, so the air cannot go through. (Suck if necessary.)
- The water runs through now.
- Put your finger on the top of the straw, then the water will stop, because no air can escape.

(21 points: 2–4, 6–11, 13–17, 19–20, 22–26)

This text differs in length as well as quality from the former: it contains more detailed descriptions, explanations and hints for readers.

It was possible to interpret almost every textual element in the manuals written by the students as a realization of one of the 29 elements in the standard manual. For an element to be counted positive, it was necessary that the jury recognized the performance of the essential communicative act (Searle 1979), be it in a direct or indirect way. For instance, an utterance was counted as an instruction if the intention to instruct could be derived from the text. A sentence 'Pour the water into the funnel' clearly contains an instruction in the imperative 'Put'. Also incomplete sentences could be interpreted as instructions, like 'Now water in the funnel'. However, 'Funnel with water', was not counted as instruction because it refers to an existing state–of–affairs rather than that it prompts to change the present state–of–affairs. Similarly, the object descriptions were scored positive if the description referred *unambiguously* to one of the objects on the table, because unambiguity is the essence of referential communication (Sonnenschein and Whitehurst, 1984). 'Take a bottle' is therefore scored as inadequate (score: zero) while 'take the medium size bottle' is considered adequate because only one object answers this description (score: one). Many readers' problems were caused by inadequate object descriptions, because readers had to search a long time for the only combination of objects which would allow a succesful experiment.

General text quality (learning effect)

The research questions can be answered by the differential effects of the independent variables on the general text quality of the final writing performance (text B). If the quality of text B is higher than of text A, it can be assumed that learning has taken place. A one–way ANOVA of the sum scores of all elements, followed by a post–hoc comparison yields these differences in general text quality. The results are displayed in table 4.2:

Table 4.2: *Mean scores, standard deviations and post–hoc comparisons for general text quality*

condition:		mean text A: ((max. 29)	s.d. text A:	mean text B: (max. 29)	s.d. text B:	quality of text B significantly higher than condition: 0	I	II	III	IV	V	VI
0	Learning by doing normal self–reflection	9.79	2.98	– –	– –							
I	Learning by doing reinforced self–reflection	9.69	2.71	10.11	3.43							
II	Learning by observation feedback on own text no written comments	10.11	3.11	14.31	1.67	*	*					
III	Learning by observation feedback on own text plus written comments	10.05	2.23	17.43	3.14	*	*	*		*	*	*
IV	Learning by observation feedback on other text	9.58	2.85	12.85	3.11	*						
V	Learning by doing inforced self–reflection no written comments	– –	– –	11.97	3.16							
VI	Learning by doing inforced self–reflection plus written comments	– –	– –	13.54	2.55	*	*					

*) Duncan: * means p < 0.05.

Duncan method for post–hoc comparisons was chosen because of its correction for differences in group size (Winer, 1971).

The reliability of the scale used for scoring, made up of all 29 elements, is 0.65 (Cronbach's alpha). If low–contributing items 1, 9 and 19 of the standard manual are removed, reliability increases to 0.71, which can be considered sufficiently high for determining between–group differences. Leaving these items out has no consequences for the reported significance of between–group differences, and thus no consequences for answering the research questions.

In column 'mean text A', the mean general quality can be found of the first versions written by the students in groups I–IV (these versions are called 'text A'). There are no significant differences between these means, so we may conclude that the groups as comparable with respect to their initial writing ability.

We now come to answering the research questions. This asks for a comparison of group means for text B. In answer to research question Q2, we see that the three learning–to–write–by–observation groups II, III and IV reached a higher score on text B than learning–by–doing groups 0 en I. That is, higher than the group 0 score on the only version group 0 wrote (which can be taken as a base line) and higher than the quality of group I's revised text. The final performance of all groups of revisors who observed readers as feedback (condition II–IV) is higher than the final performance of the writers who wrote one version only (condition 0), or of the revisors who generated comments on their own text and revised the text accordingly (condition I).

Q5 asks for the effect of 'reinforced reflection' on learning–by–doing. For writers, this is indicated by the difference between groups 0 and I. There is no difference between these groups of writers, so we must conclude that generating self–evaluative comments for the revision of a text is not an effective way to improve writing performance regarding this kind of texts. It is therefore doubtful if much learning has taken place for students in group I.

For orienters, Q15 is answered by looking at the difference between groups 0 on the one hand, and V and VI on the other hand. All groups wrote only one text, but subjects from groups V and VI were stronger oriented on the writing task because they had just read and commented on another writer's text. For group V subjects (who did not write comments) this appears not to have supported their own writing significantly; however, the orientation has worked for group VI. (See the answer to Q17.)

The difference between groups II and IV is that students in group IV observed feedback on a text that they did not write, while group II students saw a reader using their own texts. There is no difference in the quality of the texts that these groups write after observation and revision, so we conclude in answer to Q11 that the authorship of the text does not influence the effect of learning–by–

observation.

The effect of receiving summarized readers' comments (Q13) can be determined by comparing groups II and III. Both groups observed 'their own reader', but only group III subjects received written comments from two readers (the observed reader and another one) after observation. These written comments appear to be very helpful for revisors, because the writing quality after revision increased greatly. Group III outscores each of the other learning–by–observation groups. It must, however, be noted that this condition is not a 100 % observation group since the nature of the information at their disposal during revision is a combination of observations and written post–hoc commentary.

Q17 (concerning the difference between giving written comments or not in addition to the reading task) is answered by comparing groups V and VI. All subjects in these groups wrote their first version after having read and commented on another person's text. However, group VI subjects may have consolidated orientation by writing a summary of their comments. The effect is that only group VI subjects improved the quality of their text, in comparison with a plain writing task (group 0). The quality of group V texts does not surpass this level. This is an, albeit indirect, indication for a difference in effect of group V and group VI learning activities.

Specific text quality

How should we account for the increase of text quality by reader observation? To get more insight in the nature of this increase, we assigned all 29 elements of the standard manual to four categories, on the basis of the speech act (communicative act) performed in each element (Searle 1979). Four types of speech acts were used to categorize the 29 elements: Instructions, Object descriptions, Theoretical explanations, and Precautions. In the standard manual at the beginning of this chapter they are marked I, O, T or P.

Instructions are the heart of any manual text. They describe (or better: commission) the actions to be undertaken. They should be ordered chronologically, so the reader can follow the manual step–by–step.

Object descriptions supply the reader with the necessary information about the materials to be used. Many constructions or experiments will only succeed if the correct materials are used, like with this physics experiment. Without knowledge of the correct materials, the reader is likely to waste much time and energy. The writer was specifically instructed to write a manual that would enable an unproblematic task execution by the reader.

Theoretical explanations support readers in understanding what they are doing. It is not useful just 'going through the motions' when performing a physics experiment. The writers were instructed to enable the reader to understand the experiment, and had also received sufficient explanation to understand the aim of the experiment and to give meaning to their observations, so they should be aware of the necessity of such information.

Precautions make up a very useful category of hints within manual texts. When performing an experiment, some things are likely to go wrong. The writers experienced many of these misfortunes, such as air leaking from the bottle, water entering the straw etc. In good manuals, such misfortunes are anticipated on, warned for, and remedied.

Table 4.3 gives the differences in text elements within each of four speech act categories of all manuals. Categories are transformed into 10–point–scales to enable direct quantitative comparisons between categories, although the maximum number of elements within each category differs in reality. Asterisks indicate a significant post–hoc comparison (Duncan, $p<0.05$) when compared to the base line, condition 0.

Table 4.3:

Mean scores, standard deviations and post–hoc comparisons for four types of speech acts in text B. Maximum for every type = 10.

	Instructions		Object descriptions		Theory & explanations		Precautions	
condition:	mean I	sd I	mean O	sd O	mean T	sd T	mean P	sd P
0	6.43	1.24	2.23	2.60	1.64	1.66	1.64	2.46
I	6.50	1.63	3.21	3.76	1.89	1.29	0.86	1.35
II	7.18	1.15	4.86	4.18	3.18 *	1.26	3.00 *	2.30
III	7.14	1.21	8.57 *	2.33	4.29 *	1.38	5.14 *	2.61
IV	7.60	1.38	3.25	4.06	2.77	1.32	2.35	2.62
V	7.17	1.50	3.18	3.92	2.35	1.66	1.15	2.10
VI	6.68	1.59	7.14 *	3.72	2.93 *	1.59	2.14	1.83

There is considerable variability among the learners, between groups as well as within groups. This can account for the fact that certain between–group differences, even though comparably large, are not statistically significant. The large variability may be explained by the fact that no direct instruction was given on the writing task, so that each student had to 'invent' the content of learning by him/herself. This would imply that the students vary largely in their capacity to 'invent' content.

Scale relability is high for Objects (0.89), intermediate for Instructions (0.57) and and Precautions (0.59) and low for Theory (0.32). The Spearman–Brown formula for test extension shows an acceptable reliability of three of the four scales (>0.70). However, the scores on the 'theory' scale will have to be interpreted with very strong reservations.

The conditions do not differ in the number of Instructions in their manuals. As concerns the four 'writer' groups, the increase in general text quality is mainly due to an increase of object descriptions (group III), theoretical explanations (groups II and III) and precautions (groups II and III). Subjects of 'orienters' group VI have improved their texts by inserting better object descriptions and theory (group VI).

The essential speech acts in instructional texts are of course instructions. The writing proficiency of the 15–year old students is at such a level that they have no problem in writing these essential parts. Most of the instructions in the standard manual can be found in the students' texts, and this category is not susceptible to improvement by any of the instructional methods. Students do not find possible improvement in including more instructions, but in including information that supports the reader in following and understand these instructions: information about the tools to be used, about the meaning of the experiment and of their observations, and about possible pitfalls and how to cope with them. Only revisors who were confronted with authentic readers' needs for such information have improved the quality of their text according to these needs. Besides, orienters who experienced the need for such information themselves, also improved their texts on those aspects, but *only* if they had consolidated their experiences by writing summarized comments (group VI).

4.4.2 Effects on the generalization task (transfer effect)

Reader observation seems to offer a useful basis for learning to write an instructional text. But is the beneficial effect limited to the revision of the particular text, or have observers learned anything from their observations that may transfer to similar writing tasks? By having all participating students write a 'letter of advice' about manual texts (see 3.3.4), we tried to tap potential effects on knowledge about criteria for good and weak manuals. Between–group differences in the quantity or quality of the pieces of advice can be attributed to the type of feedback that each group processed or supplied.

First I will explain the scoring of the given advice. Next, the results for the experimental groups are presented.

Instrumentation and scoring

The students could expose their declarative knowledge about criteria for good manual texts in a 'letter of advice'. The request to give as many pieces of advice as possible and the open formulation of the task served to evoke a non–selective memory search and instant writing down on paper.

In all, 108 letters of advice were collected. The number of pieces of advice in each letter was assessed, as well as their nature. This was done by means of a scoring model containing a categorization of possible pieces of advice. In the construction of this form, it was assumed that the pieces of advice were either *process–oriented* or *product–oriented*; and if they were product–oriented, that they concerned either (text–independent) *style* or (text–dependent) *content*. Each of these classes was further subdivided, up to a system of quite specific and recognizable categories. The scoring form was partly constructed before collecting the data; on the basis of the collected advices, several subcategories were added (B.1.3; B.1.9; B.1.11 and B.2.5).

CATEGORY:	EXAMPLE:
A. Process–oriented advice (advice on how to organize the writing process)	
A.1 Orientation	*'start with doing the experiment yourself'*
	'first examine all the object on the table'
A.2 Text production	*'while you write, repeat the experiment in your mind'*
	'write down all you can think of'
A.3 Revision	*'re–read the text when you have finisched'*
	'at last, check if your little sister would understand the text'
B. Product–oriented advice (advice on desirable properties of the text)	
B.1 Style–oriented, regarding:	
B.1.1 – Clarity	*'what you write must be very clear'*
B.1.2 – Length	*'keep the text as short as possible'*
B.1.3 – Accuracy	*'do not forget the small details'*
B.1.4 – Completeness	*'make sure you mention all objects'*
B.1.5 – Correctness	*'beware of mistakes'*
B.1.6 – Organization	*'give the instructions step by step'*
B.1.7 – Accent	*'pay special attention to the theory'*
B.1.8 – Audience–oriented	*'someone of your age must understand it'*
B.1.9 – Goal–directed	*'do not exaggerate the details'*
B.1.10 – Spelling	*'check for spelling errors'*
B.1.11 – Other	*'make a drawing if you want'*
(These general advice can be applied to many text types)	
B.2 Content–oriented, regarding:	
B.2.1 – Instructions	*'everything the reader must do'*
B.2.2 – Objects	*'which tools you should use'*
B.2.3 – Theory	*'what the whole thing is meant for'*
B.2.4 – Precautions	*'if things goes wrong, what should you do'*
B.2.5 – Other	*'tell them to clean up the mess'*
(These categories are more specific for one text type: manuals)	

Two scorers categorized the pieces of advice according to this scoring form. The interrater–reliability on the primary level (A.1 to A.3, B,1, B.2 and B.2.1 to B.2.5) was 0.82 (Cohen's kappa). On sublevel B.1.1 to B.1.11 (subcategories which are more difficult to tell apart) the reliability was 0.70. This reliability is high enough for making between–group comparisons.

To illustrate the use of this score form, two letters of advice and their scores are presented:

Comparably weak letter:

	Dear someone,
	I am happy that you asked for my advice.
	If you want to write a good manual, you should mind these points:
B.1.3	// – Write neatly and precisely, or the person who has to do
B.1.8	the experiment will not be able to read it.
B.1.5	// – Do not write too much nonsense.
(B.1.3)	// – Write everything in detail.
A.1	// – First write down all important things you can think of.
	I hope this helps. If you still have a problem, do call me.

This letter contains some very general product–advice regarding style, which are applicable to many text types. There are two pieces of advice from category B.1.3. If the second advice seems to be just a paraphrase of the first, like in this case, they are counted as one piece of advice. The last piece of advice is process–oriented: a proposal to brainstorm before starting to write.

Comparably good letter:

score:

	Hello, here you find my promised letter with pieces of advice for your manual for the physics experiment:
A.1	// It is very handy if you start by thinking really well
B.2.2	about // the things that you need and that you do not need // and
B.2.1	about what has to be done // (If you do not make such a plan, it is better to not write at all.)
B.1.6	// Next: keep the order of the activities in mind // and also the moments that some tool must be used.
A.2	// Then it is time to start writing:
B.1.7	// Put an accent on the most important things, such as:
(B.2.1 B.2.2)	// How you must do it! // And: which object you must use.
B.1.8	// You must not forget that children your age must be able to understand your manual.
A.3	// So when you are finished, you check it yourself and
(A.3)	correct some mistakes. // Check if you would be able to do the experiment with your own manual (faultless!)
B.2.4	// Only if you are certain that you've done all to save your classmate from problems, you can hand over the manual to the teacher.

This letter–of–advice is longer and shows more variation in pieces of advice. The writer seems to have followed the course of the writing process: pieces of advice concerning orientation, writing and revision are present in their natural order (A.1, A.2 and A.3). The letter contains three content–oriented pieces of advice concerning the use of instructions, objects and precautions, albeit not very precise. Also three style–oriented pieces of advice are given.

All pieces of advice from all advice letters were categorized by using the score form. Some phrases contained more than one advice, such as:

> 'Describe chronologically all the things that need to be done.'

This advice would get a positive score on B.2.1 (instructions) and B.1.6 (organization).

If some experimental groups have acquired more knowledge about manuals than the others, this should result in a higher score on the category 'product–oriented pieces of advice – content' (B.2.1 – B.2.4). In order to increase the sensitivity of the assessment, the pieces of advice assigned to this category were rated on quality. In this way, justice can be done to differences such as:

> 'Write down what to do with the tools'

> 'Describe very accurately which of the available objects must be used, so that the reader does not have to search; also tell him which precise acts must be perform with the objects'.

Two raters rated each content–advice as '½' (content–advice given, but in a very general wording, or implicit) or '1' (content–advice given, precisely formulated, with motive or example). Interrater–reliability was 0.80 (Cohen's kappa).

Knowledge about criteria for good manuals

In order to answer the remaining research questions, an ANOVA was performed on the advice–letter data. The ANOVA was followed by pair–wise post–hoc comparisons (Duncan, since group size differs considerably). The results are presented in table 4.4:

Table 4.4: Mean scores, standard deviations and post–hoc comparisons for total number of pieces of advice, and for number of pieces of advice of each subcategory.

ALL PIECES OF ADVICE

		mean # of advice	s.d.	different from conditions: 0	I	II	III	IV	V	VI
0	(learning–by–doing + normal self–refl.)	4.93	1.73							
I	(learning–by–doing + reinf. selfreflection)	5.14	2.11							
II	(learning–by–observation, own text)	5.47	2.22							
III	(learning–by–observation, with comments)	10.86	2.73	*	*	*		*	*	*
IV	(learning–by–observation, other text)	5.50	2.33							
V	(learning–by–doing, without comments)	4.45	2.11							
VI	(learning by doing, with comments	7.50	1.70	*	*	*		*	*	

ADVICE ON PROCESS

		mean number of advice	s.d.	different from conditions: 0	I	II	III	IV	V	VI
0	(learning–by–doing + normal self–refl.)	0.57	0.85							
I	(learning–by–doing + reinf. selfreflection)	0.71	1.07							
II	(learning–by–observation, own text)	1.15	1.30					*	*	*
III	(learning–by–observation, with comments)	1.71	1.38	*	*			*	*	*
IV	(learning–by–observation, other text)	0.50	0.76							
V	(learning–by–doing, without comments)	0.30	0.57							
VI	(learning by doing, with comments	0.29	0.47							

ADVICE ON STYLE

		# advice	s.d.	different from conditions: 0	I	II	III	IV	V	VI
0	(learning–by–doing + normal self–refl.)	1.43	1.16							
I	(learning–by–doing + reinf. selfreflection)	1.50	1.16							
II	(learning–by–observation, own text)	1.79	1.47							
III	(learning–by–observation, with comments)	3.43	1.71	*	*	*		*	*	
IV	(learning–by–observation, other text)	2.00	0.79							
V	(learning–by–doing, without comments)	1.25	1.16							
VI	(learning by doing, with comments	2.71	1.06	*	*	*			*	

ADVICE ON CONTENT

		# advice	s.d.	different from conditions: 0	I	II	III	IV	V	VI
0	(learning–by–doing + normal self–refl.)	2.93	1.77							
I	(learning–by–doing + reinf. selfreflection)	2.92	1.64							
II	(learning–by–observation, own text)	2.53	2.17							
III	(learning–by–observation, with comments)	5.71	1.50	*	*	*		*	*	
IV	(learning–by–observation, other text)	3.00	2.49							
V	(learning–by–doing, without comments)	2.90	1.68							
VI	(learning by doing, with comments	4.50	2.21			*			*	

Subscale reliabilities are as follows: Process (3 items): α = 0.61; Style (6 Items): α=0.44; Content (4 items): α=0.5n explain the low reliability of the scale 'style' from the fact that the items making up this scale are not related to each other. Advice on (non–textspecific) style can be given independently; there is no cohesive mental 'set' of style advice, such as the three items making up the 'Process' scale (orientation, writing, evaluation) or the 'Content' scale (Instructions, Objects, Theory, Precautions) which can generally be found in any manual. The intermediate alpha of the 'Content' scale is partly due to its small number of items (four), and partly to the influence of the 'instructions' item, which is scored in 90 % of all advice letters, both by good and by weak writers.

We can now answer the remaining research questions concerning 'transfer effects', or better: the generalization of declarative knowledge acquired in the intervention. Research question Q8 (see section 4.2) refers to the difference in transfer–effectivity of the learning–by–observation and the learning–by–doing revisor groups. This question cannot be answered unambiguously, because condition II and IV did not have any advantage over learning–by–doing, while condition III showed a great advantage. Apparently the effectivity of these learning activities are modified by other variables. These other variables are addressed in the more specific questions Q12 and Q14.

Q9 asks for the effect of 'reinforced reflection' on the transferability of the acquired knowledge during learning–by–doing. 'Revisor' group 1 show no increase over group 0. This was not as expected, because can now be accounted for when we see that the quality of their final writing performance had been equal as well (Q5). For orienters, the effect of 'reinforced reflection' is assessed by comparing group V and VI with group 0. We see that the reading task that preceded writing for groups V and VI does not result in a higher number of pieces of advice (Q16).

Hardly any effect on knowledge was found in the comparison of groups II and IV (research question Q12). It makes no difference whether the observed reader uses a text written by the observer, or written by an unknown person. There are no effects regarding the total number of pieces of advice given. Regarding the subcategories, only 'process' advice shows a significant difference in favour of group II, but this may be an artefact since the scores in this category indicate bottom–effects, and it is not easy to theoretically account for the difference between groups II and IV. (Although maybe a stronger personal involvement of group II subjects in the observation led them to remember how important made them remember how important orientation and revision activities are).

On the other hand, strong effects were found resulting from the addition of written comments to both the revisor's task and the orienter's task (Q14 and Q18). When comparing group III to II, we find a considerable advantage of group III regarding the total number of advice given. This result confirms our idea that summarized comments may 'consolidate' the writer's knowledge acquired during writing.

It is noteworthy that learning–by–observation in itself does not contribute much to the acquisition of declarative knowledge about 'writing good manual texts' (results of groups II and IV), but that strong positive effects of this instructional method can be found by adding the some written comments. This effect is mirrored in the comparison of groups V and VI. The difference is that group VI subjects wrote summarized information after the reading task. Accordingly, they give many more pieces of advice, specially with regard to style and content. The progress with respect to content–oriented pieces of advice is most significant, because this category represents cognitions that are specific for the type of text that was written.

In general, students from conditions III and VI wrote the most pieces of advice. This can be concluded with respect to the total number of pieces of advice (which is 50–100% higher than in the other groups) and each of the subcategories of pieces of advice. We do not know whether the feedback and revision activities have allowed the students to acquire *more* knowledge about instructional texts, or that the activities result in knowledge that is *more readily retrievable* from memory (which is in itself a beneficial result of learning too).

4.5. Summary and discussion

Three variants of a learning–by–doing method of instruction (including base–line condition 0: four variants) and three variants of learning–by–observation–as–feedback were experimentally compared in a study on writing instructive texts. Expectations were based on a theory of observational learning, adapted to communication (see chapter 3). This theory predicts effective learning and transfer resulting from observation of task execution processes. In applying this theory to the domains of reading and writing, we distinguished *learning by observation of models* from *learning by observation as feedback* on one's own task performance. The present experiment is about observation as feedback. By observation of comprehension processes of authentic readers', revisors may derive knowledge about reader's needs and potential problems, which they can use in the revision of an existing text, or which they may anticipate when performing new writing tasks. These theoretical ideas have been investigated by Schriver (1987; 1991; 1992), who found that writers, by studying authentic readers' processes, could enhance their texts as well as their knowledge about typical reader's problems with certain text types. This experiment adds some new variants to the observation method of instruction, and is aimed at secondary students learning to write by revision. Moreover, this study is concerned with the learning effects for the 'orienters', who perform a reading and commenting task before writing themselves. Lastly, revisors and orienters are tested on their declarative knowledge regarding instructive texts, or the extent to which they are able to generalize across their writing experiences (transfer effect), by formulating what they have learned as advice to other writers.

I will briefly summarize and discuss the answers to the research questions.

EFFECTS ON LEARNING TO WRITE (learning effects)

The question whether learning–by–observation–as–feedback is more effective than learning–by–doing (Q2) can be answered positively. All writer groups who observed readers before revising were able to improve their text significantly, while the learning–by–doing writers of group I failed to establish such improvement. Apparently, these writers have not been able to detect the flaws in their texts. An explanation for this uneffective self–revision is that the writers have a too–well understanding of their own text and of the subject they write about, so that they are no longer able to imagine realistic readers' needs and problems. The confrontation with realistic readers, such as in the observation groups, is in this respect very informative.

Of the three variants of learning–by–observation we compared, the most succesful was group III who observed a reader using their own text, and received written evaluative comments from this reader. It makes no difference if the observed reader uses one's own text or someone else's text (Q11: group II does not outscore group IV) but it does matter if the observations are followed by written readers' comments or not (Q13: group III outscores group II).

We have also tested three (including condition 0: four) variants of learning–by–doing activities. As reported above, the 'reinforced reflection' of group I subjects, operationalized as prompted revision, did not aid in the quality of their performance (Q5), which remained on the same level as the 'normal reflection' subjects. It was assumed that the orienters' task would prepare them better for a consecutive writing task. In comparison with the writing task performed without such orientation (group 0), orienters from group V did not do better, while orienters from group VI did better than group 0 and better than group V. The difference is in the act of writing down their comments on the reading task, which calls for processing of the information they gained during the orientation task. In all, writing comments or receiving written comments appeared to add much to the effectivity of the observations.

An analysis was made of the revised texts in order to account for the increase in textual quality. Although instructions are the central speech act in instructive text, the quality of the manuals did not increase on this aspect. Not the number of instructions had changed in the revised manuals, but non–

instructive parts: object descriptions, theoretical explanations and precautions. Apparently those information elements were identified as 'reader needs' in the observations.

EFFECTS ON GENERALISATION TASK (transfer effects)

As a result of the learning activities regarding this particular text, orienters and revisors will acquire knowledge about *how to write instructive texts in general*. This knowledge is a generalization across the learning experiences in the interventions. It is essential to understand how, and to which extent the construction of such knowledge develops in order to understand how transfer of the learning activities to new tasks takes place.

The posttest, a 'letter of advice' is an attempt to measure the subjects' declarative knowledge. The conditions that had scored highest on the writing assessment, also scored highest on this 'knowledge assessment': condition III for the revisors (anwering Q14) and condition IV for the orienters (Q17). Other differences that were expected (Q8, Q9) or not (Q12) did not occur. It again seems that the production and reception of written comments are the active elements: they may well affect the generalization level of the constructed knowledge, or at least offer a way for the learner to condensate his experiences into more easily accessible cognitions.

VALIDITY

There are some factors that may influence the internal validity of the results. One may object that the comparisons made are unfair, since the time–on–task differs between the conditions. This is true. The difference is substantial *between* the revisors' and the orienters' conditions: the orienters work in two sessions, the experimental revisors work in three sessions. However, *within* the group of revisors' conditions the differences in time–on–task between conditions are comparably small. The written comments added to condition 3, for instance, did not prolong the session for more than five minutes on a total of 40. Generating self–comments lasted, due to the method of brainstorming, elaboration, and testing of the manual, not much shorter than observing the readers on videotape. The experimental effects are very unlikely attributable to the differences in time–on–task. It should be added here that it was not my aim to make all possible comparisons between the conditions on all possible variables. I have confined myself to answering the research questions. Since there is no need to directly compare orienters and revisors interventions, the difference in time–on–task is not a problem.

The data for conditions II, IV and V were collected one year after the other data. During every data recording session the procedure was standardized (read from paper), the test assistant (the same person in all sessions) was trained only to stimulate and not to help the student, and the scoring procedure was anonymous. Thus it is unlikely that a significant part of the effects found can be attributed to between–group differences for the first and second year. It was already mentioned (in section 4.4.1) that an ANOVA did not yield any significant difference on the quality of the first versions written by the the four writing conditions, and that students were randomly assigned to the reading or writing sessions, which makes between–group differences, although not impossible, very likely either.

Some experimental groups are very small. One of the most succesful groups, III, consisted of no more than seven subjects. Generalizations of these results must therefore be made with reservations. On the other hand, the fact that the results of even some small groups differ significantly from other groups, indicates that the experimental intervention is probably quite effective.

I have partly discussed the topic of concept validity concerning the assessment of 'transfer'. I have chosen to do so indirectly, by trying to measure the declarative knowledge. Of course having such knowledge at one's disposal does not guarantee a succesful implementation in a new writing situation. Under the assumption that the knowledge needed in a new domain (writing instructive texts) is initially acquired in a declarative way, as Anderson (1987) says, I have tried to determine to what extent this first step op learning has been taken. It is supportive of this idea that the quality of the last writing task, and the number of pieces of advice given three weeks later, were so much in agreement.

The main conclusions of this experiment can be summarized as follows. Within the writing domain, learning–by–observation of readers may have two advantages over learning–by–doing. First, the revisors are better able to improve the quality of their text after observing the reader's needs and problems. Improving the text on the basis of self–evaluations contribute less, or nothing at all, to text quality. The advantage of learning–by–observation is valid for each of the three variants, although revisors profit most by observing a reader using their own text, and receiving written comments in addition. Second, this last variant of learning–by–observation is effective regarding the acquisition of declarative knowledge about how to write the particular category of texts. Such knowledge is constructed through observation and reflection activities, that may transfer to new writing tasks. Similarly, students reading such a text, experiencing its weaknesses and commenting on its qualities that caused these experiences, construct an effective orientation base that can be applied in consecutive writing tasks.

Schriver had reported on the effectiveness of a reader–oriented writing course with college students. The present experiment shows that secondary students are also capable of getting the most out of their observations.

Fig. 4.1: *Structure of independent variables in this experiment*

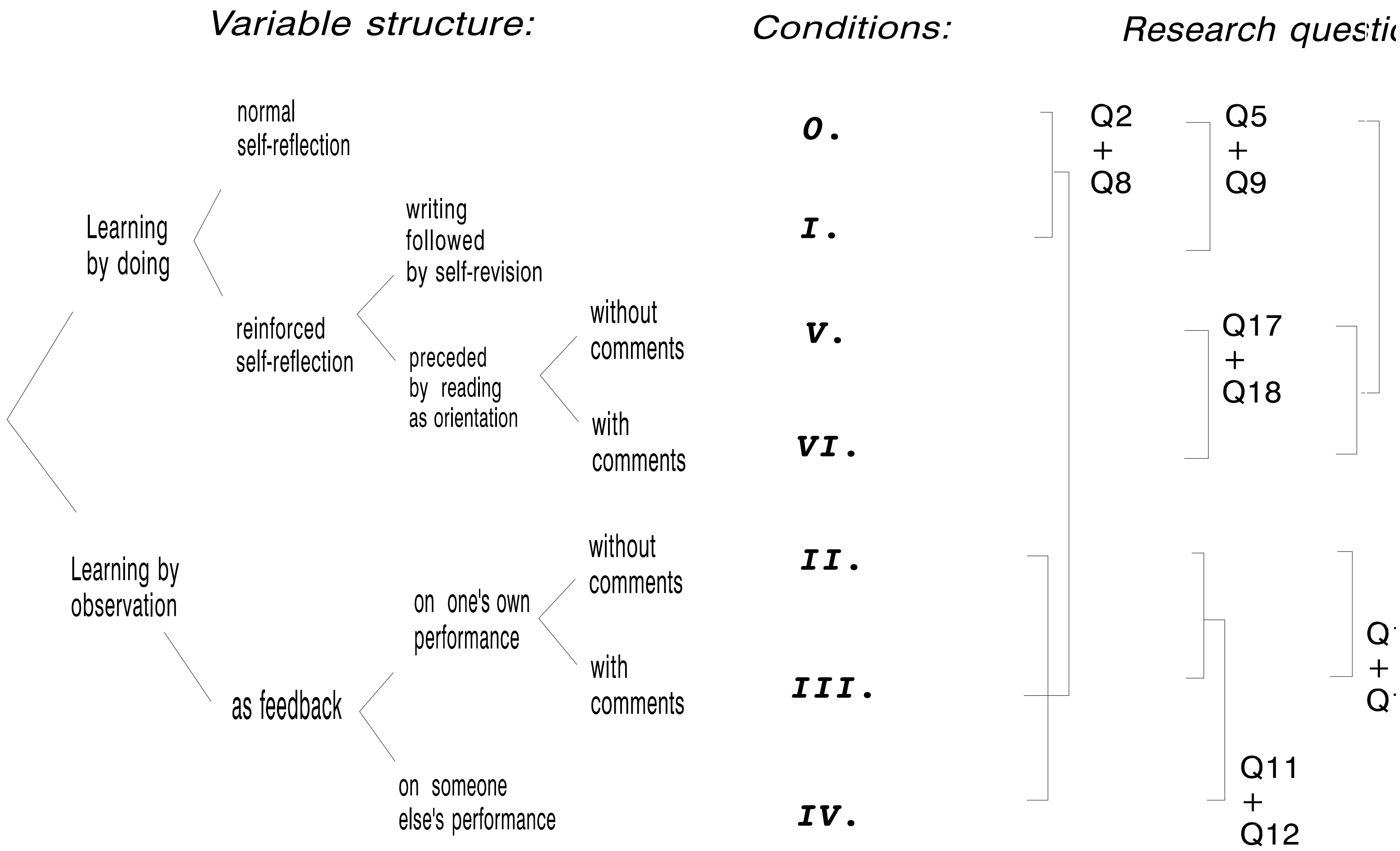


Fig 4.2 *Construction of bottle, cork and funnel.*

(hardcopy)

Fig 4.3 *Water, poured into the funnel, will not run in the bottle.*

(hardcopy)

Fig 4.4 *A straw, inserted in the bottle through the funnel, will let the air flow out and the water run down.*

(hardcopy)

CHAPTER 5: Observing writers and readers of argumentative text: Learning to write – transfer to reading

The effectiveness of three variants of learning–by–observation is compared to learning–by–doing. All variants are aimed at learning to write argumentative text. Learning–by–doing subjects learn by doing writing exercises. The learning–by–observation variants are: observing writers as a model, observing writers & readers as models, and observing readers as feedback on one's own writing performance.

Participants are third–grade secondary students, who followed one of four short experimental courses in writing argumentative text. Observations were made by means of authentic video–tape recordings (model conditions) or by 'live' confrontations (feedback condition). The participants are pretested and posttested on intelligence, reading skill and writing skill regarding argumentative text.

Results show that the 'observation' conditions are more effective than the learning–by–doing condition. 'Observing writers' and 'Feedback from readers' are more effective than learning–by–doing because they yield larger learning and transfer effects. 'Observing writers & readers' is more effective because it yields a large combined learning and transfer effect on both modes.

5.1 Introduction

The experiment reported in the previous chapter was conducted to test a theory which favors observation as a learning activity over individual practice, because of its presumed focus on monitoring and evaluative activities. Some of the main research questions – as presented in section

3.7 – were answered with respect to the writing mode, when applied to instructive texts. Questions were answered concerning the difference in effectiveness between learning–by–observation–*as–feedback* and learning–by–doing, and difference in transfer of the acquired knowledge to a generalization task. The results were generally in agreement with the theory. Moreover, it was found that positive learning and transfer effects of learning–by–observation could be enhanced by a supplementary task, aimed at processing written evaluative comments.

In the present experiment, the previous one is taken three steps further. First, the overall scope of this study as presented in section 3.7 includes two kinds of observation: not only observation–as–feedback, but also observation–of–models (models for one mode as well as models for both modes). Therefore, learning effects of both types of observation will now be assessed, so that we can answer the research questions Q1 to Q4, Q7 and Q8.

Second, the kind of transfer we seek to promote is intermodal transfer, that is transfer from writing practice to reading skill, or transfer from reading practice to writing skill. In the previous experiment, an attempt was made to assess the declarative knowledge that the participants constructed, regarding specific criteria for successful communication using instructive texts. This declarative knowledge base was regarded as a first step towards transfer (to other writing tasks or to reading tasks). In the present experiment, it is attempted to test for genuine intermodal transfer.

Third, all conditions in the previous experiment were tested for learning and transfer of one mode only (writing), while the present experiment and the experiment reported in the next chapter together offer an investigation of bilateral (intermodal) transfer, as predicted by an interactive model or reading–writing transfer (see section 3.2). The present experiment is aimed at learning effects in the writing mode and transfer to reading, while the next chapter has as its object learning effects in the reading mode and transfer to writing. The research design embracing both this and the following experiment is the 'mutual transfer' design presented in figure 2.5 (chapter 2).

The theory will again be tested with secondary students from the third grade (average 15 y.) of intermediate and high level. The text type, however, is different. In the present and the following chapter the theory is applied to *argumentative text*. This is a less 'appellative' and more 'cerebral' text type than manual texts; 'less appellative' in that it gives less occasion to visible readers' responses, from which an observing writer may derive information about the comprehensibility of the text. (In saying so it is understood that argumentative texts also make an essential appeal to the readers – that they should become convinced of the expressed opinion.) This information should now come from the thinking–aloud activities of the readers. And of the writers, as I will explain further on.

The observation methods–of–instruction that will be tested and compared to a learning–by–doing method are based on Schriver (1989; 1992; observation–as–feedback) and on Sonnenschein & Whitehurst (1983; 1984; two kinds of observation–of–models). In sections 3.5 and 3.6 these two perspectives for observational learning were presented and their differences discussed. In short, learners who observe models aim their attention at monitoring, understanding and evaluating the observed task execution activities. They evaluate in order to determine if the observed activity is an example worth following, or one that should be avoided. It is specific for this type of observation that the observed behavior matches the observer's learning goal, and that the learner does not take part in the observed communication.

Observation–as–feedback, however, relies on the observers' participation in the communication. Their initial task execution is followed by observation of a communicative partner who performs the communicative complementary task. Thus, the observers acquire authentic information about the adequacy of their performance which they cannot acquire in another way. It is specific for this type of observation that the observed behavior is complementary to the observer's learning goal, and that the observer takes part in the communication.

Both learning–by–observation methods have been presented as theoretically superior to learning–by–doing, although for different reasons. Observation–of–models would be effective because of the realistic, exemplary function of the models, and the learner's special attention for monitoring and evaluating their performance processes. Observation–of–feedback would be effective because of the authenticity of the feedback and the personal involvement of the observer. There is no theory favoring one of these methodss over the other. Nevertheless, we will check in the analysis for differences in effectivity of these rivaling methods.

In imitation of Sonnenschein & Whitehurst (1984), two variants of the observation–of–models approach are included: observation of only one communicative role (in the present experiment: writers) and observation of complementary communicative roles (writers as well as readers, who perform a complete information 'transfer'). It is expected that if communication rules are observed in both writing and the reading contexts, they will be acquired in a more abstract way (i.e. abstracted from the specific mode level) that enables active use in each of the modes. In other words, the acquired rule will possibly be available for writing and for reading tasks. Such is indicated by the successful first condition in Sonnenschein & Whitehursts study (see table 3.2).

As concerns the effectiveness of this instructional method: in comparison with observation of only one role (either writing or reading), this method offers only 50 % of the number of trials. For instance, if a student can observe six writers in one hour (when focusing on writers only), he can in that same hour observe only three writers and three readers (when observing complete communcation transfers). His experience regarding each single mode is less (so smaller learning effects are expected), but this may be compensated by his broader experience concerning the usage contexts of the acquired informtation. This instructional method would, for instance, be favorable if learning effects were equal to, or little lower than the 'one role' observations, but 'transfer' to the other mode would be higher (which can be expected since this other role was also observed).

In sum, this experiment investigates the effectiveness of three variants of writing instruction based on learning–by–observation, in a comparison to writing instruction based on learning–by–doing. All writing instruction variants aim at the domain of argumentative text. The variants are: 1) observing writers as a model, 2) observing writers & readers as a model, and 3) observing readers as feedback on one's own writing performance. These learning activities will be further explained in detail in section 5.3.3.

Like in the previous experiment, I distinguish two possible effects of the independent variables: a learning effect and a transfer effect. They are defined in a more traditional way than in the previous experiment. 'Learning' stands for the acquisition of skill within the same mode as practice or learning activities were aimed at. 'Transfer' stands for acquisition of skill in the complementary mode. Since learning–by–doing writing exercises, learning–by–observation of writers, learning–by–observation of both writers and readers, and learning–by–observation of readers as feedback are all aimed at the acquisition of writing skill, the effect on the subjects' writing ability is a learning effect. The effect on their reading skill is a transfer effect. This experiment investigates both effects.

5.2 Research questions and expectations

The following graph represents the structural relations between the independent variables in the present experiment:

[insert Figure 5.1 about here]

Figure 5.1: *Structure of independent variables in this experiment.*

The main research questions as presented at the end of chapter 3 are aimed at learning and transfer effects of learning–by–doing and of two types of learning–by–observation (observation of models, and observation–as–feedback). Observation–of–models has a one–mode and a two–mode variant.

The following research questions are addressed in this study: They can be found at the right side of figure 5.1, in connection with the independent variables and conditions in which these are operationalized.

Regarding learning effects (on writing):

Q 1: Is learning–by–observation of models (one mode) more effective than learning by doing?

Q 2: Is learning–by–observation as feedback more effective than learning by doing?

→ (Expectations: both types of learning–by–observation are more effective)

Regarding transfer effects (on reading):

Q 7: Does learning–by–observation of models (one mode) promote more transfer than learning by doing?

Q 8: Does learning–by–observation as feedback promote more transfer than learning by doing?

→ (Expectations for Q7 and Q8: observation promotes less context–specific learning and thus more transfer (but see Sonnenschein & Whitehurst, 1984))

Regarding a mixed learning- and transfer effect (on writing):

Q 3: Is learning–by–observation–of–models of both modes more effective than learning–by–doing?.

Q 4: Is learning–by–observation–of–models of both modes more effective than observation–of–models (one mode)?

→ (Expectations for Q3 and Q4: undetermined. Sonnenschein & Whitehurst, 1984 report no differences between these methods, but there have been ceiling effects; observation of models for both modes may also led to accumulation of the learning effects (with less trials per mode), and to the best scores on the criticism task.)

Operationalization of the independent variables:

Type of instruction: learning–by–doing is operationalized as individually doing short writing exercises of various types, as part of a four–lesson course on 'argumentative text'. Learning–by–observation–of–models (one mode) is operationalized as observing peer students who think aloud while they perform these writing exercises. Learning–by–observation–of–models (both modes) is operationalized as first observing a writer (thinking aloud while performing a writing exercise) and then observing a reader (thinking aloud while reading and analyzing the writer's text). Since subjects observe both modes, the progress in each of these modes is most likely due to within–mode *learning*, but can also result from between–mode *transfer*; hence the 'mixed' effect. Learning–by–observation–as–feedback is operationalized as writing an argumentative text, observing an authentic reader who thinks–aloud while reading and analyzing the text, and evaluating one's own performance on the basis of this observation.

There are no supplementary questions in this experiment.

Operationalization of the dependent variables:

There are two dependent variables: 'learning effect' and 'transfer effect'. The 'effectiveness' of an instructional method refers to the learning effect, that is, to progress within the same mode as at which the learning activities were aimed. I have already pointed out (at the end of section 5.1) that the acquisition of argumentative writing skill is the learning effect in this chapter and that the acquisition of argumentative reading skill is the transfer effect in this chapter.

Several subvariables or 'indicators' of the variable 'argumentative writing' are distinguished. In agreement with the learning goals of the intervention programs (which will be explained in section 5.3.3; cf. Table 5.2), I will attempt to assess the following subskills as indicators for argumentative writing skill (the learning variable):

1) the ability to express so–called *social parameters* in the text, thus placing the text in the social context of an argumentative discussion, by explicitizing standpoints, naming the parties involved and ascribing the correct standpoints to them, and expressing the communicative goal of the text;
2) the ability to organize argumentative texts using a *standard structure*, which asks for specific subdivisions of introduction, body and ending;
3) the ability to express hierarchical *argumentation structures* in prose, which can be reconstructed by readers;
4) the ability to use verbal means to *enhance the presentation* of argumentation and of text structure.

These abilities will be measured in three posttests (see section 5.3.4). There are no specific expectations as to which of these abilities will profit most from the experimental interventions.

The other dependent variable, transfer, is defined as the progress within the complementary mode, in this case reading. Thus, for a course aimed at development of writing, effects on reading skill are regarded as transfer effects. In analogy to the learning variable, the indicators of the transfer–variable 'argumentative reading' can be specified as:

1) the ability to identify so–called *social parameters* in the text, by which the text is placed in the social context of an argumentative discussion;
2) the ability to analyze argumentative texts in terms of a *standard structure*, which asks for specific subdivisions of introduction, body and ending;
3) the ability to analyze hierarchical *argumentation structures* in prose, as expressed by the writer;
4) the ability to identify verbal means by which the *presentation* of argumentation and of text structure has been enhanced.

Investigation of transfer effects will require quantification of the concept 'transfer'. A viewpoint advocated by Singley & Anderson (1989) was adopted in which transfer is defined as the ratio between the progress made in a certain domain, and the maximally feasible progress in that domain, within the same time, under the same circumstances. In this way, transfer can be defined and interpreted as a proportion (or percentage) of a meaningful maximum (see section 2.6).

The selection of a criterion that can represent this 'maximum feasible progress' must be motivated. To obtain meaningful transfer scores, I will use the learning–by–doing posttest scores for the complementary mode as transfer criteria. The motivation is that these learning activities are a) 'standard' because of their general use and b) straightforward in that they are directly aimed at acquisition of the behavior–to–be–learned. In other words, the amount of transfer is measured against the standard of 'what one would normally do to learn these skills'. It is also in agreement with the 'mutual transfer design' presented at the end of chapter 2 and the transfer formula belonging to it. Further explanation can be found in section 2.6 or 6.4.2 of the next chapter.

5.3. Method

5.3.1 Design

An experiment was set up in order to test our hypotheses. The research design can be schematized as in table 5.1:

Table 5.1: ***Experimental design for 'learning to write argumentative texts'***

Condition:	*n:*	*Pretests:*	*Learning activities:* (4 lessons of 1 hour each)	*Posttests:*
DW	30	X	Doing Writing exercises	X
OW	30	X	Observing Writers as models	X
OWR	30	X	Observing Writers & Readers as models	X
FW	30	X	Observing Readers as Feedback on Writing	X

This can be characterized as a pretest–posttest control–group design, in which the 'learning by doing' group is denominated as control group. Pretest scores, assessing intelligence and pre–experimental skill in reading and writing argumentative text, will be used as covariates in the analyses of posttest data, in order to correct for potential initial differences between the groups, and to measure the effect of conditions as exact as possible. Since subjects were randomly attributed to the groups, these differences will most probably be rather small (see 'subjects').

The experimental conditions differ with respect to the type of learning activities. Posttests, however, are the same for all conditions: measurements of both reading and writing skill. The writing posttests thus constitute *intra–modal learning* measurements for the DW, OW and FW conditions (because training and testing are within the same mode). Similarly, the reading posttests are *transfer* tests for DW, OW and FW.

The position of the OWR condition is different, since its training is aimed at both modalities. Thus, the writing and reading posttests will measure *learning* in the writing mode, but also learning in the reading mode. For this reason, the data from this condition will also be used in the next chapter (which deals with learning to read). It is not necessary to assume between–mode transfer effects in addition to the learning effects within each mode. Even if such transfer would occur, it might sink into insignificance beside the learning effects. The reason why OWR is interesting, is that the learning effects regarding each mode seem to accumulate, even though there are less 'trials' in each mode, in comparison to e.g. OW.

5.3.2 Subjects

In all, 120 students who had just finished the 9th grade (intermediate and high level) took part in the experiment. The average age was 15.5 years. 65 % of the participants were female; boys and girls were almost equally spread across the conditions. The students came from 8 different city schools and participated voluntarily in the experiment. The schools had been approached by telephone and post and were willing to distribute subscription forms among the students, which they could send post–paid to the research institute. For their participation the students received a modest financial reward. The data collection took place during their summer holidays.

Atypical in the sample is the two–third majority of girls. Internal validity is not threatened since students were randomly distributed. External validity may however be limited, if sex is a relevant variable in this subject or this kind of learning method.

Assignment of subjects to the four conditions was semi–random. That is, a stratification was applied first with respect to level of education (intermediate vs. high); further assignment within the strata was random. As a result, in each condition precisely 12 students took part from intermediate level and 18 students from high level. The pretest measurements on IQ, writing and reading skill enable a check on equality of the groups.

5.3.3 Training materials: experimental writing courses on argumentative texts

A communication course was required dealing with one text type, which could be adapted for both writing and reading instruction. This adaption had to be done by adding a particular type exercises that would evoke the learning activities typical for a certain condition. Further, it was necessary that in the theoretical part of this course, properties of the text type be treated 'neutrally', i.e. with a modest and balanced attention for application in reading and writing (because the real focus on reading or writing would be made in the exercises). Moreover, the text type had to be relatively new to the students. Finally, the level of the course had to be sufficiently high to avoid ceiling effects, so that potential experimental effects would become observable.

For these reasons 'argumentative texts' were taken as the subject on which a four–lesson course was developed. These are texts in which a speaker or author defends a standpoint by supplying argumentation. In The Netherlands, the ability to analyze or compose argumentative text will be obligatory for the highest streams of secondary school (CVEN, 1992). Although 9th grade students have already had some experience with verbally expressing and explaining their opinions, they have not yet received formal instruction on the composition or analysis of argumentative text. The subject–matter is generally abstract in nature, since it calls on the ability to invoke complex speech acts and thinking skills, such as arguing, refuting, comparing and contrasting, selecting main ideas and paraphrasing.

Learning goals and theoretical contents

In selecting subject–matter for the argumentation course, we joined the pragma–dialectical perspective on argumentation that Van Eemeren and Grootendorst advocate (Van Eemeren & Grootendorst 1983; 1992). The advantage of this perspective is the explicit framing of argumentation within the social situation of a (critical and problem–solving) discussion. I expect that this social perspective with its distinct, but related communicative roles will allow for integration of receptive and productive skills in a way that is meaningful to the student.

The learning objectives for the argumentation course are shown in table 5.2:

[insert table 5.2 around here]

Table 5.2: *Learning objectives for the short courses in 'argumentative texts'*

There are four objectives of the short course, each corresponding with the content of one lesson. The objectives for reading and for writing are mirror–images: the strategy that a writer applies in structuring his text can be reversed and used by the reader to analyze this structure. For instance, if the writer has used verbal markers for his argumentation ('First... second....') then readers will use such markers to enhance their analysis. Or readers may look for a clear presentation of the issue ('the main problem is that...') which the writer has intentionally formulated in this way.

A summary of the theoretical contents of the four lessons is displayed in table 5.3.

[insert table 5.3 here]

Table 5.3: *Theoretical contents of the course on 'Argumentative texts'*:

For our selection of objectives and contents of the four–lesson course I am obliged to Van Eemeren & Grootendorst (1983) and Koetsenruijter & Slot (1990).

The complete course consists of four lessons and is self–instructive: all subjects work individually without a teacher's help. Each lesson lasts one hour. There is a workbook for each lesson, containing theory and exercises. The theory is divided in small parts of 1 – 1½ page and is explained using many examples. The construction of the course is cumulative, so that in each consecutive lesson the theory from the previous lessons is repeated and extended. In order to stimulate the cumulative acquisition of knowledge by the students, each lesson starts with a summary of previous theory which must be actively studied and completed by the subject.

Instructional sequence

The theory on argumentative texts forms the backbone of the four different courses that are developed for the experimental conditions. Nevertheless, the subjects spend about 70 % of the time on the exercises in which the theory must be applied. The nature of a course as a learning–by–doing course or a learning–by–observation course is therefore not at all determined by the theory, but only by the type of exercises. Figure 5.2 shows the similarities and differences between the four courses (DW, OW, OWR and FW) with respect to the instructional sequence of theory and exercises.

[insert Figure 5.2 about here]

Figure 5.2: *Instructional sequence of the courses*

The chronological order is from the top downward. Subjects in all conditions study the same theoretical part, and subsequently answer one or two 'control questions'. These questions ask for the gist of the part that has just been studied and are intended to stimulate active reading of the theory.

Next, subjects apply the theory in *one* of four different types of exercises: individual writing exercises (DW), observation of writers (OW) or communicative dyads (OWR) , or observation as feedback on one's own writing exercises (FW). After completing one or more exercises, subjects continue with the next portion of theory, the next control question, the next exercise, and so on. The subject–matter presented in the courses is identical, but the nature of the exercises differs: the theory must be applied to either writing exercises or observation exercises regarding writing argumentative text.

Exercise types

We will describe the differences between the types of exercises by means of an example from the first lesson. In the theoretical part, the two characteristics of argumentative text have been introduced: an opinion is stated, and one or more reasons for having this opinion are supplied. Some examples are given:

> *'I think we should go to Italy for our holidays, because the whether is always fine and the food is great.'*

> *'You must really put the volume of your music down. I cannot work with all that noise in my ears.'*

The learning that takes place is a form of concept learning (Mayer, 1983). Subjects learn the concept 'argumentative text', and to identify a text as belonging to this subclass of texts, according to a conceptual rule:

> *'S and A are the essential parts/properties of a type B text'*

As we have seen in section 3.2, such 'neutral' conceptual rule can be used to build productions that can be used in either receptive or productive communication. I will first give a 'receptive' formulation of this rule, which is aimed at the identification of properties from which class membership is inferred.

< **if** (goal = typify text as type B)
then (actions = (identify characteristic A) **and** (identify characteristic B))>

In this form, the conceptual rule can be used in reading tasks (identification of text elements) and for self–checking in writing tasks. Its counterpart, the productive form, can be used in writing tasks for generating activities:

< **if** (goal = produce text type B)
then (actions = (produce characteristic A) **and** (produce characteristic B))>

After the subjects reproduced the characteristics on paper (control question), they start the first exercise.

DW (learning by doing writing exercises)

The subjects from the DW condition do the following assignment:

DW:

Check again the three examples on page 2 and then **write three new examples of argumentative texts.**
1. ....

....

The DW subjects must use the rule productively. In the workbooks a limited space is reserved for the answer, so they must confine themselves to application of the rule. More specifically, they must inductively give meaning to the characteristic concepts 'opinion' and 'reason for having this opinion', aided by the examples. Secondly, they must understand that *both* characteristics are necessary to meet the rule, so that opinions only, however floridly presented, will not suffice. Finally, they must generate new instances of the characteristic concepts. Like in all the other conditions, the subjects reveived no external feedback.

OW (observation of writing)

OW subjects do not have to do the writing exercises themselves. Instead, they observe age–group students doing these exercises. Authentic videotape recordings of the age–group students are used for the observations. The observed students think aloud while writing, so the observer can closely follow the writing process. The assignment runs:

OW:

Read the following assignment, and imagine how you would answer it.

> "Check again the three examples on page 2 and then **write three new examples of argumentative text**"

You are going to see two students doing this assignment.
It is your task to find out what they do well, and what they do wrong.
When you have observed both students, you may advance to the next page.

(...)

(next page):
You saw two students doing the assignment. They wrote the following texts:

Student 1	***Student 2***
"I don't need a dog any more because I already have three"	*"Dogs are more fun than cats, but they need much more attention"*

===>>> Which student did better, according to you? Student .....

===>>> Explain briefly why you think the other student did worse.

Student ..... did *worse*, because ....

....

The subjects get oriented on the observation exercise by reading the writing assignment. Next they are explicitly instructed to aim their attention at evaluating the observed students' task performance, which should stimulate engaged and therefore instructive observation. Observation thus holds that the subject checks the application of the rule by the observed students.

After having observed two different student writers (see section 'Procedures') the subject must determine if one did worse than the other, and explain what exactly made this performance less successful. In this way the subjects are forced to designate 'good models' and 'worse models'.

It should be noted that the subjects in this condition not only *observe* writing processes, but also *perform* comprehension processes. In order to evaluate the texts of the observed writers, they must analyze them in terms of the argumentative characteristics.

OWR (observation of writers and readers)

Whereas the evaluation in the previous condition is focused on only one communicative role, for the OWR subjects it is aimed at both roles. They observe a complete communicative process: the construction of a text by a writer and the reconstruction of the writers' intention by a reader. Their exercise runs like this:

OWR:

Read the following assignments:
The first student you will observe is the **writer** who was instructed to:

"Write a short argumentative text'

After the text was written, the second student or **reader** was asked to:

'Determine if this is an argumentative text. Tell us why."

You are going to observe both the writer and the reader.
It is your task to find out what each of them does well, and what they may be doing wrong.
When you have observed both students, you may advance to the next page.

(....)

(next page:)
You saw two students doing writing and reading assignments. Their answers were:

Writer:	**Reader:**
"'I think I will enjoy reading this book, because I already like the introduction.'	*'Yes, that is an argumentative text. That's because she gives her opinion about the book."*

Explain briefly on which aspects the communication was successful or not.

===➢➢➢ Did the **writer** do well?
O Yes O No
because ……………………………………………………………………………..

===➢➢➢ Did the **reader** do well?
O Yes O No
because ……………………………………………………………………………..

OWR subjects must divide their attention between the two communication modes. More than for the other subjects it may become visible for them how strongly writing and reading – or the construction and reconstruction of meaning – are related through the use of the conceptual rule for 'argumentative text'. The subjects must evaluate writers and readers by their use of this rule; or, more precisely, of the two variants of the conceptual rule mentioned above. Because of this varied representation of the rule, the theoretical element may become more flexible and therefore more readily transferable to both reading and writing. On the other hand learning and transfer effects are inseparable in this condition. Therefore we have called the resulta 'mixed learning and transfer effects'.

FW (reader observation as feedback for writers)

In the previous two conditions, observation of models was a means to gather information about successful and unsuccessful reading and writing processes. By having the subjects focus on process evaluation, it is expected that learning takes place by imitation of good examples and avoidance of worse examples.

In the present condition, the observations serve a different goal. FW subjects start with application of theory in a writing exercise, then hand over the text to an age–group reader, and observe this reader while he/she performs an analysis task by the same criteria as with which it was constructed. More about the function and selection of this 'test reader' can be read in section 6.3.

An illustration of this type of assignments can be found on the next page.

The FW subject does a writing task in the same manner as a DW subject. But instead of using all practice time for writing exercises, 50 % of the time is spent on observing readers' as feedback on the writing task. So the time–on–task is the same for these conditions.

An FW subject also has to process the two variants of the concept rule: the productive variant while writing and its receptive counterpart while following the observed reader's activities. The observation is aimed at discovering possible reader's comprehension problems, attribution of these problems to the quality of the text, and proposing possible remedies.

FW:

Check again the three examples on page 2, and
write a new example of argumentative text
When you have finished your text, you present it to the reader.
....

(....)

Now you will see a student analyzing your text while reading aloud. It is the reader's task to find out:
a) whether your text is argumentative or not
b) which part represents the opinion, and which the reasons for it

Now observe this reader's performance. Don't say anything.

It is your task to check if the reader can fluently perform these tasks, and if not, to find out *why* not.

(....)

Now answer these questions:

a) Has the reader recognized your text as *argumentative* ?
O No O Yes

b) Was the reader able to correctly identify the *opinion* and the *reasons* for it ?
O No O Yes

c) Can you change the text in any way, in order to make it (even) clearer for the reader?
O No
O Yes, namely ....
....

The Videotape Recordings

Videotape recordings of students executing writing and reading processes are required for the OW and OWR conditions. There recordings were made on another school, with students of the same age as students in the sample. The recordings were not staged, but authentic. They were made by having 16 students think–aloud while doing all four lessons of the argumentation course in front of the videocamera.. 8 students had made the writing course (DW version) and 8 students had made the reading course (DR version). The students had experienced two small thinking–aloud tasks prior to the recordings. During the recordings, they were prompted to think aloud whenever they had to do an exercise. Thus, a collection of videotaped, successful and less successful task executions could be made, from which we could choose in editing the tapes for the experimental sessions. The recordings were made using microphones, which helps to get the necessary clear sound and comprehensible speech.

The videotapes for the experimental sessions were edited such, that for almost every exercise two different processes or solutions can be observed. This was done in order to provoke active interest from the observer and to offer him/her a problem to solve: to choose the best from the two realistic solutions to the task.

In Appendix 5.1 an index to the videotapes is given.

Prior testing of the lessons

The workbook–lessons of condition DW were priorily tested in order to check for comprehension problems in the text, and in order to check for timing problems to see if the work would take up a full hour. Due to the prior test, some small corrections in the text were made and one exercise was removed from the last lesson in order to not make this lesson not too long.

We can make a comparison between the four conditions with respect to the type of cognitive activities they perform. By doing so, it becomes clearer to which differences in activities we may attribute possible differences in effectivity.

Table 5.4: *Comparing learning activities in the DW, OW, OWR and FW conditions*

	DW: 'Write a correct argumentative text.'	**OW**: 'Check if P writes a correct argumentative text.'	**OWR**: 'Check if P writes a correct argumentative text, and if Q analyzes it correctly.'	**FW**: 'Write a correct arg. text and check if Q analyzes it correctly.'
ORIENTATION:	Reading, interpreting and conceptualizing the coding rule: B $\rightarrow$ S + A.	same	same	same
EXECUTION:	Concretizing: writing S and matching A	Observe writers concretizing by writing S and matching A	1a: Observe writer 2a: Observe reader	same as DW
MONITORING & EVALUATION	Self–monitoring & self–evaluation (?)	Observe writing and give evaluative comments	1b: Comment on writing 2b: Comment on reading	Self–evaluation by means of reader observation
REFLECTION:	Accommodate: change or confirm conception of coding rule	same	same	same

We see that the 'orientation' and 'reflection' steps in the exercises are similar for each condition, although there may be small differences regarding the 'reflection' stage (the last question of FW excercises can be considered a reflection prompt). In these steps, the learner's cognition is construed (making an initial conceptualization of the rule in the first step) or accommodated (changed or confirmed as a result of experiences in the exercise). The steps in between are 'practice steps' in which the execution of a linguistic process is either undertaken or observed, and evaluated. As advocated in section 3.3, these (self)observation and evaluation activities yield the information that is the basis for learning. 'Reflection' steps are not prompted; that is, the extent to reflection is left up to the subjects. By invoking these different learning activities in the four conditions, we attempt to find if there are any differences in learning associated with them (as theory suggests).

5.3.4 Test materials

The tests are aimed at the measurement of the dependent variables, as operationalized in four indicators (see section 5.2). Adding pretests may enhance the quality of posttest measurement, if pretest scores can be successfully used in a covariance analysis to filter out undesirable disturbing effects, such as the influence of pre–experiment abilities which may not be equally divided across the conditions in spite of randomization.

In the posttests, all indicators for writing and reading skill are operationalized. This is – for practical reasons – not the case in the pretests. Therefore I will first explain and exemplify the measurement of dependent variables in the posttests, and then relate this to the pretest measurements.

Posttests for resulting writing ability and for resulting reading ability

Six posttests were administered: three for the measurement of writing skill (the learning measure) and three for reading (the transfer measure). The four indicators of either writing or reading skill (see section 5.2), which match the learning objectives and lesson contents (section 3.3.1) are measured by one or more parts of the three posttests for that skill (Fig 5.3):

[insert Figure 5.3 about here]

Figure 5.3: *Variable structure of experiments reported in chapters 5 and 6.*

Figure 5.3 represents the variable structure that underlies both experiments reported in this and the next chapter. The figure may clarify the mirror–image relationship between the two experiments. The arrows between independent and dependent variables indicate causal influence, as predicted by the theory. The diagonal arrows represent transfer effects, the straight arrows learning effects. The small arrows on the right side of the figure indicate that the four indicators for each mode operationalize the dependent variable for that mode. Thus it may become clear that operationalizations of reading and writing skill have a highly related content. On the very right side one can see in which posttest each indicator is measured. For instance, the explicitation of 'social context' is measured in part D of writing test W1; writing on the basis of a text structure is measured in part A of W1 and in part B of W3, and so on. I will explain and exemplify each of these indicators.

1a) **Writing – social context (W1D):** the ability to put the text in its social context, by explicitizing the issue, the defended viewpoint (positive, negative, neutral), other parties in the discussion and their viewpoint. These items, taught in the first lesson, are scored in an argumentative text that the subjects had to write on a given subject. For instance, all items are present in the following introduction:

> *"Do you think that students should give grades to their teachers? Well, the teachers certainly don't think so. But I think it is a great idea."*

1b) **Writing – text structure:** (W1A, W3B): the ability to organize argumentative texts using a *standard structure*, which asks for specific subdivisions of introduction, body and ending.
The following twelve parts are taught in lesson 2 and scored in the texts written by the subjects: Introduction (consisting of: attractor, issue, parties and their standpoints), Body (consisting of: standpoint, argumentation, refutation) and Ending (consisting of: conclusion, most important argumentation, consequence or exhortation). For instance, a good attractor is:

> *"You have failed your test!" Huh, I wonder, what would that mean? I was soon to find out..."*

and a good exhortation:

> *"If you agree with me, come on an write a letter to the head of the school."*

1c) **Writing – argumentation structure (W1C, W3B):** both simple (singular, subordinate, compound argumentation) and complex (any combination of these argumentation types). Tests are administered in which the student is shown a hierarchical argumentation structure, which he should write out in words such that a capable reader can reconstruct the original structure, which has been practiced in lesson 3. For instance, the structure:

(1) It is time to go home

↑ ↑

(1.1) Your mother is waiting (1.2) You have become tired

cannot be recognized if it is written down as:

> *"I think it is time to go home. You have become tired, so your mother will be waiting for you."*

1d) **Writing – means for presentation (W1B, W3A):** the ability to use verbal means to *enhance the presentation* of argumentation and of text structure. Texts written by subjects are scored on three items which were taught in lesson 4: use of paragraphing to highlight textual organization, use of markers to indicate specific textual parts, use of connectors to explicitize the relation between textual parts. For instance, in the sentence:

"The cupboards are all empty; it's time to do shopping"

the use of standpoint and argumentation markers aids to the comprehensibility, namely::

"I think that the cupboards are all empty, since it's time to do shopping"

as opposed to:

"I think it's time to do shopping, since the cupboards are all empty."

Indicators 1a, 1b, 1c (simple) and 1d were measured in two tests for writing argumentative texts or text fragments (W1 and W3). Only indicator 1c (complex) requires a specific test (W2). We have no specific expectations as to which of these abilities will profit most from the experimental interventions.

In analogy, we can specify the indicators of the variable 'argumentative reading', which is assessed as intra–modal transfer measure (cf. Table 5.2)

2a) **Reading – social context R3A,B,C,D:** the ability to identify so–called *social parameters* in the text, by which the text can be placed in the social context of an argumentative discussion.
For instance, subjects are required to identify issue, parties and their standpoints in the following fragment:

"Well, they are simply too lazy to work". That is what I keep hearing when I ask people what we should do about the growing army of the unemployed. Recently I found more and more people talking about the question whether the labour act of 1963 shouldn't be sharpened. There is quite some disagreement: our government seems to be quite fond of the idea, and the Parliament has reacted rather moderate – but hasn't condemned the plan either. The unions, however are furious because they oppose every form of forced labour. Personally I feel little sympathy for a change of law, and I will gladly explain why."

2b) **Reading – text structure (R2D, R3F):** the ability to analyze argumentative texts in terms of a *standard structure*, which asks for specific subdivisions of introduction, body and ending;
Two texts (400 and 500 words) must be analyzed using the same structure as was presented above under 1b). The texts have been specially constructed for the purpose, which makes the job not too

difficult.

2c) **Reading – argumentation structure (R2C, R3E):** the ability to analyze hierarchical *argumentation structures* in prose, as expressed by the writer. There are two variants. The first is *simple* argumentation; we chose a subset of the instruments developed by Oostdam and Couzijn (Oostdam & Couzijn 1989; Oostdam 1991), aimed at the analysis of singular, compound, and subordinate argumentation. These instruments were developed to be administered to students who never received any instruction in argumentation; besides, they have proven to be very reliable (< .85). An example:

> *"That old newspaper is always handy to set the barbecue to fire, so I think we really shouldn't throw it away. And we can use it to wrap up the flowers, too!"*
>
> Does this text contain argumentation? O YesO No
>
> If yes, underline the argument(s).

The second variant is *complex* argumentation: argumentation structures have to be drawn by argumentative analysis of two pieces of text. The final structure includes single, multiple, and subordinate argumentation.

2d) **Reading – means for presentation (R2A,B, R3G):** the ability to identify verbal means by which the *presentation* and thus the comprehensibility of argumentation and of text structure has been enhanced.

In two texts, the subject has to identify all textual markers and all argumentative connectors. A sample:

"I<u>n my opinion,</u> changing the law is certainly worth it, <u>because</u> the unemployment in a country should be as low as possible, <u>for the simple reason</u> that we shouldn't wast manpower <u>and</u> we shouldn't waste money.

Measurements of indicators 2a, 2b, 2c (complex) and 2d was done by analysis of two comprehensive texts (posttests L2 and L3). The other posttest for reading (L1) measured indicator 2c (simple) and consisted of multiple choice items as described above.

For scoring of the posttests, see Appendix 5.2. For reliability coefficients see section 5.4.1.

Pretests (covariates) for IQ, initial writing ability and initial reading ability

Covariate analysis requires pretest measurement of relevant variables, which may be - unintentionally –

included in the posttest measurement and which may – unintentionally – influence the experimental effects. The initial skill level in reading and writing argumentative text is such relevant variable, for instance. The problem is that the initial skill level can be considered rather low (in the lower streams there is not much systematic attention for argumentative texts at school) which impedes measurement with the same instrument as used to measure the resulting level. Therefore we have measured not all indicators, and added pretests for intelligence as an alternative explanatory factor for differences in posttest performance. The following indicators were measured in the pretests:

1a) **Writing – soc. context**

1b) **Writing – text structure**

1d) **Writing –** means for presentation

These indicators for writing were measured in the pretests by asking the subjects to write two short argumentative essays, in which the standpoint 'the bicycle is the best means of transport' and 'school uniforms must not be introduced in our school' had to be defended on the basis of some documentation. The texts were scored on the categories mentioned above under 1a, 1b and 1c. For reliability assessment, see section 5.4.1.

2c) **Reading – argumentation structure**

This indicator was measured by two multiple choice tests (2Ca (simple) and 2Cb(complex), which are parallel to the test used to measure the indicator 2c) on the posttest (the same item construction, but a different content).

Added were pretests for the assessment of (relevant aspects of) IQ, since one can interprete the analysis of argumentation as an ability to discern abstract relations between verbal units (Oostdam 1991). Therefore we chose two validated CMR tests (Cognition of Meaningful Relationships) and one CMU (Cognition of Meaningful Units) test for the measurement of 'intelligence'. The tests were 'Conclusions' (Elshout 1966), Word list (DAT 198..) and Verbal Analogies (DAT 198..).

The variable list can now be summarized as follows:

		pretest	posttest
1a)	**Writing – social context:**	1A	W1D
1b)	**Writing – text structure:**	1B	W1A, W3B
1c)	**Writing – argumentation structure (simple):**	–	W3
	Writing – argumentation structure (complex):	–	W1C, W2
1d)	**Writing – means for presentation:**	1D	W1B, W3A
2a)	**Reading – social context:**	– –	R3 ABCD
2b)	**Reading – text structure:**	– –	R2D, R3F
2c)	**Reading – argumentation structure (simple):**	2Ca	R1
	Reading – argumentation structure (complex):	2Cb	R2C, R3E
2d)	**Reading – means for presentation:**	– –	R2AB, R3G
iq1	**Intelligence – CMR 'Conlusies III':**	IQ1	– –
iq2	**Intelligence – CMU 'Word list':**	IQ2	– –
iq3	**Intelligence – 'Verbal Analogies':**	IQ3	– –

Thus, the instruments used were: 3 IQ pretests; 1 writing pretest; 1 reading pretest; 3 writing posttests (W1, W2, W3) and three reading posttests (R1, R2, R3).

5.3.5 Procedures

For each subject, participation in the experiment took place in two sessions during two consecutive mornings or two consecutive afternoons. On the first day, the pretests were administered during the first two hours, after which students followed lesson 1 and lesson 2. On the second day, the course continued with lesson 3 and lesson 4. Immediately after lesson 4, the posttests were administered during the last two hours. One hour of the posttests was necessary to measure several aspects of reading ability, and one hour for the measurement of several aspects of writing ability. The order in which the three subtests for reading and the three subtests for writing ability were made varied, so that not all writing tests followed or preceded all reading tests.

All subjects from conditions DW, OW, OWR and FW worked individually from a workbook, in which theory and exercises were combined. Several subjects worked at the same time in one room; it was not allowed to conversate during the tests or training. Only condition FW required co–operation between the subjects, so these sessions were limited to small groups. Time–on–task was scheduled to be the same for every condition. Students were informed about the time every fifteen minutes so they would not be surprised by a sudden deadline. Moreover, the video–conditions were also timed by the length of the videotape which was between 57 and 63 minutes, and an on–screen timer in between the fragments for observation.

Learning-by-doing condition (DW)

Sessions of the learning–by–doing conditions were rather straightforward. After groupwise administration of the pretests, subjects had a workbook and a pen at their disposal and could work individually in a normal tempo until the hour was over. Due to the fact that the course had been pretested, the time estimation of one hour appeared to be sufficient.

Learning-by-observation of models (OW and OWR)

The sessions for learning–by–observation of models were more complicated. Subjects had a workbook and pen at their disposal, and were seated facing a videoplayer. Each subject had a 'personal' videoplayer and headphones. At the start of each lesson, a tape would be inserted and started. By means of an on–screen timer and on–screen messages the subject was informed about how much time was left for each acitivity: reading a piece of theory in the workbook, answering a control question, and doing a reading or writing exercise. Short beeps made the student aware that an observation exercise was approaching; this to make sure that the fragments would not be missed.

After two or three video fragments of students doing an exercise, there was ample time for the subject to write down the comments on the observed performances.

In this way, the time spent on the various parts of the lessons was highly controlled. Because the tempo was not too high and the tape could be stopped if necessary, this appeared not to entail more stress for the subjects than in the learning–by–doing conditions.

Learning–by–observation–as–feedback (FW)

The procedure for the FW condition was very different. Subjects of the FW conditions worked in tryads: one student, the so–called 'feedback supplier', who did not take part in these four conditions, served as the 'test reader' for two FW writers, who came in turns to offer their texts for analysis.

All subjects worked from their workbooks. After an FW writer had completed a writing exercise, the text was presented to the test–reader. The writer was seated oppositie the test reader, at about 1.5 meter distance, in order to observe the reading. The reader had a certain amount of time to complete the reading/analysis task. After that time, it was the second writer's turn to have his/her text analyzed; the first listener went back to his/her place and continued with the work. The tryads thus worked according to a strict time schedule. They timed themselves using a stopwatch and a workplan described by the minute. The turn–taking went on throughout the course and was supervised by a research assistant. Precautions were taken so that the test–reader's thinking–aloud would not disturb the writers at work.

The organization of the sessions in this group is so demanding, that 6 groups (of 15) had to work 5–10 minutes longer in order to complete the lesson. Since most of the delay was caused by waiting, moving places etc. we did not find it necessary to make a correction in the data..

5.4. Results

The results of this study will be presented in two parts. First we will report on the instrumentation for the measurement of pre– and posttest variables. Quality assessment is necessary because the instruments differ in nature and length, and because most of them were constructed for the purpose of this study and thus not tried out elsewhere. We will also pay some attention to pretest scores by reporting and discussing their intercorrelations. Differences in pretest scores between groups are not statistically tested, since we will attempt to use pretest scores as covariates in the analysis of posttest data. The reason is that we want to estimate the net effect of the treatments, stripped of the influence of other factors that are likely to affect the posttest score. These other factors are intelligence and initial writing ability or reading ability concerning argumentation. In such a covariance analysis, the posttest scores must be regressed first on the relevant covariating pretest scores; then the part of the posttest score that can be safely attributed to the pretest is subtracted from the posttest scores, and a variance analysis is performed on these corrected posttest scores. To this end, quantitative relations between the posttests are tested and discussed using a correlation matrix.

In the second part, the research hypotheses will be statistically tested, and a report is given on the multivariate variance analyses performed on the posttest data using the relevant pretest data as covariates. The first of the two sections in this part is about the learning measures: to which extent do the various types of practice yield skill acquisition in the same mode as practiced? The second section concentrates on the transfer measures: to which extent is transfer to the opposite mode stimulated by each of the types of practice?

5.4.1 Instrumentation

A variety of tests had to be used as instruments for the measurement of dependent variables and covariates (see section 5.3.4). We will list these instruments, give a short description and some psychometric data: number of items ('standard items including rejected' and 'rejected items' after item–analysis using $R_{item\text{–}total} \geq 0{,}15$ as a criterion) and homogeneity (after removal of non–fitting items)..

Table 5.5: *Number of items, number of rejected items, and reliability indices for pretests and posttests*

	Indicator:	**Pretest**	# items totaal	#items rejected	alpha	**Posttest**	# items totaal	#items rejected	alpha	correlation pre–post
1a)	**Writing – social context:**	1A	6	1	.55	W1D	3		.71	.097
1b)	**Writing – text structure:**	1B	18	4	.63	W1A, W3B	9	2	.62	.210
1c)	**Writing – argumentation structure (simple):**	– –				W3	11	1	.86	
	Writing – argumentation structure (complex):	– –				W1C	12	3	.64	
						W2	32		.92	
1d)	**Writing – means for presentation:**	1D	6	1	.62	W1B, W3A	14	2	.73	.091
2a)	**Reading – social context:**	– –				R3ABCD	4		.55	
2b)	**Reading – text structure:**	– –				R2D, R3F	14	3	.85	
2c)	**Reading – argumentation structure (simple):**	2Ca	33	3	.86	R1	48	5	.93	.450
	Reading – argumentation structure (complex):	2Cb	20	5	.74	R2C, R3E	19	2	.83	–.016
2d)	**Reading – means for presentation:**	– –				R2AB, R3G	18		.51	
iq1	**Intelligence – CMR 'Conclusies':**	IQ1	40	1	.88	– –				
iq2	**Intelligence – CMU 'Word list':**	IQ2	75	10	.89	– –				
iq3	**Intelligence – 'Verbal Analogies':**	IQ3	50	1	.91	– –				

Posttest measurement of indicators 1b, 1d, 2b, 2c (complex) and 2d took place with more than one test. The relevant parts of the test were taken together in the analysis. The psychometric data reported in table 5.5 are also based on these parts together.

The quality of each test is indicated by its homogeneity (or reliability) and other aspects of its validity. Since we cannot validate the tests in another way than by face–validity, we must confine ourselves to the assessment of homogeneity.

The homogeneities of the tests (Cronbach's alpha reliability of each test without its rejected items) are satisfactory in most cases (> 0.60), with the exception of the pretest measurement of indicator 1a and the posttest measurements of indicators 2a and 2d. It is not surprising that these tests all have a low number of items. When corrected for test length (with the Spearman–Brown formula), the relibiability of these tests falls within the acceptable range as well.

Items with an item–total correlation of less than 0.15 were rejected from the tests. This concerned mostly items with an unclear or ambiguous formulation, thus functioning as trap questions, or items with an extraordinarily high p–value, which could not discriminate between overall high scorers and overall low scorers.

With the help of a pretest–posttest correlation table (see Appendix 5.3) we can determine which pretest variables may function as covariates. Pretest variables have only a function as covariates if they show a significant correlation with the corresponding posttest scores (Edwards, 1985). That is, if pre– and posttest appear to have a theoretical (=interpretable) and empirical overlap. In the application of covariance–analytical methods one should be careful not to take up too many covariates in the model, since they decrease the degrees of freedom in the final analysis of variance. They may in this way decrease test power, even though they might not filter out any undesirable variance from the posttest scores.

It can be read from this correlation table that none but one of the writing or reading pretests correlates significantly with the posttest that supposedly measures the same construct (only indicator with posttest). Only the pre– and posttest measures of indicator R3 (single) correlate, which is not surprising because they are paralleltests. The other theoretically related pre– and posttests do apparently not measure the same construct. A possible explanation is that the knowledge that students gained during the course is really very new to them, so that they may acquire a very new behavior in coping with argumentative texts, as compared to the way they wrote and read before the experiment. It must be anyway noticed that most of the reading and writing pretests cannot function as covariates in the analysis of posttest data.

In the upper half of this table (and specidally in the top left hand and bottom right hand of this upper half) one can see that the average intercorrelation of posttest measures *within* each of the modi (both reading and writing) is much higher than the average intercorrelation *between* these modi (in the bottom left hand corner). The consequence is that the five indicators for each dependent variable may not be considered independent. For this reason we will use multivariate analysis of variance (MANOVA), a technique that can take the phenomenon into account of dependent variables influencing each other. Since our design consists of only one independent variable (type of practice), we will perform a multivariate one–way analysis of variance.

The lower half of the table is used to select the covariates that can be included in the analysis. Note that the first IQ–pretest does not correlate with any posttest variable, while the second and specially the third do, with two and with seven posttests respectively. Covariates will only be included in the analysis of posttest measures with which they significantly correlate.

The mean pretest scores for the conditions do not show differences that are strong enough to decide that the groups differ on initial skill (see Appendix 5.4). This could be expected, because the students were randomly assigned to the conditions. An analysis of variance for each indicator also yielded no significant between–group differences, so we may conclude that the groups are comparable regarding iq, and prior learning and writing skill concerning argumentation.

We report on the intercorrelations of the pretests in order to see if some of them can be expected to make up a scale.

[insert table 5.6 here]

Table 5.6: *Correlations between pretest scores*

Several conclusions can be drawn from this table. In the first place, the three IQ subtests do not measure the same components of intelligence. IQ3 (verbal analogies) is obviously an outsider, while IQ1 (logical operators) and IQ2 (word list) share some variance, although not enough to join them (their intercorrelation is lower than half of each pretest's reliability, which is in itself a kind of generalized split–half). Each of these measures may be independently added as covariates to the analyses of posttest data, in as far as they share significant variance with these tests.

The high correlation between IQ3 and PR1 (the identification of argumentative relations) is remarkable. A more general skill like 'the identification of abstract semantic relations' may underlie the strong relation between the two tests. However, this must remain speculative at the moment.

5.4.2 Posttests data analysis: Effects on learning and transfer

Table 5.7 contains the mean posttest scores and standard deviations for each condition:

[insert table 5.7]

Table 5.7: *Means and standard deviations for posttest scores across conditions*

In contrast to the pretest scores, there is much between–group variance among the mean posttest scores. The groups apparently differ on most of the measures. On the other hand, the within–group variance is considerable in comparison to the difference in mean scores. Therefore we must determine if the reported between–group differences can be generalized, which can be done by means of the hypothesis–testing MANOVA procedure. We will test the null hypotheses in multivariate procedures because the dependent variabele 'writing skill' is made up of several correlating variables (the indicators). Separate testing would camouflage this correlation.

Learning effects

We will start with answering the research questions regarding learning effects, that is: the effect on writing skill. In table 5.8, the results of the MANOVA procedures on the posttest data are shown.

[insert table 5.8 here]

Table 5.8: *MANOVA tests for between–group differences on the dependent variable 'learning to write'.*

It is important to note that in the following hypothesis testing procedures, there is more than one dependent variable in each hypothesis. All group scores on all indicators are taken simultaneously in the analysis as dependent variables, taking their intercorrelations into account.

Research question Q1 asks for the comparison of conditions OW versus DW: is **learning–by–observing** students who are doing writing exercises more effective than **learning–by–doing** these exercises? The answer, in the upper line in table 5.8, is that the mean score of the observation condition on the writing posttest is significantly higher than the mean of the individual practice conditions. This means that in the writing mode, learning–by–observing of models turns out to be more effective, which was expected.

In condition FW, the observing student is personally involved in the communication: first the student performs a communicative act (writing a short text) and then observes the communicative consequences immediately afterwards. The transfer effect that this 'reader observation' may yield, is the motive for raising research question Q2: is **learning–by–observing** a communicative partner after writing more effective than **learning–by–doing** writing exercises? From the table (second row) it appears that at least for the writing mode, the combination of exercises with observation of authentic feedback is more successful than the standard working method (doing exercises only). Note that the number and content of the tasks, and the time–on–task has been kept constant across the conditions.

In research question Q3, learning–by–doing is compared to observations of both writer and reader: a complete communicative transaction. A MANOVA (on line 3 in table 5.8) shows the effect of the 'observing both roles'–observations in comparison to the learning–by–doing approach. No experimental effect can be found; the groups' scores were not significatnly different with respect to the acquisition of writing skill. This is noteworthy, because subjects in the OWR condition have had 50 % less practice on writers' tasks since they were forced to invest time in the observation of readers' tasks. Nevertheless, they attained a level that almost surpasses that of DW subjects (mind the low p–value of this test). Our expectation concerning this difference was neutral: Sonnenschein & Whitehurst found no differential effect either.

Research question Q4 asks for a direct comparison of the OW and OWR conditions, because it is important to know if observation of two roles enhances or impedes learning that would otherwise take place if only one role was observed. The MANOVA procedure (fourth row) shows that there is no significant difference in learning–to–write between the groups. This is in agreement with our expectations as well.

In sum, when it comes to learning to write argumentative texts, the two experimental learning activities OW and FW appear to offer an advantage compared to the standard condition DW. Thus, learning to write this text type is best served with either the 'observation–of models' (one mode) method of instruction, or with the 'observation as feedback' supplied by readers. The OWR condition, although it is only half devoted to writing (observation), attains a skill level equal to the DW condition.

Transfer effects

The two 'distant observation' modes OW and OWR can together make up an experimental successor to the Sonnenschein & Whitehurst (1983, 1984) series of experiments described earlier. They found that transfer from the productive mode to the receptive mode (or vice versa) was hard to induce, and that only participants placed in an observer's position (observing and evaluating other students' task behavior) were able to transfer their observation experiences to the performance of new tasks in both modes. We will see if this finding can be repeated with another age group, and quite a different task, which does not include the external feedback which was necessary for Sonnenschein & Whitehurst to make the intervention effective.

Transfer effects are also expected from the 'engaged observation' group FW, because of the precise match between 'performing in the one mode' and 'observing and evaluating in the other mode'. If transfer can be obtained by translating 'encoding rules' into 'decoding rules' or the other way around – as discussed in chapter 3 – the FW learning activities would make a good chance to stimulate such transfer. It is important to note that neither of the groups had been told in advance that the posttest would include writing as well as reading tasks for every participant.

There are two types of information regarding transfer that we want to offer: between–group differences in the extent to which transfer is promoted, and a quantification of the transfer effects of each condition, using a standardized transfer measure such as announced in section 5.2. Together these types of information can answer the questions: which condition promotes transfer more than others – and how large is the advantage gained by the transfer?

In this chapter only information from this first category will be given. It will be determined if there are differences between experimental groups in their performance on the transfer tasks. These tasks are administered at the end of their respective writing courses. We will not yet give a quantitative overview over transfer effects of the various conditions. The reason is that in order to quantify transfer using the measure presented in section 2.6 and section 5.2, we would need a new measurement (the criterion, as explained in section 2.6) that will be presented in the following reported experiment (chapter 6). Therefore the reader will have to wait until the end of Chapter 6 for a presentation and comparison of transfer effect sizes.

For the present purpose we used MANOVA procedures again to determine if between–group differences exist on the transfer task. The only difference is that the writers were assessed on their posttest reading skill. It was assumed that between–group differences regarding the reading task can be considered as transfer effects of the experimental interventions. The reading scores are in table 5.7, the results of the MANOVA in table 5.9.

[insert table 5.9 here]

Table 5.9: *MANOVA tests for between–group differences on the dependent variable 'transfer to reading'.*

Question Q7 is the transfer–oriented pendant of Q1: does learning–by–observing students doing writing exercises lead to higher transfer to the opposite mode than learning–by–doing these exercises? Although some counterevidence can be found in Sonnenschein & Whitehurst's studies (see section 3.5, table 3.2), it has been suggested that observation may have a beneficial effect on transfer because the quality of knowledge construction and acquisition may be higher. In table 5.9 we see that OW subjects outscore DW subjects with respect to reading skill, so the OW learning activities have apparently prepared them better. Thus, question Q7 must be answered positive.

The last research question, Q8, addresses transfer effects of observing one's communicative partner. This means here: observation of readers instead of writers. Much appears to be gained from this observation of real readers; not only did it afflict the writers' writing skill (Q2) but it also enhances their reading as can be seen in the second line of table 5.9. Our expectation was confirmed.

No transfer effects have been assessed or presented of the OWR condition. The decision not to do this may need some clarification. The OWR condition is the only condition in which practice is not limited to one mode. Since only complete communication transactions are observed, the OWR students observed exactly as many writers as readers. It was the observer's aim to collect information about good and weak writing and reading processes. According to our definitions of 'learning' and 'transfer', we must consider the increased reading or writing skill resulting from these observations as 'learning gains', with possibly some mutual influences between the modes that qualify as transfer effects. This OWR condition can principally not yield any transfer gains in the narrow sense, since all effects may result from learning that was deliberately aimed at both modes.

It is, however, important to compare the OWR score on the reading posttest to the reading score of the DW condition. As we have seen, their scores on the writing posttest do not differ significantly. A difference on the reading posttest would be an important argument to favour one condition over the other. A MANOVA using the reading posttests as dependent variable and condition (DW vs. OWR) as independent variable) yields a large significant result (table 5.9) in favour of the OWR group. We conclude that observation of complete communications is equally effective regarding writing as the traditional learning–by–doing approach DW, and is more effective regarding reading. Thus, the 'hidden strength' of this instructional method lies in the learner's ability to adapt the knowledge in communicative complementary situations: both as a writer and as a reader.

In sum, we found that transfer from writing practice to reading skill was promoted more by two types of learning–by–observation than by learning–by–doing activities, and that also the observation of both modes is in this respect advantageous. At this point we can only establish that *more* transfer takes place; in order to establish *how much more* we must use a kind of quantification which 1. enables incorporation of the five indicators into one construct 'writing skill' or 'reading skill', and which 2. is informative in that it expresses the achieved amount of transfer in relation to some meaningful criterion (see 5.2). Such criterion will be acquired in the next chapter. The quantification of transfer scores in this experiment and in the next experiment will be reported there.

5.5 Summary and discussion

SUMMARY

The aim of this experiment is to test a theory about effective activities for learning to write, in this case learning to write argumentative text. It was expected that two types of activities, which are both instances of 'observational learning', would be effective regarding learning to write as well as regarding transfer to reading. The rationale for the learning activities has been presented in chapters 2 and 3 of this study.

The expectations were tested by setting up an experiment, using a full–between pretest–posttest design, in which four groups of thirty ninth–grade secondary school students took part. The four treatments consisted of short experimental courses aimed at learning to write argumentative text. The presented subject–matter was the same for each group, but the learning activities varied systematically: doing writing exercises, observing writers, observing both writers and their readers, and doing a writing exercise and observing a reader as feedback. After a pretest session and four one–hour training sessions, the same set of posttests measuring reading and writing skill were administered to all participants.

A multivariate analysis of variance was used in order to test the hypotheses regarding learning effects, using 'intervention' as an independent variable and a set of five indicators for writing skill as complex dependent variable. Hypotheses regarding transfer effects were tested in the same way, but with a set of five indicators for reading skill as dependent variable in the MANOVA.

The main findings are that both types of learning–by–observation (observation–as–model and observation–as–feedback) are more effective (reagarding learning output) than learning–by–doing (research questions Q1 and Q2). Two variants of observation–of–models were distinguished: observation of writers only (who create a text for a virtual reader), and observation of complete communicative transactions (both writers and their readers). It was found that observation of both writers and readers did not differ from the effectivity of learning–by–doing writing exercises (Q3), but that it adds to this an important extra effect: a strong effect (in fact: a mixed learning and transfer effect) on reading. Moreover, the effectivity of observing complete reading–writing processes does not significantly differ from observing writing processes only (Q4). Finally, the transfer effects of observation–of–models (one mode) and observation–as–feedback are stronger than the transfer effects of learning–by–doing (Q7 and Q8).

Part of this experiment can be seen as a replication of Sonnenschein & Whitehurst (1984), since they also found transfer of observing/commenting on tasks to performing the tasks themselves. In their study, however, not much transfer is found from observation of one role only to performing in the complementary mode (their conditions 4 and 5), while such transfer is found here. It is not sanctioned by receiving oral feedback by a teacher, such as was the case with Sonnenschein & Whitehurst. This may be due to the very different group of learners and a totally different subject.

A second main finding is that some between–mode transfer can be observed in all of the conditions, although transfer is by far strongest in the observation conditions. This is most likely due to the many clues that the students have to link learning and application situation. They have situational clues (a new task, but in the same time & place & with the same 'teacher' & materials) and clues regarding content (the same concepts & terminology). It would indeed be very strange if the various learning activities did not transfer to reading at all. On the other hand, it is difficult to assess the amount of transfer from learning–by–doing, since we have no valid assessment of a bottom–line performance (pretests measures did not correspond with postttest measures).

In sum, the experimental effects support the theoretical model–variant of learning–by–observation (Sonnenschein & Whitehurst) as well as the feedback–variant (Schriver). The learning–by–observing approach, in the two shapes that we have put it in, can be very effective. When applied to communication skill training, it can also be quite efficient because the observations appeal to the transferability of the skill. Besides, the similarity here betweem the observandum (comprehension and construction processes) and the act of observing itself (a process of comprehension and mental construction) may be a key factor in its effectivity.

Explanations for the experimental effects were given in the analysis of cognitive activities performed in the various conditions (table 5.4). The more succesful conditions, OW and FW, were different from learning–by–doing in the executive and evaluative activities. The OW task was convergent ('decide which observed reader is better') as opposed to the divergent DW task ('write an argument on...'). By making comparisons, OW subjects could mould and sharpen their image of what 'good task strategies' are. On the other hand, for FW subjects the divergent writing tasks was changed into a more convergent task: their writing was connected with the consecutive comprehension tasks. FW sujbects purposefully wrote in such a manner that the comprehension task would be successful too. They are more personally involved, and may therefore have a stronger motivation They also have to replace themselves more often in the reader's position. Some students, on the other hand, may prefer to the more 'neutral' learning–by–observation of models.

VALIDITY

When an effect–study appears to favour the experimental conditions, it is time for the researcher to pucker his brows and check for factors that may threaten the validity of the experiment. The experiment has been designed in such a way that several alternative explanations have been ruled out. The experimental groups can be considered comparable, the time–on–task was almost similar for all groups, there is no influence of teachers, research assistants etc. since all courses are self–instructive, the students were motivated because they would receive a small reward, the treatments and the tests correspond equally well for every condition because the theory was the same for everyone.

However, criticism to the validity of the results is of course possible. For instance, an important difference between the learning–by–doing and the learning–by–observation conditions is, that the former are very familiar for the student and the latter not. It may be that the novelty of observations, the use of video, the observation of 'live' models, has interested the participants from these conditions to such extent that they worked harder and were more personally involved, which can account for part of the experimental effects. On the other hand, research assistants from all conditions have noticed both enthousiast and tedious reactions of students doing the tests or working on their workbooks. Tediousness was not necessarily greater in the learning–by–doing conditions – although it is difficult to check this. An indication may be that the number of not completed workbooks or not completed tests (a possible symptom of disinterest) does not vary across the groups. However, it must be added that a more attentive working attitude may be specific for learning activities that call for special attention, such as observations and evaluations.

Due to organizational requirements, the workings conditions were not equal for all conditions. Subjects in the learning–by–doing group worked individually, while seated in a large room with 3 to 8 people at a table, leaving more than enough space to work.. It was not allowed to co–operate or to conversate during the lessons. Subjects in the 'model' condition, who had to use a videoset, were seated in a middle–size room with a table for themselves. Only six persons were at the same time in the room. Subjects in the 'feedback' condition worked in a large room, with only 2 writers, the proof–reader, and the research assistant present. If group size influences performance, then this worked to the advantage of the feedback condition. On the other hand, subjects in this condition had to cope with more organizational problems (walking to and from the proof–reader, keeping a very strict time schedule).

There are also some weaknesses in the experimental and statistical design of the study. In the first place, the pretest–posttest design is not genuine, since the pre– and posttests are not equavalent. We had to use different pre– and posttests, because two quite different levels of mastery had to be reliably measured without bottom– or ceiling effects. Thus pre– and posttests were globally aimed at the same skills (aspects of reading and writing argumentative text), but different subskills may be assessed.

This is related to the problem of the covariates. Pretests were included in the design to enable covariance analysis which would filter out undesirable effects in the posttest measurements. However, the majority of the pretests did not correlate with posttests that were aimed at the same construct. It is uncertain what the pretests , which in themselves have a sufficient homogenity, have measured. It is anyway unwise to use the non–correlating pretests as covariates, so we have left them out.

There are some threats to the external validity as well. Due to the organization of the experiment, the posttests were administered almost immediately after the training had taken place. We can therefore not be certain about the durability of the results. Although durability is an important feature in real educational settings, the development of long–lasting skills for the students was not given highest priority. Highest priority was given to answering the research questions – about the effectivity of learning activities – under experimental control. Problems arise if the results of learning–by–observation appear to be less durable than results from learning–by–doing. However, we have no reason to assume such differences in durability.

The use of the word 'reading' and 'writing' is problematic. In the course of this report the word 'analyzing' is sometimes used as a synonym for 'reading'. Of course these terms are no true synonyms, because the concept of reading contains much more than only analytical activities, just like writing is much more than combining sentences and inserting conjunctions. On the other hand, the kind of 'reading' and 'writing' that is taught at school is also more technical/analytical than students would do in their leisure time. In process–oriented language education, analysis of reading and writing tasks is an important activity, because one or more qualities of the process or product must be demonstrated and/or practiced.

Although in this experiment learning–by–observation turns out to be more effective, we would not use the results to discredit learning–by–doing in general. We believe that learning–by–doing remains indispensable in, and essential to language skill education, but that it is not the 'only true instructional method'. It would not surprise us if the great majority of the learning activities that students must do today fall into this category: the individual execution of language processes which have been explained by someone else (a teacher or a book) and which will be evaluated by someone else too. In the present study, we have only tested this type of learning in contrast to three types of observation, and as a new field of application we chose argumentative text, which had to be read and written by students for whom the subject was comparably new and difficult. In such a situation, the students can be considered in need of good examples (models) who demonstrate what the behavior–to–be–acquired is like, along with examples demonstrating the pitfalls to avoid; pitfalls they are likely to make since they are novices to the task.

In this situation, learning–by–observation showed to be advantageous. However, once a basic cognitive level of knowledge and skill has been acquired, the need to proceduralize and flexibilize arises (Salomon & Perkins, 1989; Anderson, 1990). This calls again for learning–by–doing activities. These activities can now profit from the observation experience because criteria for self–evaluation have become more explicit.

In the end, we expect most of a well–balanced interplay of learning–by–observation and learning–by–doing activities. Each of these instructional methods has its qualities and drawbacks. It is up to the educator to compose learning programs in which the qualities are combined and the drawbacks compensated. That the qualities of learning–by–observation deserve to be studied in close detail, is what we hope to have demonstrated.

Figure 5.1: *Structure of independent variables in this experiment*

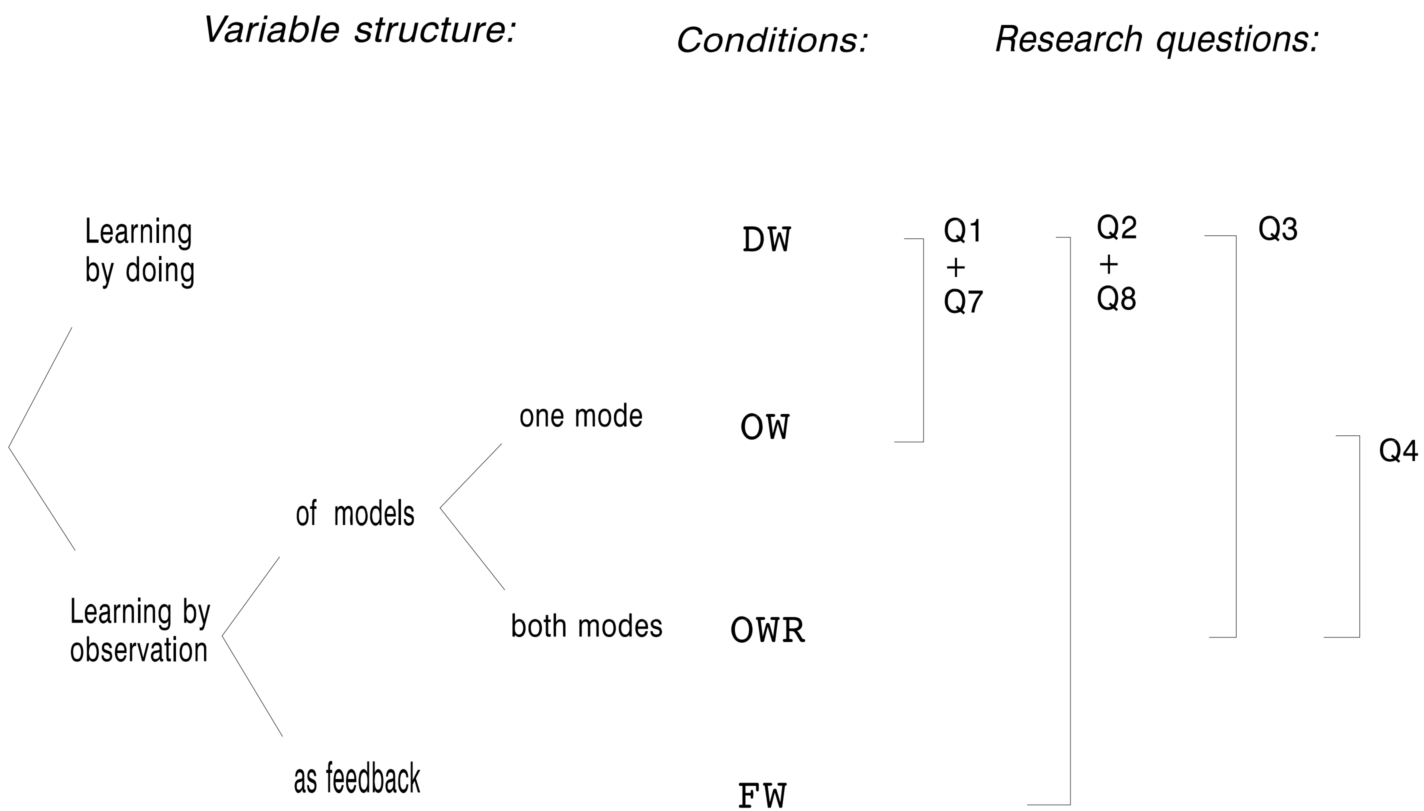


Figure 5.2: *Instructional sequence of the courses*

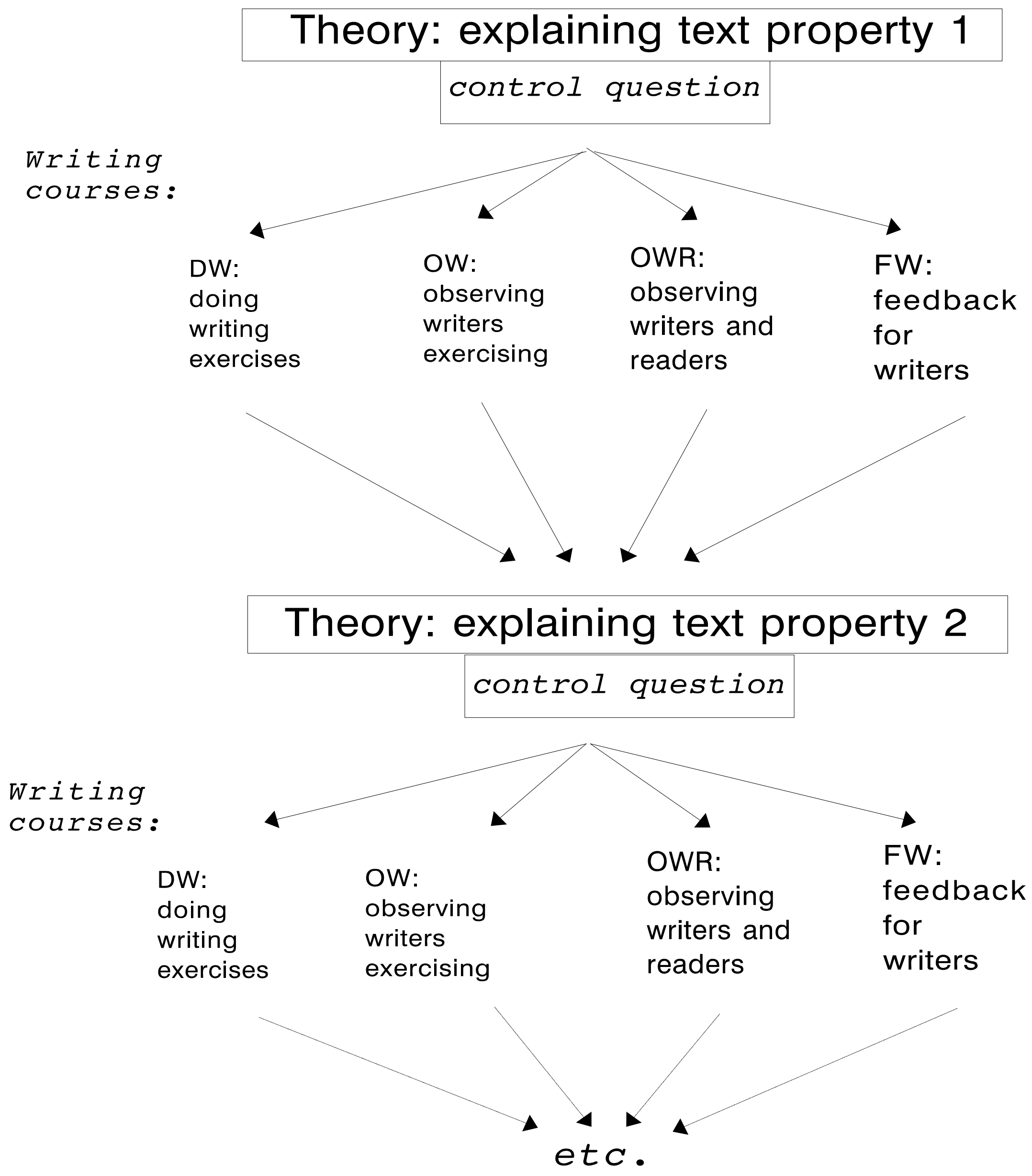


Table 5.2: *Learning objectives for the short courses in 'argumentative texts'*

Writing skill		**Reading skill**
Expliciting the social parameters of the discussion: issue, parties, communicative goal	*lesson 1*.	Recognition of the social parameters of the discussion in the text
Composing a well–strucured text on the basis of a model	*lesson 2*.	Analyzing the text structure on the basis of a model
Writing on the basis of simple and complex argumentation structures	*lesson 3*.	Analyzing complex argumentation structures and its simpler constituents
Applying various means for presentation of text elements	*lesson 4*.	Identification of means for presentation of text elements

Table 5.3: *Theoretical contents of the course on 'Argumentative texts'*

Lesson 1: 'Argumentative texts and discussions'

Introduction of five main concepts:

- *standpoint* (opinion): positive, negative, neutral
- *argument* (reason)
- *argumentative text*
- *issue*
- *discussion*

In an inductive fashion, the concept of argumentative texts is explained by means of its constituting elements: standpoints and arguments (opinions and reasons for having these opinions). The genre of argumentative texts is placed in the social context of discussions aimed at resolving a dispute, which centers around the acceptability of a certain proposition (the 'issue'). (cf. Van Eemeren & Grootendorst 1983, ch. 1 & 2).

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 of counter–argumentation
- *Ending* : conclusion; most important arguments; consequence.

The well–known global text structure consisting of introduction, body and ending is specified for argumentative texts. The function of each subpart is discussed in relation to the discussion goal. Various examples help to give meaning to the concepts.

Lesson 3: 'The argumentation in argumentative texts'

Presentation and discussion of several types of argumentation:

1. *singular argumentation*
2. *compound argumentation*
3. *subordinate argumentation*

and the *complex argumentative structures* of which these are the constituents. Moreover, a simple notation system for schematization of complex structures is taught. (cf. Van Eemeren & Grootendorst 1983, ch. 9 & 11)

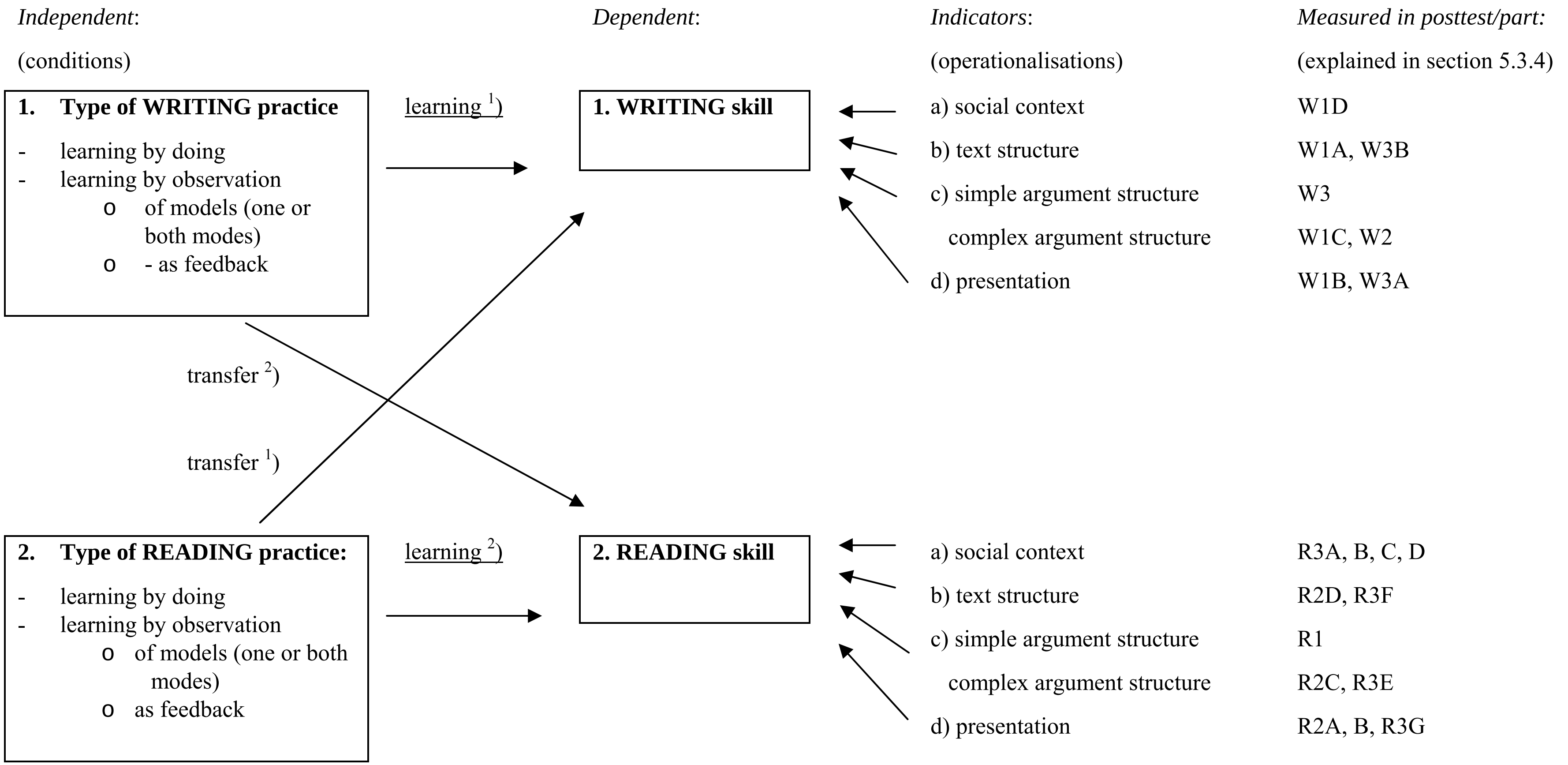
Lesson 4: 'The presentation of argumentative texts'

Presentation and discussion of three means for the clarification of the text structure

1. *paragraphing* and the rhetorical model
2. using verbal *structure markers*
3. argumentative *connectors*

It is demonstrated how each of these three means is helpful in recognizing or expressing the global text structure, the parts that make up this structure, or the complex structure of the pro–argumentation.

Figure 5.3: *Variable structure of experiments reported in chapters 5 and 6*



[1]) covariates IQ (pre1, 2, 3) and Writing skill (pre-S1, pre-S2)

[2]) covariates IQ (pre1, 2, 3) and Reading skill (pre-L1, pre-L2)

Appendix 5.3: *Pretest–posttest correlation table*

Posttests:	W1	W2	W3–1	W3–2	W4	R1	R2	R3A	R3B	R4
Posttests:										
W1 (writ)										
W2	.1017									
W3–1	.4556**	–.2602*								
W3–2	-.0054	.0763	-.0741							
W4	.2996**	.0734	.4283**	.0383						
R1 (read)	.1183	-.0227	-.0496	.0849	.1770*					
R2	.0890	.1780	.0166	.1411	.1909*	.4523**				
R3A	.2166*	.0191	.1133	.1226	.1549	.0584	.1925*			
R3B	.0377	.0342	.0733	.1479	.2029*	.4269**	.6523**	.1560		
R4	.1036	.0826	-.0167	.1117	.1137	.4297**	.6342**	.2199*	.5773**	
Pretests:										
IQ1 (iq)	.0147	-.0751	.1010	-.0972	.1223	.0963	.1994	-.0141	.1724	.1470
IQ2	-.1163	-.0250	-.0392	.0231	.0296	.1271	.2423**	.0611	.2905**	.1355
IQ3	.2420**	.0277	.1432	.3670**	.2108*	.1541	.4268**	.4514**	.2995**	.4311**
1A (writ)	.0968	-.0632	.1291	.1347	.1579	.0717	.0649	.0374	.1518	.0494
1B	.1293	.2102	.1130	-.0496	.0775	.0501	.0906	-.0109	.0978	-.0211
1D	.0672	.0839	-.1044	-.0464	.0911	.1292	.0350	-.0181	.1113	.0147
2Ca (read)	.1897	-.0064	.1601	.3275**	.2192*	.1918*	.4310**	.4495**	.3278**	.3369**
2Cb	.0345	.0805	–.0988	.1608	–.0314	–.0342	.0205	.0761	–.0155	.0692
# of covariates:	1	0	0	2	2	1	3	2	3	2

* = $p < 0.01$ ** = $p < 0.001$

Appendix 5.4: *Means and standard deviations for pretest scores across conditions*

	Pretests for IQ:						Pretests for READING			
CONDITION:	IQ1		IQ2		IQ3		PL1		PL2	
	CMR–test	*CMU–test*		*CMR–test*		*argu. simple*	*argu. complex*			
Learning by Doing Exercises	19.53	8.92	46.50	7.78	22.51	10.19	22.10	6.15	5.55	2.77
Learning by Observation (1 mode)	20.60	7.22	50.55	7.82	23.31	10.01	22.03	4.72	6.06	3.18
Learning by Observation (2 modes)	20.79	7.66	52.31	11.33	25.13	10.86	23.13	6.03	4.58	2.89
Learning by Observation as Feedback	18.82	6.93	52.55	9.71	25.51	10.17	23.62	5.57	4.00	2.29
Max. score:	40		75		50		33		20	

	Pretests for WRITING:					
CONDITION:	W1		W2		W4	
	soc.context		*text structure*		*presentation*	
Learning by Doing Exercises	2.82	2.22	9.30	5.45	3.82	1.76
Learning by Observation (1 mode)	4.92	1.52	9.81	3.31	5.03	2.13
Learning by Observation (2 modes)	3.75	1.70	9.89	3.46	4.62	2.37
Learning by Observation as Feedback	3.85	1.26	8.13	3.53	4.20	2.57
Max. score:	8		18		6	

Table 5.6: *Correlations between pretest scores*

(n=119) (*= p<0.01, **= p<0.001)

Corr.	IQ1	IQ2	IQ3	PW1	PW2	PW3	PW4	PR1
IQ1	1.0000							
IQ2	.2735**	1.0000						
IQ3	.0235	.1510	1.0000					
PW1	.0638	-.0027	.1320	1.0000				
PW2	.0447	.0489	.0160	.1316	1.0000			
PW3	-.0217	-.0367	-.0092	.3168**	-.2763**	1.0000		
PW4	.0347	.0138	-.0180	.3685**	.0033	.3697**	1.0000	
PR1	.0485	.1980*	.9244**	.1298	.0381	-.0399	.0131	1.0000
PR2	-.0781	.0130	.0172	.0390	-.0631	-.0152	.0595	-.0568

Table 5.7: *Means and standard deviations for posttest scores across conditions*

Posttests for WRITING:

CONDITION:	n	W1 *soc.context*		W2 *text structure*		W3–1 *argu. simple*		W3–2 *argu. complex*		W4 *presentation*	
DW: Learning by Doing Writing Exercises	29	2.51	(2.11)	7.51	(5.82)	21.51	(8.52)	16.55	(7.17)	4.13	(2.85)
OW: Learning by Observation (1 mode)	30	4.13	(2.32)	9.34	(5.63)	27.44	(8.06)	22.65	(5.82)	8.89	(3.53)
OWR: Learning by Observation (2 modes)	30	3.58	(2.35)	8.75	(5.16)	26.03	(8.05)	20.27	(7.92)	5.65	(3.65)
FW: Learning by Observation as Feedback	30	3.79	(2.09)	8.51	(4.68)	26.55	(7.51)	22.55	(6.83)	8.72	(3.82)
Max. score:		6		14		30		32		12	

Posttests for READING:

CONDITION:	R1 *soc.context*		R2 *text structure*		R3–1 *argu. simple*		R3–2 *argu. complex*		R4 *presentation*	
DW: Learning by Doing Exercises	5.72	(3.25)	8.97	(5.69)	23.32	(8.15)	8.58	(5.50)	5.36	(1.82)
OW: Learning by Observation (1 mode)	8.06	(2.64)	17.26	(4.59)	27.65	(10.05)	24.27	(4.62)	12.08	(2.17)
OWR: Learning by Observation (2 modes)	8.75	(3.18)	15.54	(5.05)	28.62	(9.74)	21.44	(5.43)	12.54	(2.80)
FW: Learning by Observation as Feedback	8.72	(2.21)	18.10	(4.84)	29.34	(8.75)	24.84	(4.39)	12.26	(1.96)
Max. score:	12		22		43		34		18	

Table 5.8: MANOVA tests for between–group differences on the dependent variable 'learning to write'. In the statistical design, all five indicators are included in the construct 'writing skill'.

question	conditions	posttest	covariates	n	F	∝	testing:
Q1	DW – OW	Writing	IQ3, PR1	59	2.67	< 0.04 **	one–sided
Q2	DW – FW	Writing	IQ3, PR1	59	2.93	< 0.02 **	one–sided
Q3	DW – OWR	Writing	IQ3, PR1	59	1.76	0.08 *	two–sided
Q4	OW–OWR	Writing	IQ3, PR!	60	1.13	0.17	two–sided

** = significant at $\alpha = 0.01$
* = significant at $\alpha = 0.05$

Table 5.9: MANOVA tests for between–group differences on the dependent variable 'transfer to reading'. The design includes all five indicators of the construct 'reading skill'.

question	conditions	posttest	covariates	n	F	∝	testing:
Q7	DW – OW	Reading	IQ2, IQ3, PR1	59	3.34	< 0.01 *	one–sided
Q8	DW – FW	Reading	IQ2, IQ3, PR1	59	3.12	< 0.02 *	one–sided
– –	DW – OWR	Reading	IQ2, IQ3, FR1	59	3.64	< 0.01 *	two–sided

* = significant at α = 0.05

CHAPTER 6: Observing readers and writers of argumentative text: Learning to read – transfer to writing

The reported experiment is a replication of the experiment in chapter 5, with the difference that the learning activities are aimed at reading instead of at writing. The effectiveness of three variants of learning–by–observation is compared to learning–by–doing. The learning activities are aimed at learning how to read argumentative texts. The variants are: observing readers as a model, observing writers & readers as a model, and observing writers as feedback on one's own reading performance.

Participants are ninth–grade secondary students, who followed one of four short experimental courses in reading and analyzing argumentative text. Observations are made by means of authentic video–tape recordings (model condition) or by 'live' confrontations (feedback condition). The participants are pretested and posttested on writing and reading skill related to argumentative text.

Results show that learning–by–observation activities lead to better performance and to more transfer. Transfer of the learning–by–doing method of instruction is comparably low. Transfer from writing to reading, as obtained by the observation method, is stronger than transfer from reading to writing. The findings confirm the effectivity of monitoring– and evaluation–oriented learning activities, and support an interactive, unsymmetrical conception of the reading–writing relationship.

6.1 Introduction

The previous chapter was concerned with learning to write, and with transfer from learning–to–write to reading performance. It was found that learning effects as well as transfer effects were enhanced by certain observation activities as part of the learning process. The results yielded threefold support: support for a learning theory applied to the writing domain which favours observation activities over application activities (see section 3.3); support for a task–oriented transfer theory that permits transfer from the productive mode to the receptive mode (see section 3.2); and support for a transfer theory stressing the importance of instructional factors: the occurrence and amount of transfer appeared to be dependent on the type of cognitive activities elicited in the learning situation (see section 2.5).

In this chapter, the same theoretical issues are addressed with respect to the complementary domain: reading. The same research questions are addressed in an experimental setting that closely resembles the one described in chapter 5. Also the same population is chosen to test the hypotheses. Thus, the experiment reported in this chapter might be considered a replication study in order to test the theory in a different, but related domain. Moreover, the experiment supplements the one reported in the previous chapter in an attempt to collect evidence for an interactive or bilateral model of the reading–writing relationship (or, more precisely, for rejection of unilateral models; see section 3.2). If the same learning activities that foster transfer from writing to reading, will yield transfer from reading to writing, the interactive model is supported. If the learning activities result in the absence of transfer, or a substantial difference in the amount of transfer obtained from reading to writing compared to writing to reading, the two skills integrated in the model are apparently not equivalent mirror–images.

In section 3.2 it was discussed how linguistic knowledge specifying the relationship between verbal forms and meanings can be integrated in language production rules as well as in language reception rules. It was suggested that transfer could develop in two ways: the linguistic knowledge is initially learned at an abstract level, and is transformed to one or more mode–specific rules (the top–down order; the learner starts as it were on the 'high road'), or the learner can acquire a coding/decoding rule in a mode–specific manner first, and mindfully abstract this knowledge later, enabling application in another mode (the bottom–up order, starting from the 'low road'). This theoretical perspective allows for differences in the likeliness of transfer between skills. Is top–down or bottom–up translation easier in the writing mode or in the reading mode? For instance, if a student has learned that enumerations in a text like 'first', 'second' etc. are used to distinguish arguments, is it more likely that he transfers this knowledge to reading when it was trained in writing, than to transfer it to writing when trained in reading? We have no data or specific expectations on such unequivalence of transfer to reading and writing, but such knowledge is evidently useful for the development of effective instruction. Thus, I will compare the amounts of transfer obtained from learning–to–write to reading performance and from learning–to–read to writing performance.

The present experiment will address the research questions Q1–Q4, Q7 and Q8 with respect to learning–to–read argumentative texts. Moreover, questions Q5 and Q9 (about the effect of reinforced self–reflection on learning and transfer) will be addressed since a variant of the learning–by–doing method of instruction will be added, which is explained in this section. Finally, questions Q6 and Q10 can be answered since the effects can now be compared of the task factor 'domain' (writing vs. reading) on learning and transfer obtained by the various instructional methods.

The instructional methods which are experimentally put to the test are the standard method, learning–by–doing, opposed to the variants of learning–by–observation (feedback and model). The operationalizations will be explained in the next section.

We should pay attention here to one extra condition that is added to the design. In the previous experiment, subjects in the observation–as–feedback condition learned to write by observation of their 'own' readers, who read and analyzed the text aloud and commented afterwards on its comprehensibility. Those 'readers' were not involved as a group in that experiment, since the activities that they undertook were not aimed at learning–to–write. Their activities (doing reading exercises while thinking aloud and commenting on the comprehensibility of the texts) were part of a learning–to–read condition, which can be called a special case of learning–by–doing. These activities are comparable to the readers in the experiment reported in chapter 4, who read and commented on the manual texts for the physics experiment. The thinking–aloud and evaluation activities they perform may stimulate self–reflection. Therefore this group or readers is denominated as the condition 'learning–by–doing with reinforced self–reflection' and will in this experiment be compared to the condition 'learning–by–doing with normal (that is: unprompted) self–reflection'. Both groups do the same type, and almost the same number of reading exercises. The 'self–reflection' group, however, is thinking–aloud during performance and answers an evaluative questions directly after.

One word about the use of the term 'reading' in this chapter. 'Reading' stands for 'analytical reading'. The type of reading that is developed by the experimental minicourses is not mere technical reading or reading–for–understanding; but it is reading aimed at text decomposition: recognizing the essential elements that the text is made of and understanding their interrelations. The reader of this study ought to keep in mind that the terms 'reading', 'reading activities' or 'learning–to–read' in this chapter refer to such text–analytical activities.

6.2 Research questions and expectations

The following graph represents the structural relations between the independent variables in the present experiment:

[insert figure 6.1 about here]

The main research questions as presented at the end of chapter 3 are aimed at comparing the effectiveness of learning–by–doing and of two types of learning–by–observation. The present experiment includes all these instructional methods, so the following research questions are addressed:

Regarding learning effects (on reading):

Q 1: Is learning–by–observation of models (one mode) more effective than learning by doing?

Q 2: Is learning–by–observation–as–feedback more effective than learning by doing?

→ (Expectations: both types of learning–by–observation are more effective).

Q 5: Is learning–by–doing with reinforced self–reflection more effective than learning–by–doing with normal (= unprompted) self–reflection?

→ (Expectation: reinforced self–reflection promotes learning)

Regarding transfer effects (on writing):

Q 7: Does learning–by–observation–of–models promote more transfer than learning by doing?

Q 8: Does learning–by–observation–as–feedback promote more transfer than learning by doing?

→ (Expectations: both types of learning–by–observation promote more transfer).

Q 9: Does learning–by–doing with reinforced self–reflection promote more transfer than learning–

by–doing with normal self–reflection?

→ (Expectation: reinforced self–reflection will promote abstraction and thus transfer)

Regarding a mixed learning- and transfer effect (on reading):

Q 3: Is learning–by–observation–of–models (both modes) more effective than learning–by–doing?.

→ (Expectation: no difference (Sonnenschein & Whitehurst, (1984)).

Q 4: Is learning–by–observation–of–models (both modes) more effective than observation of one mode (model)?.

→ (Expectation: no difference (Sonnenschein & Whitehurst, (1984)).

(Intermodal transfer effects of the condition 'Observation–of–models (both modes)' cannot be distinguished from its learning effects. Therefore no specific question will be asked as to the relative size of these learning effects and these transfer effects.)

Operationalization of the independent variables:

Type of instruction: learning–by–doing is operationalized as doing exercises in reading and analyzing argumentative texts as part of a short course on 'argumentative text'. Learning–by–doing with reinforced self–reflection is operationalized as doing these reading exercises while thinking aloud and commenting on the comprehensibility of the texts. Learning–by–observation–of–models (one mode) is operationalized as observing student readers who think aloud while they perform such reading and analyzing tasks. Learning–by–observation–of–models (both modes) is operationalized as first observing a writer (thinking aloud while performing a writing task) and then observing a reader (thinking aloud while reading and analyzing the writer's text). N.b. this is *exactly* the same operationalization as in the previous experiment, i.e. *not* any 'complementary activity', due to the fact that a complete communicative transfer cannot be mirrored. As a consequence, the same group of subjects and measurements will be added to the data of this experiment. Learning–by–observation–as–feedback is operationalized as reading and analyzing an argumentative text and then observing the original writer on video (who is thinking–aloud while constructing this text) followed by evaluating the adequacy of one's analysis.

This latter condition, learning–by–observation–as–feedback is a bit unusual. The rationality underlying this instructional method holds that writers or readers who observe their communicative partner may detect their partners intentions, needs and thinking activities, may detect flaws in their own performance, and may discover that the linguistic knowledge they use is also used – in a mirror–like way – by their partner. In chapter 5, this condition was operationalized by writing a text and then observing an authentic reader, which leaves the natural temporal sequence untouched. In the present experiment, however, the reader should be allowed insight in the writer's composition process. The reader gets this insight by looking at video recordings of the writer, who is thinking aloud while composing the text that the reader had just analyzed. Subsequently, the reader/observer uses this insight to evaluate his reading performance.

As said, the group of subjects who functioned as readers/commenters in the experiment reported in chapter 5 are added to the design as a special case of learning–by–doing. In this way, Q5 and Q9 were addressed here.

Moreover, by comparing the results of both experiments in chapters 5 and 6, we can address the remaining questions Q6 and Q10:

Q 6: Do the abovementioned questions 1–5 yield different answers for <u>learning to read</u> and <u>learning to write</u>?

→ (Expectation: no difference)

Q10: Do the abovementioned questions 5–10 yield different answers for <u>learning to read</u> and for <u>learning to write</u>?

→ (Expectation: no difference; Sonnenschein & Whitehurst, 1984)

Answering these questions requires that the between–group differences found for each dependent variable are compared between the modes (reading) and (writing). This must be done regarding the learning measures (Q6) and the transfer measures (Q10).

Operationalization of the dependent variables.

There are again two dependent variables: *learning* to read argumentative text, and *transfer* from reading argumentative text to writing such texts. The design underlying the writing and reading experiments contains the same pre– and posttests for all groups. This means that in this experiment, the dependent variables will also be measured with the same indicators (= argumentative pre– and posttests) that were presented in section 5.2. It must be noted that the indicators for 'argumentative reading' will now indicate learning effects, and the indicators for 'argumentative writing' will indicate transfer effects.

6.3. Method

6.3.1 Design

A similar experiment as in chapter 5 was set up in order to test our hypotheses. The research design can be schematized as in table 6.1:

Table 6.1: ***Experimental design for 'Learning to Read argumentative texts'***

Condition:	*n:*	*Pretests:*	*Learning activities:* (4 lessons of 1 hour each)	*Posttests:*
DR	30	X	Doing Reading exercises	X
DRS	15	X	Doing Reading exercises with Reinforced Self–Reflection	X
OR	30	X	Observing Readers as models	X
OWR	30	X	Observing Writers & Readers as models	X
FR	30	X	Observing Writers as Feedback on Reading	X

This is a pretest–posttest control–group design, similar to that presented in table 5.1. The 'learning by doing' group is denominated as control group, and the other groups are experimental groups. Pretest scores will again be taken as covariates in the analysis of posttest data, in order to correct for potential initial differences between subjects.

The reading posttests constitute *intra–modal learning* measurements for the DR, DRS, OR and FR conditions (because training and testing are within the same mode). The writing posttests are *inter–modal transfer* measurements for DR, DRS, OR and FR.

Data for the OWR condition, measuring *intra–modal learning* for both the writing and the reading mode, is taken from exactly the same group of subjects and the same measurements as included in the previous experiment. That is because the operationalization of this instructional method is identical for learning to write and for learning to read. Thus, the data from these subjects are used twice: in the previous experiment and in the present one.

New is the DRS condition (Doing Reading exercises with reinforced Self–Reflection), consisting of the readers/commenters who took part in the previous experiment by supplying feedback to writers (FW condition). These DRS subjects do the same reading exercises as the DR subjects, during all four lessons, while thinking–aloud and comparing the analyzed texts on comprehensibility. Since their learning activity is aimed at learning to read, their group is added to the other groups in the present study. The group size is half the size of the other groups, because each DRS subject served in the previous experiment as a reader/commenter to two FW writers; hence the 1:2 ratio. This enabled DRS subjects to make comparisons between two reading experiences with two different texts. Time–on–task for the exercises was comparable between DR and DRS subjects: for instance, in the same time that DR subjects did five reading exercises, the DRS subjects could do 2 x 2 exercises (two texts written by two FW subjects).

6.3.2 Subjects

In all, 135 students who had just finished the 9th grade (intermediate and high level) took part in the experiment. The average age was 15.6 years. 63 % of the participants were female; boys and girls were almost equally spread across the conditions. The students came from the same 8 schools as before. The same voluntary enlistment and semi–random assignment procedures were applied, on the understanding that the assignments of subjects to the DRS condition had already taken place during the previous experiment. Again, the pretest measurements on IQ, writing and reading skill enable a check on equality of the groups.

6.3.3 Training materials: experimental reading courses on argumentative texts

Experimental minicourses were developed aimed at reading argumentative text, containing exactly the same subject matter as the writing minicourses used in the previous experiment. The exercises, however, were replaced by ones that 1. would provoke the learning–to–read activities typical for a certain condition, and 2. would mirror the complementary exercises in the writing course. For instance, where the writing course has the assignment to write a text containing one standpoint and three compound arguments, the reading course contains an assignment to analyze the argumentation structure of a text containing one standpoint and three compound arguments.

Learning goals and theoretical contents

As in the writing courses, the selection of the theoretical content or subject matter of the lessons was in agreement with the learning goals (see table 5.2). I refer to section 5.3.3.1 for a more thorough description of theoretical content. It is important to notice that no changes to the theory or its presentation were needed to transform the writing course into a reading course. Again, the only differences were the type of exercises. The reading minicourses all consisted of four lessons and were all self–instructive.

Instructional sequence

The theory on argumentative texts forms the backbone of the five different courses that are developed for the experimental conditions. Nevertheless, the subjects spend about 70 % of the time on the exercises in which the theory must be applied. The nature of a course as a learning–by–doing course or a learning–by–observation course is therefore not at all determined by the theory, but only by the type of exercises. Figure 6.2 shows the similarities and differences between the five courses with respect to the instructional sequence of theory and exercises.

[insert figure 6.2 here]

Fig. 6.2: *Instructional sequence of the courses*

Notice that two learning–by–doing courses are on the left hand, and three learning–by–observation courses on the right hand. The subject–matter presented in the courses is identical, but the nature of the exercises differs: the theory must be applied to either reading exercises, or to observation exercises regarding reading argumentative text. (For more explanation on this figure, see section 5.3.3.)

Exercise types

I will describe the differences between the types of exercises by means of the same example ((from the first lesson) that was used in chapter 5, but now applied to the reading mode. Once the definition of 'argumentative text' is given, clarified by examples, and the students may have some notion of the conceptual rule:

> *'S and A are the essential parts/properties of a type B text'*

and after they reproduced these characteristics on paper (control question), they start the first exercise.

DR (learning by doing reading exercises)

The subjects from the DR condition do the following assignment:

DR:

Jackie says to his father: *'We haven't been to the cinema for ages. So I think it is high time that we go and see 'Jurassic Park' with Christmas'*

Question: Is this an argumentative text?

O **NO !**

O **YES, because** I can find the following characteristics of argumentative texts:

opinion: ………...

reason for having this opinion: ………………………………………………………………………………………

The DR subjects must use the rule in a receptive way, in the analysis of given texts. They must check for the occurrence of each characteristic, and then determine if the whole text matches the rule. Answering 'yes' is not enough: in order to check if subjects mix up opinions and arguments, they must write down their analysis.
In doing such an exercise, the abovementioned rule is applied to the receptive mode, which may yield a production aimed at the identification of properties from which class membership is inferred:

< **if** (goal = typify text as type B)
then (actions = (identify characteristic A) **and** (identify characteristic B)) >

In this form, the conceptual rule can be used in identification tasks

OR (observation of reading)

The observation exercises for the OR subjects resemble those of the OW subjects in the previous experiment, with the difference that reading or analyzing processes are observed (such as performed by the DR subjects) instead of writing processes. The observed students think aloud as well. OR subjects received the following exercise:

OR:

Read the following assignment, and imagine how you would answer it:

"Read the following example:

Jackie says to his father: *'We haven't been to the cinema for ages. So I think it is high time that we go and see 'Jurassic Park' with Christmas'*

Question: Is this an argumentative text?
O **NO !**
O **YES, because** I can find the following characteristics of argumentative texts:
opinion:..
reason for having this opinion: ....

You are going to see two students doing this assignment.
It is your task to find out what they do well, and what they do wrong.
When you have observed both students, you may advance to the next page.

(....)

(next page:)
You saw two students doing the reading assignment. Their answers were:

Student 1:	**Student 2:**
'<u>Yes, because</u> *opinion: "That it is time to go to the cinema again"* *reason for opinion: "He has not been there for a long time"*	*'<u>Yes, because</u>* *opinion: "That it was too long ago that he was in the cinema".* *reason for opinion: "That he wants to go at Christmas".*

===>>> Which student did better, according to you? Student .....

===>>> Explain briefly why you think the other student did worse.

Student ..... did *worse*, because ....

OR subjects are also prompted for orientation, and go through the same stages of observation, comparison and evaluation as DR subjects. They are granted only a short time for orientation on the exercise, because their main learning should follow from observation, not from normal practice. It is important to note that the subjects were not allowed to read the written answers by student 1 and 2 *before* the observations had ended. This is because the experimental instructional methods stress the observation of *processes* rather than *products*. If the subjects would have had the final 'solutions' at their disposal before or during the observations, it is well imaginable that they would evaluate these products only and not be patient enough to observe the complete reading processes that yielded these products.

It is the OR subjects' task to evaluate the use of the conceptual rule by the observed readers, who try to analyze the text. In this example it is clear that the second observed student has mixed up standpoint and argument. The observer must detect the differences in the analysis, and can attribute them to differences in working–method of the two readers.

OWR (observation of writers and readers)

Whereas the evaluation in the previous two conditions is focused on only one communicative role, for the OWR subjects it is aimed at both roles. They observe a complete communicative process: the construction of a text by a writer and the reconstruction of the writers' intention by a reader. Their exercise runs like this:

OWR:

Read the following assignments:
The first student you will observe is the **writer** who was instructed to:

"Write a short argumentative text'

After the text was written, the second student or **reader** was asked to:

'Determine if this is an argumentative text. Tell us why."

You are going to observe both the writer and the reader.
It is your task to find out what each of them does well, and what they may be doing wrong.
When you have observed both students, you may advance to the next page.

(....)

(next page:)
You saw two students doing writing and reading assignments. Their answers were:

Writer:	**Reader:**
"I think I will enjoy reading this book, because I already like the introduction."	*'Yes, that is an argumentative text. That 's because she gives her opinion about the book."*

Explain briefly on which aspects the communication was successful or not.

===➤➤➤ Did the **writer** do well?
O Yes
because
O No...
....

===➤➤➤ Did the **reader** do well?
O Yes
because
O No...
....

OWR subjects must divide their attention between the two communication modes. More than for the other subjects it may become visible for them how strongly writing and reading – or the construction and reconstruction of meaning – are related through the use of the conceptual rule for 'argumentative text'. The subjects must evaluate writers and readers by their use of this rule; or, more precisely, of the two variants of the conceptual rule: the productive and the receptive transformation. Because of this varied representation of the rule, the theoretical element may become more integrated or 'connected' and therefore more readily applicable to both reading and writing.

FR (writer observation as feedback for readers)

This mirror–image of the FW condition consists of readers receiving feedback from the original writers. An artificial intervention had to be applied, since texts can of course not be analyzed before they are written. Therefore we used again authentic videotape recordings, which allowed the reader to 'go back in time' and observe how the text had once come about. An example:

FR:

Answer this question first. Is the following text argumentative or not?

"I think I will enjoy this book, because the introduction pleases me."

O **NO !**

O **YES, because** I can find the following characteristics of argumentative texts:

opinion: ....

reason for having this opinion:
....

(....)

Now look at the video.
You will see the writer working on the text that you have just tried to analyze.
It was the writer's task to:
'Write a small argumentative text'

It is your task to check if you have understood the writer well.
You must listen carefully for any indications that the writer may give about:
* the type of text (s)he wanted to write;
* the parts of which the text consists.

(....)

You saw the writer composing an argumentative text. It may have become clearer to you (from what the writer did or said) what his/her intentions were. So now you can check if you understood him/her well.

==>> Do you still think that your initial answer was correct?
O ....yes / no, because
....

Subjects from the FR condition also gain experience with both modalities: active reading first, followed by 'passive writing'. By observation of writing they can acquire knowledge about the strategic ways in which writers construct texts. They are instructed to use this knowledge as feedback to their previous reading performance. They may make this knowledge profitable by incorporating it in other reading, comprehension and analytical activities.

DRS (doing reading exercises with reinforced self-reflection)

This condition is on the one hand very similar to DR, and on the other hand very different. Similarities between DRS and DR are that the type of exercises is equal, namely direct application of theory to reading exercises. The most important differences are that in DRS, the texts to be analyzed have just been written by two peers from the FW condition (so each DRS subject analyzes texts which are different from the other DRS subjects) and that the subjects must do the analytical exercises by reading and thinking aloud, while being observed by the writer. This social situation is for this reason different from the relatively independent DR subjects.

DRS:

1a) Your partner gives you a text that he or she wrote.
Question: Is this an argumentative text?

O **NO !**

O **YES, because** I can find the following characteristics of argumentative texts:

opinion: ....
reason for having this opinion: ....

[think aloud !!] *[tell how you find your answer]*

1b) Your other partner also gives you a text. Question: Is this an argumentative text?

O **NO !**

O **YES, because** I can find the following characteristics of argumentative texts:

opinion: ....
reason for having this opinion: ....

[think aloud !!] *[tell how you find your answer]*

1c) You have just read two texts. Which of those two was best?

O the first O the last

1d) Tell why you did not like the other as much. How would you improve it?

....

Videotape recordings

The conditions OR, OWR and FR make observations by means of videotape recordings. The same type of self–instructional video lessons were developed as was used by conditions OW and OWR in the previous experiment, composed of the same set of recordings (see end of section 5.3.3).

The videotapes for the experimental sessions were edited in such a way, that for almost every exercise two different processes or solutions could be observed. This was done in order to provoke active interest from the observer and to offer him/her a problem to solve: to choose the best from the two realistic solutions to the task.

In Appendix 6.1 an index to the videotapes can be found.

Prior testing

The workbook–lessons of condition DR were also tested prior to the experiment, in order to check for comprehension problems in the text. This testing was done with a group of students similar to those who participated in testing the DW version. As a result, one exercise was removed from the last lesson in order to not make this lesson not too long.

Comparing learning activities in the DR, DRS, OR, OWR and FR conditions

We can make a comparison between the five conditions with respect to the type of cognitive activities they perform. By doing so, it becomes clearer to which differences in activities we may attribute possible differences in effectivity. Again, the example of defining argumentative text ($S + A \rightarrow B$) is used.

Table 6.4: *Comparing learning activities in the DR, DRS, OR, OWR and FR conditions*

	DR: 'Identify argumentative text'	**DRS**: 'Identify argumentative text while thinking–aloud'	**OR:** 'Check if P identifies argumentative text correctly'	**OWR**: 'Check if P writes a correct arg. text, and if Q analyzes it correctly.'	**FR**: 'Identify argumentative text; check if it matches P's intentions.'
ORIENTATION:	Reading, interpreting and conceptualizing the decoding rule: S + A $\rightarrow$ B.	same	same	same	same
EXECUTION:	Concretizing: checking for properties S and A	same as DR, plus thinking– aloud and prompted evaluation	Observe readers concretizing by identifying S and A	1a: Observe writer 2a: Observe reader	same as DR
MONITOR/ EVALUATION	Self–monitoring and self–evaluation (?)	same	Observe reading and give evaluative comments	1b: Comment on writing 2b: Comment on reading	Self–evaluation by means of writer observation
REFLECTION:	Accommodate: change or confirm conception of decoding rule.	same	same	same	same

The 'orientation' and 'reflection' steps in the exercises are the same for each condition. In these steps, the learner's cognition is construed (making an initial conceptualization of the decoding rule in the first step) or accommodated (changed or confirmed as a result of experience during the exercise). The steps in between are 'practice steps' in which the application of this rule is either undertaken or observed, and evaluated. As advocated in section 3.3, these monitoring and evaluation activities are supposed to yield the information that is the basis for learning. 'Reflection' steps are not prompted; that is, the extent to reflection is left up to the subjects. By invoking these different learning activities in the four conditions, we attempt to find if there are any differences in learning associated with them (as theory suggests).

6.3.4 Test materials

The tests used for the measurement for the resulting writing ability and the resulting reading ability have been discussed in section 5.3.4. Exactly the same learning goals, indicators, and tests have been used in this experiment. The only difference is that posttest measurement of Reading skill is a measure for learning in this experiment, while it was a measure for transfer in the previous one. Accordingly, the posttest measurement of Writing skill is a measure for transfer now, instead of a measure for learning. This means that all information on posttests and covariates given in sections 5.3.4.1 and 5.3.4.2 is valid for this experiment as well.

6.3.5 Procedures

The scheduling of sessions, pre– and posttests and lessons was equal to the previous experiment.
All subjects from conditions DR, DRS, OR, OWR and FR worked individually from a workbook, in which theory and exercises were combined. As we said earlier: condition DRS was different in that the groups were smaller, because one DRS subject functioned as a reader for two FW writers. Two such triads worked at the same time in the same room, under supervision of two research assistants.

Learning–by–doing – normal self–reflection (DR)

Sessions of this learning–by–doing condition was comparable to DW. After groupwise administration of the pretests, subjects had a workbook and a pen at their disposal and could work individually in a normal tempo until the hour was over. Due to the fact that the course had been pretested, the time estimation of one hour appeared to be sufficient.

Learning-by-doing - reinforced self-reflection (DRS)

Sessions of this condition were unlike any other, because of the demand for thinking–aloud activity on the subject. The organization of the lessons and the nature of the texts to analyze were similar to the DR condition, with the difference that when a DR subject would arrive at an exercise containing pre–printed text in the workbook, the DRS subject would receive a handwritten text by another student (from the FW condition, see chapter 5), and would have to perform the reading task with this text. Once this task was performed, the text was returned to its author, and a second student would come with another text. After reading two texts, the DRS subject was asked to compare both texts on comprehensibility and to motivate a choice for the better one. Because they had to spend time on answering this question, DRS subjects could read 20% fewer texts than the DR subjects.

Learning-by-observation & evaluation (OR and OWR)

The learning–by–observation & evaluation sessions were a bit more complicated. They are procedurally similar to OW and (of course) OWR, for which explanations are given in section 5.3.5. Subjects worked with a workbook and were seated facing a videoplayer. Each subject used a 'personal' videoplayer and headphones. With these devices, the models were observed; comments were written in the workbook.

Learning-by-observation-as-feedback (FR)

Subjects from the FR condition also worked with videotaped fragments; the procedure was similar to that of OW, OR and OWR. Instead of observing readers, these subject observed writers as feedback to their reading performance.

6.4. Results

The results of this study will be presented in the same way as in the previous chapter. Thus, the first part will report on the instrumentation of pre– and posttest measurement. In the second part, the research hypotheses will be statistically tested by multivariate variance analyses performed on the posttest data using the relevant pretest data as covariates. The first of the two sections in this latter part is about the learning measures: to which extent do the various types of reading practice lead to an increase in reading skill? The second section concentrates on the transfer measures: to which extent is writing performance stimulated by each of the types of reading practice?

6.4.1 Instrumentation

I will list the pre– and posttests instruments, give a short description and some psychometric data: number of items ('standard including rejected items' and 'rejected' after item analysis using $R_{item–total} \geq 0,15$ as a criterion) and homogeneity (after removal of non–fitting items).

The homogenities of the tests (Cronbach's alpha reliability of each test without its rejected items) are acceptable (> 0.60), with the exception of the pretest measurement of indicator 1a and the posttest measurements of indicators 2a and 2d; these indicators had also low reliability indices in the previous experiment. When corrected for test length (with the Spearman–Brown formula), the relibiability of these tests becomes acceptable as well.

Items with an item–total correlation of less than 0.15 were rejected from the tests. The number of removed items is visible in the table.

Table 6.2: *Number of items, number of rejected items, and reliability indices for pretests and posttests*

		Pretest	# items total	#items rejected	alpha	Posttest	# items total	#items rejected	alpha	correlation pre–post
1a)	**Writing – social context:**	1A	6	1	.52	W1D	3		.69	.134
1b)	**Writing – text structure:**	1B	18	3	.61	W1A, W3B	9	2	.67	.261
1c)	**Writing – argumentation structure (simple):**	– –				W3	11	1	.82	
	Writing – argumentation structure (complex):	– –				W1C	12	3	.67	
						W2	32		.88	
1d)	**Writing – means for presentation:**	1D	6	1	.60	W1B, W3A	14	1	.71	.118
2a)	**Reading – social context:**	– –				R3ABCD	4		.59	
2b)	**Reading – text structure:**	– –				R2D, R3F	14	3	.82	
2c)	**Reading – argumentation structure (simple):**	2Ca	33	3	.83	R1	48	4	.93	.513
	Reading – argumentation structure (complex):	2Cb	20	6	.76	R2C, R3E	19	2	.82	.072
2d)	**Reading – means for presentation:**	– –				R2AB, R3G	18		.56	
iq1	**Intelligence – CMR 'Conclusions':**	IQ1	40	1	.86	– –				
iq2	**Intelligence – CMU 'Word list':**	IQ2	75	8	.91	– –				
iq3	**Intelligence – 'Verbal Analogies':**	IQ3	50	1	.87	– –				

Posttest measurement of indicators 1B, 1D, 2B, 2C (complex) and 2D took place with more than one test. The relevant parts of the test were taken together in the analysis. The psychometric data reported in this table is also based on these parts together.

When trying to determine which pretests can function as covariates in the analysis of posttest data, we run into the same problem with low correlations as in the previous chapter (see the pretest–posttest correlation table in Appendix 6.2). It is only indicator R3 of which the pretest (2Ca) and posttest measures correlate *and* are theoretically related. The other theoretically related pre– and posttests do not appear to measure the same construct.

The average intercorrelation of posttest measures *within* each of the modi (both reading and writing) is higher than the average intercorrelation *between* these modi. As a consequence, the five operationalizations of each dependent variable cannot be considered independent. For this reason I will use multivariate analysis of variance (MANOVA), a technique that can take the phenomenon into account of several dependent variables influencing each other. Since our design consists of only one independent variable (type of practice), a one–way multivariate analysis of variance will be performed.

The covariates that will be included in the analysis are displayed in the lower half of appendix 6.2. Of the IQ–measures, only the second and third will be included in the analysis. Another covariate will be PR1 on the analysis of both reading (correlating with R1 to R4) and writing (correlating with W3 and W4).

Randomization appears to have taken care of an absence of between–group differences concerning reading and writing skill, as measured in the pretest scores (see Appendix 6.3). Correction for existing individual differences will take place using the relevant pretest scores (IQ2, IQ3, PR3–1) as covariates.

6.4.2 Posttest data analysis: Effects on learning and transfer

Table 6.3 contains for each condition the mean posttest scores and standard deviations for each indicator:

[insert table 6.3]

Table 6.3: *Means and standard deviations for posttest scores across conditions*

The mean posttest scores for the indicators vary among groups. However, within–group variance is not low either. With the use of MANOVA procedures, as applied in the previous chapter, it will be decided whether the between–group differences that are observed may be generalized.

Learning effects

First I will address research questions regarding learning effects. That is, for all conditions: an effect on reaiding. learning to read. In the following table 6.4, the results of the MANOVA procedures on the posttest data are shown.

[insert table 6.4 here]

Table 6.4: *Results of MANOVA tests for between–group differences on the dependent variable 'reading'.*

It should be kept in mind that in the following hypothesis testing procedures, there is more than one dependent variable in each hypothesis. All group scores on all indicators are taken simultaneously in the analysis as dependent variables, taking their intercorrelations into account.

Research question Q1 asks for the comparison of conditions OR versus DR. In other words: is **learning–by–observing** students who are doing reading exercises more effective than **learning–by–doing** these exercises? The mean score of the observation condition on the reading posttest is significantly higher than the mean of the individual practice conditions. This means that learning–by–observing is more effective regarding reading argumentative text. This finding confirms our expectation. It corresponds with the answer on Q1 regarding writing (previous chapter). In this respect, there is no difference between learning–to–read and learning–to–write (research question Q6).

In condition FR, the students are involved in the observed communication. Feedback on one's reading activities by observation of a writer positively affects the students' reading skill acquisition: the MANOVA on the second line (Q2) is significant. This result is again in agreement with what was found regarding writing (previous chapter), so reading and writing don't seem to differ on this point (Q6).

Observation–of–models is again investigated by means of Q3, comparing learning–by–doing to learning–by–observation of both writer and reader: a complete communicative transaction. A MANOVA (line 3 in the table) shows that observation of both models leads to better results on the reading posttests. Since no such expectation was formulated, tests were two–tailed. In the writing mode, an experimental effect was not found (previous chapter; the probability of the difference under H0 was 0.08). The relative effectivity of OWR activities is apparently different regarding reading than regarding writing (Q6).

Q4 asks for a direct comparison of the OR and OWR conditions. The MANOVA procedure (fourth line) shows that there is no significant difference in learning–to–write between the groups. This is in agreement with our expectations. It is also in agreement with this same comparison in the writing mode (research question Q6).

Q5 asks for a comparison of both learning–by–doing conditions: making reading exercises silently, or reading–aloud and comparing texts. Results appear to favour the DRS condition, thus confirming the expectation.

In sum, when it comes to learning to read argumentative texts, each of the four experimental learning activities (OR, OWR, FR and DRS) appears be advantageous in comparison with the standard condition DR. Learning to read argumentative texts may be best served by the 'observation–one–mode'–method (highest F–value), but to assess this smaller effect further investigation is needed with larger and/or more homogenous groups. These results are in agreement with what was found regarding writing (previous chapter), except the OWR condition, which appeared more effective regarding reading, but not regarding writing.

Transfer effects

The other research questions (Q7 – Q10) are concerned with inter–modal transfer effects of the respective instructional methods. Intermodal transfer from writing to reading was found in several conditions in the previous experiment. We can investigate now whether the writing–reading transfer is bilateral or not.

In section 5.4.2, I promised to report on two types of information about transfer effects: tests for between–group differences in the extent to which they promote transfer, and a quantification of the amount of transfer realized by each type of learning activities. In the previous chapter, only information from this first category was given, and only with respect to writing–reading transfer. This was because a criterion was lacking to which transfer scores could be related. In the present experiment, data has been collected that will offer such a criterion. Thus, in the next section I report on the amounts of transfer promoted by the various instructional methods.

Our first concern is with testing for between–group differences, in the extent to which the experimental and control groups promote transfer. It was hypothesized that the experimental learning activities 'observation of models' and 'observation as feedback' may stimulate transfer, since they call for a more 'mindful' acquisition of the linguistic rules governing the communication: observation of models and commenting on them (the Sonnenschein & Whitehurst variant) require a certain distance and abstraction, while observation of communicative partners as feedback offers the opportunity to see a communicative rule function in its productive as well as in its receptive context.

For the present purpose MANOVA procedures were chosen again to determine if between–

group differences exist on the transfer task. The only difference is now that the reading students have been tested on their posttest writing skill in order to measure transfer. It was assumed that between–group differences regarding the writing task can be attributed to the experimental interventions as transfer effects. The results of the MANOVA's are in table 6.5:

[insert table 6.5 here]

Table 6.5: *MANOVA tests for between–group differences on the dependent variable 'transfer to writing'.*

Question Q7 is the transfer–oriented pendant of Q1: does learning–by–observing students doing reading exercises lead to higher transfer to writing than learning–by–doing these exercises? In the table we see that OR subjects outscore DR subjects with respect to writing skill, so the OR learning activities have apparently prepared them better. Thus, question Q7 must be answered positive. The same was the case in chapter 5, so this transfer mechanism appears to work bilateral.

The second research question, Q8, addresses transfer effects of observing one's communicative partner. Here, this means: observation of writers after doing a reading exercise. Much appears to be gained from this observation of authentic writers–at–work; not only did it affect the readers' reading skill (Q2) but it also enhances their writing as can be seen in the second line of table 6.5. Our expectation was confirmed; it is again in line with the findings in the writing domain (chapter 5) (Q10).

The intermodal transfer of DRS activities can be made visible by comparing their posttest scores for writing with posttestscores of the DR group (Q9). In the bottom line of table 6.5 one can see that these two methods are not equally effective, although similar in the nature and number of the performed reading exercises. The expectation of more transfer due to stronger self–regulative activities was confirmed.

Also in this experiment we could not assess genuine 'transfer effects' of the OWR condition: the learning effects of this condition are actually mixed learning–and–transfer effects. The OWR condition can principally not yield any intermodal transfer effects as defined by us, since the effects result from practice within both modes. Nevertheless, it is necessary to compare the OWR posttest score for writing with the transferscores of DR. Since OWR was more effective than DR (see answer on Q3), the amount of transfer to the other mode could be a further argument tipping the balance on which instructional method to use in education. A MANOVA using writing performance (=writing posttests) as dependent variable and condition (DR vs. OWR) as independent variable) yields a significant result ($F=3,37$; $p<0.01$; $n=29$) in favour of the OWR group. So it can be concluded that observation of complete communicative transfers is more effective regarding learning than method DR, and that a 'hidden strength' of the OWR method lies in the learner's ability to adapt the knowledge in communicative complementary situations: from the reader's perspective to a 'writers' perspective.

In conclusion, transfer from writing practice to reading skill is promoted more by the learning–by–observation method than by traditional learning–by–doing activities. In a learning–by–doing method of instruction, supporting self–reflective activities by thinking–aloud (condition DRS) also contributed to learning and to subsequent transfer.

Comparing amounts of learning and transfer

The hypothesis testing procedures reported in the previous section have the function of decision–making: to decide which instructional methods can be said to be advantageous over other. It helps us not only to establish *that* more transfer was promoted, but in order to establish *how much more,* a kind of quantification must be used which 1. enables incorporation of the five indicators into one construct 'writing skill' or 'reading skill', and which 2. is informative in that it expresses the achieved amount of transfer in relation to some meaningful criterion (see 5.2 and 2.6). In this section, an attempt will be made to quantify the achieved amounts of transfer in the various conditions.

In section 2.6, several ways to calculate 'amounts of transfer' were discussed, each leading to a more or less different concept and operationalization of 'transfer'. The preferable way to measure transfer was presented in formula 3 (or variants in 3A and 3B):

$$(3) \qquad T_{\%\ \text{improvement}} = \frac{E_{B1} - C_{B1}}{C_{B2} - C_{B1}} \times 100\ \%$$

The numerator represents the advantage of the experimental group over the control group on a task B. The denominator represents the progress of the control group after a certain amount of practice. The formula is meant to be used in situations where control and experimental groups have had an equal amount of practice (time–on–task, number of trials, number of assignments etc.). The complete formula drafts transfer as the *lead* that the experimental group has over a control group, and expresses this lead as a percentage of the progress found after a determined amount of practice in the control task.

In our experiments on writing–reading transfer, there are three questions to answer before transfer amounts of the various conditions can be presented.

The first question concerns the denomination of the groups. With 'experimental group' a condition is meant in which the subjects perform activities, aimed at learning skill A, and promoting transfer to a certain skill B. With 'control group', or reference group, the control condition is meant with learning–by–doing activities directly aimed at skill B (cf. section 2.6). For instance, a group of 'writers' may function as a control group to a group of 'readers' if we want to assess the transfer of their reading activities to writing. In our experiment, it means that DW is the control group for the experimental reading conditions, and that DR is the control group for the experimental writing conditions.

Second, the formula requires a measurement of skill prior to the treatment. Unfortunately, the pretest measurements that were obtained do apparently not measure the same concepts as the posttests (low intercorrelations). This leaves us two options: either to put C_{B1} at some fictitious level, or remove it from the formula altogether. In this presentation, a choice is made for the last option. It means that C_{B1} is set at zero, as if none of the skill that is acquired during the treatment was present at the start of the experiment, neither in the control group not in the experimental group. This assumption is not true in an absolute sense, of course. However, since the groups may be considered equal, and since the subject matter is completely new to this age–group, and since none of the two groups is harmed more than the other, we consider it not a bad choice. In making this choice, another problem is avoided (namely accounting for any arbitrary estimate of the starting level C_{B1}). By setting C_{B1} to zero, the formula (and hence the concept of transfer) changes into the ratio of experimental posttest performance and control posttest performance, from which 100 % will be subtracted in order to obtain net effects. The amount of transfer is then expressed as a percentage of the *lead* over, or arrears or drawbacks to, the control posttest score:

$$(5) \qquad T_{\%\ \mathrm{improvement}} = \left(\frac{E_{B1}}{C_{B2}} \times 100\ \% \right) - 100\%$$

Third, it seems that the formula can only be applied to posttests yielding a singular measure. In the present experiments however, the constructs 'writing skill' and 'reading skill' are not measured directly, but indicated by five subvariables or indicators. Since the indicators as operationalized in the posttests have comparably high intercorrelations and reliabilities, we will attempt to scale them by calculating a mean transfer effect across the five indicators for each condition. Cronbach's alpha's for these scales are .68 for reading and .57 for writing, calculated across all participants in the experiment.

For its dependency on absolute scores, the presented transfer measure will not allow for comparisons between domains. But it seems a good enough instrument to compare transfer effects assessed by the same instrumentarium. Thus we are able to quantify the transfer effects of every condition, and within one mode the transfer measures may be compared.

In table 6.6 the learning and transfer effects on each indicator of each reading condition is displayed. Learning effects are assessed in a comparable way as transfer effects, by the ratio of posttest scores of experimental and control condition (measured in the same mode) and subtracting 100 %. On the right hand, the mean effects across indicators are displayed.

Since we have now data on condition DR, a similar table for the experiment reported in chapter 5 can be constructed, displaying learning and transfer effects of conditions DW, OW, OWR and FW (Table 6.7).

Tables 6.6 and 6.7 should be read as follows. The upper half displays the learning effects, in which the standard learning–by–doing condition is taken as a point of reference or control–group. The lower half of each table displays the transfer effects, now with the learning–by–doing condition of the complementary mode taken as control group. The middle five columns present learning and transfer effects for each of the five indicators, as calculated by the above measure. The right hand column displays the mean learning or transfer effect that an experimental condition adds to the standard condition, calculated across the indicators (unweighed). It is important to note that only the effects marked with an asterisk are significant. The learning effect of the OWR condition in the writing domain was not significant and must therefore be considered non–existent. Thus, no learning size for writing is assessed or displayed.

Looking at the tables, the following points are noteworthy. In the upper part, there are large differences in learning effects between the indicators. Apparently some learning goals of the argumentation course (R2, R4,W1, W4) benefit more from the experimental instructional methods than other goals (R1, R3–1, W2). These last goals are still more effectively acquired by the observation methods, but the difference is not large.

Furthermore, the mean learning effects of the observation methods do not differ much, which is in agreement with the general positive answer to the research questions Q1, Q2, Q3 and Q5 and the non–significance found in answer to Q4 (OW vs. OWR). The mean learning effects of the writing methods do differ, specially since the size of the OWR effect is about half the size of the OW effect regarding learning–to–write. This is not in agreement with answering Q4 negatively; apparently the comparably large contribution of indicator W4 to the effect size accounts for the seemingly more effective OW condition.

Looking at the lower, 'transfer' effect part, two conclusions can be drawn. First, it is evident that the learning–by–doing conditions deviate from the observation conditions in yielding transfer. All transfer scores are lower than the matching learning–by–doing scores of the complementary mode (which can be expected), while the transfer scores of the other conditions are usually above this level. Second, transfer effects from learning–to–write (by observation) to the reading mode are much stronger than transfer effects from learning–to–read (by observation) to the writing mode, while within a certain mode, the effects are of a similar size. This is in agreement with the anwers to research questions Q7, Q8 and Q9. It should be noticed that the transfer effect of OWR is not genuine, since the OWR subjects observed both writers and readers with the purpose of learning from their behavior. It is striking that the DRS group attain a (genuine) transfer level that seems to equal the learning of the OWR group.

The amounts of transfer yielded by the various variants of the learning–by–doing and learning–by–observation methods are graphically displayed in figures 6.3 and 6.4. Figure 6.3, which corresponds with table 6.6, shows in the upper part the learning effects of the experimental writing conditions as leads over the control condition DW; in the lower part the transfer effects of all writing conditions (experimental and control) over the control conditions DR. In figure 6.4, the communicative counterpart effects are displayed: learning effects of the observation methods of instruction in the upper half, and transfer effects in the lower half.

6.5. Summary and discussion

Summary

It was the aim of this study to repeat the experiment reported in chapter 5, but now aimed at learning to read (and transfer to writing) instead of learning to write (and transfer to reading). The same theoretical considerations underlie this study, and in fact the experiments in chapters 5 and 6 can be integrated in a larger experimental design testing for mutual transfer between related cognitive skills (section 2.6). Furthermore, both experiments are aimed at the same learning content (argumentative texts) to be acquired by the same group of students (ninth–grade secondary students of intermediate and high level).

In the study, the expectation was tested that two types of activities, which are both instances of 'observational learning', would be effective regarding learning to read as well as regarding transfer to writing. To this purpose an experiment was set up, using a full–between pretest–posttest design, in which four groups of thirty, and one group of fifteen 15–year old secondary students took part. The four treatments consisted of short experimental courses aimed at learning to read and analyze argumentative text. The presented subject–matter was the same for each group, but the learning activities varied systematically: doing reading exercises, doing reading exercises while thinking and commenting aloud, observing readers, observing both writers and their readers, and doing reading exercises and observing writers as feedback. After a pretest session and four one–hour training sessions, the same set of posttests measuring reading and writing skill were administered to all participants.

A multivariate analysis of variance was used in order to test the hypotheses regarding learning effects, using 'intervention' as an independent variable and a set of five indicators for reading skill as compound dependent variable. Hypotheses regarding transfer effects were tested in the same way, but with a set of five indicators for writing skill as dependent variable in the MANOVA (dependent variables were indicated in the same way as in chapter 5).

The main findings are that both types of learning–by–observation (observation–as–model and observation–as–feedback) are more effective (regarding learning) than learning–by–doing (Q1 and Q2). Two variants of observation–as–model were distinguished: observation of readers only (who analyze given texts), and observation of complete communicative transactions (both writers and their readers). It was found that observation of both writers and readers surpassed the effectivity of learning–by–doing (Q3), and that it adds to this an important extra effect: a strong transfer to reading (Q7).

Moreover, observation of complete reading–writing processes does not yield more or less transfer than observation of writing processes only (Q4). Also, the transfer effect of observation–as–feedback is stronger than that of learning–by–doing (Q8). Finally, the learning–by–doing condition including thinking–aloud and commenting on text difficulty appeared to be more effective than learning–by–doing with respect to both learning and transfer (Q5 and Q9). This is in agreement with the learning and performance effects we found in chapter 4, with subjects thinking aloud while reading and commenting on a manual text.

Again, intermodal transfer is found that is not sanctioned by receiving oral feedback by a teacher, such as was the case in the Sonnenschein & Whitehurst study (1984). It was already suggested that this may be due to the very different group of learners and a totally different subject. Intermodal transfer can be observed in all of the conditions, through observation or through self–reflection.

In sum, the experimental effects support the theoretical model–variant of learning–by–observation (Sonnenschein & Whitehurst) as well as the feedback–variant (Schriver). The learning–by–observing method, in the two shapes that we have put it in, can be very effective, for either of the communication channels: productive or receptive. In that respect, the findings also support the interactive reading–writing model, which predicts the possibility of transfer from either mode to the other. I will come back to this in the final chapter.

Validity

Since this experiment was set up in a similar way to the previous experiment, the same considerations are valid with respect to matters of validity. Several alternative explanations have been ruled out by making the groups comparable, keeping the time–on–task constant, making the lessons self–instructive, motivating the students, and developing interventions that are highly similar in subject–matter content.

The main validity problems concern the relative 'novelty' of the experimental conditions, which may have motivated the students more, and the not completely equal task settings with fewer subjects in one room. A minor point is that the DRS subject had about 10–15% more time for completing their lessons (due to organizational circumstances); time that was mainly spent on group organization and not on learning.

The problem must be added with the pretest measurements which would not correlate with the posttest measurements. It was not possible to avoid this problem that was experienced in the previous experiment, because the data for this experiment was collected – using the same pretests and posttests – some time before the data of the previous experiment were analyzed and this problem was detected.

It should not be forgotten that the two experiments are carried out not only to test an instructional theory predicting learning and transfer from observation activities, but also a task–oriented transfer theory predicting transfer between skills that have a certain cognitive overlap. The mutual transfer effects that are found make a strong case for the interrelationship between the human activities of reading and writing. We still need to know much more about the actual thinking activities during observation and understanding, and about thinking activities during the application of subject matter to reading and writing exercises, in order to learn more about the precise relationship between these skills.

Figure 6.1: *Structure of independent variables in this experiment*

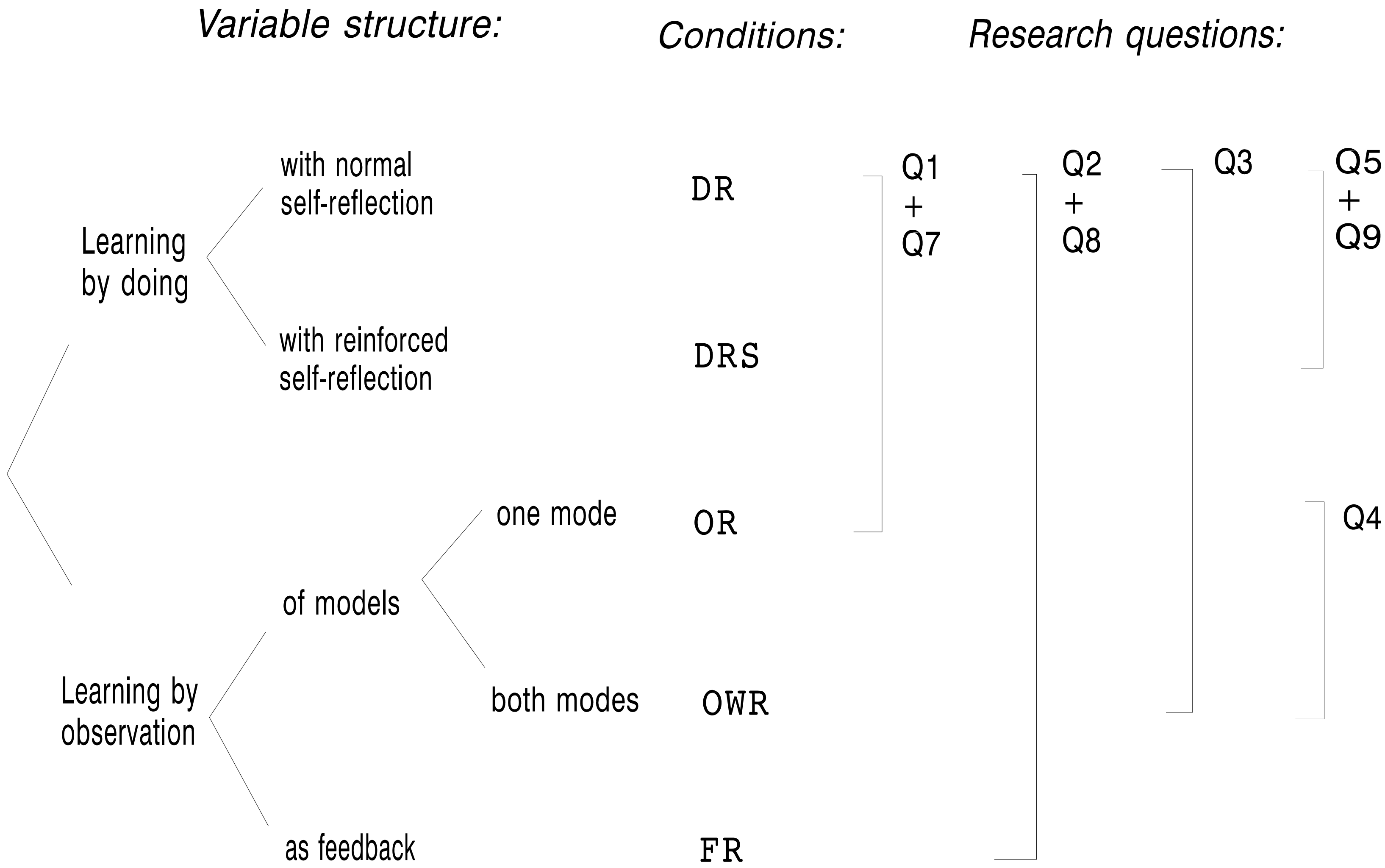
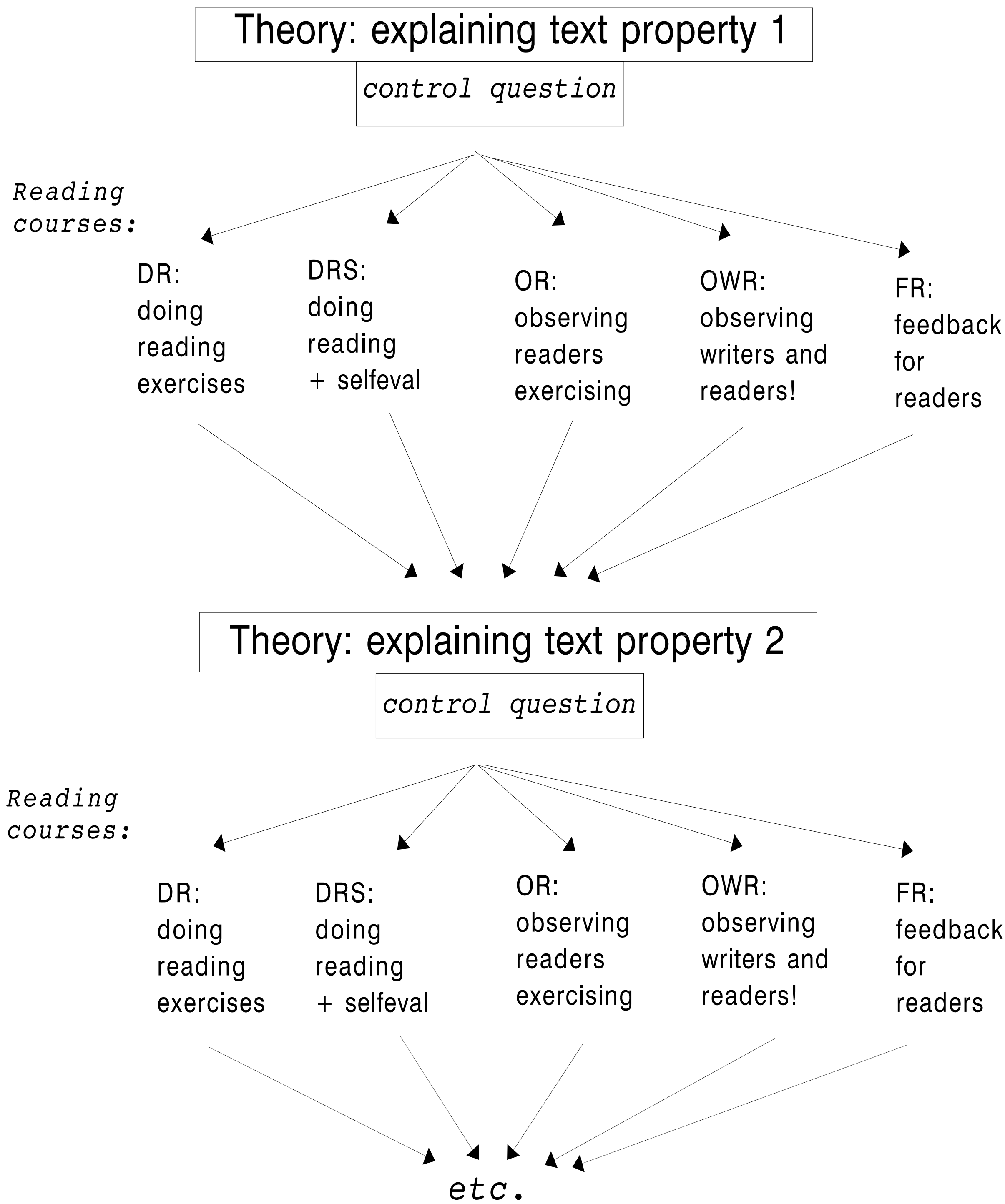


Figure 6.2: *Instructional sequence of the courses*



Appendix 6.2: *Pretest–posttest correlation table*

Posttests:	W1	W2	W3–1	W3–2	W4	R1	R2	R3A	R3B	R4
Posttests:										
W1 (writ)										
W2	.1457									
W3–1	.4262**	-.2416**								
W3–2	.0853	.1066	-.0541							
W4	.2131**	.0934	.3984**	.1184						
R1 (read)	.0953	.0527	.1096	.0442	.2575*					
R2	.1148	.1482	–0176	.0416	.2218*	.3521**				
R3A	.2526*	–.0591	.1133	.0923	.1545	.2583	.2225*			
R3B	.1077	.0144	.0633	.1539	.1729*	.3869**	.7127**	.2263		
R4	.1516	.1126	.1267	.1115	.0837	.3797**	.6202**	.1972	.4972**	
Pretests:										
IQ1 (iq)	–.0737	.0551	.1311	-.1272	.0993	.1263	.2794	.0141	.1483	.1670
IQ2	.1463	-.0356	.1392	–.0731	.0496	.1371	.2271*	–.1611	.3805**	.1255
IQ3	.2390**	–.0877	.1531	.3230**	.2208*	.1871	.4713**	.4814**	.3664**	.3617**
1A (writ)	– .0318	-.0532	.1494	.1244	.1349	.0617	.0149	.0474	.1478	–.1294
1B	.0983	.1908	–0990	-.0796	.0745	.0204	.1246	.0111	.1633	.0211
1D	.1143	–.0638	.0547	.0265	–.1110	.0897	.0351	-.0149	.1419	.1147
2Ca (read)	.2137	-.1064	.0901	.3474**	.3092**	–.0968	.4578**	.4595**	.4178**	.4161**
2Cd	.1065	.0608	.1189	.1403	-.0514	-.0742	–.0425	–.1061	.0555	.0612
# of covariates:	1	0	0	2	2	1	3	2	3	2

* = $p < 0.01$ ** = $p < 0.001$

Appendix 6.3: *Means and standard deviations for pretest scores across conditions*

	Pretests for IQ:						Pretests for READING			
CONDITION:	IQ1		IQ2		IQ3		PL1		PL2	
	CMR–test		*CMU–test*	*CMR–test*	*argu. simple*		*argu. complex*			
Learning by Doing Exercises	20.35	6.73	44.10	8.15	23.52	11.55	21.54	7.45	5.12	3.12
Learning by Doing Exercises (Reinf. Selfreg.)	21.12	5.89	45.82	9.12	22.36	9.97	23.15	4.34	4.47	2.78
Learning by Observation (1 mode)	20.72	6.28	48.53	8.01	23.74	10.27	21.16	3,84	5.71	3.14
Learning by Observation (2 modes)	21.12	7.58	51.61	9.72	25.82	9.87	23.02	5.27	5.12	3.01
Learning by Observation as Feedback	19.53	6.32	47.17	8.17	24.31	10.81	22.73	5.34	4.79	2.64
Max. score:	40		75		50		33		20	

	Pretests for WRITING:					
CONDITION:	W1		W2		W4	
	soc.context		*text structure*		*presentation*	
Learning by Doing Exercises	2.13	1.87	9.75	3.55	4.56	2.62
Learning by Doing Exercises (Reinf. Selfreg.)	3.43	2.24	8.33	3.31	3.99	1.38
Learning by Observation (1 mode)	5.62	1.67	9.12	4.67	5.71	1.75
Learning by Observation (2 modes)	4.12	2.13	8.94	3.52	5.26	2.44
Learning by Observation as Feedback	4.07	1.57	8.45	4.72	4.16	2.18
Max. score:	8		18		6	

Table 6.3: *Means and standard deviations for posttest scores across conditions*

	Posttests for READING									
CONDITION:	R1		R2		R3–1		R3–2		R4	
	soc.context		*text structure*		*argu. simple*		*argu. complex*		*presentation*	
DR: Learning by Doing Reading Exercises	7.27	(2.57)	12.34	(5.62)	24.41	(7.43)	16.63	(4.78)	7.53	(1.62)
DRS: Learning by Doing Exercises (Reinf. Selfrg.)	8.88	(3.31)	15.82	(5.74)	29.52	(8.34)	24.73	(4.13)	11.89	(1.25)
OR: Learning by Observation (1 mode)	9.12	(3.62)	17.11	(5.15)	28.15	(8.73)	23.57	(5.14)	13.73	(1.93)
OWR: Learning by Observation (2 modes)	8.54	(2.66)	16.64	(5.74)	27.32	(8.45)	21.73	(5.44)	12.59	(2.64)
FR: Learning by Observation as Feedback	8.63	(1.89)	16.11	(5.13)	29.32	(6.74)	21.84	(4.83)	9.73	(2.71)
Max. score:	12		22		43		34		18	

	Posttests for WRITING:									
CONDITION:	W1		W2		W3–1		W3–2		W4	
	soc.context		*text structure*		*argu. simple*		*argu. complex*		*presentation*	
DR: Learning by Doing Reading Exercises	1.63	(1.89)	5.14	(4.72)	18.78	(7.73)	10.12	(6.52)	2.93	(3.04)
DRS: Learning by Doing Exercises (Reinf. Selfrg.)	3.37	(2.03)	6.79	(5.63)	22.55	(6.82)	19.64	(6.55)	4.89	(2.02)
OR: Learning by Observation (1 mode)	2.86	(2.14)	6.12	(5.55)	21.52	(8.25)	18.25	(5.83)	4.42	(3.41)
OWR: Learning by Observation (2 modes)	3.23	(2.47)	6.43	(4.83)	21.76	(7.79)	17.01	(6.52)	4.82	(3.72)
FR: Learning by Observation as Feedback	2.82	(1.89)	6.73	(4.58)	23.18	(6.12)	17.86	(4.33)	6.21	(2.42)
Max. score:	6		14		30		32		12	

Table 6.4: *MANOVA tests for differences betweeen reading and observation groups on the dependent variable 'learning to read'. In the statistical design, all five indicators are included in the construct 'reading skill'.*

question	conditions	posttest	covariates	n	F	signific.	testing
Q1	DR – OR	Reading	IQ3, PR1	58	4.12	< 0.01 **	one–sided
Q2	DR – FR	Reading	IQ3, PR1	59	2.31	< 0.02 **	one–sided
Q3	DR – OWR	Reading	IQ3, PR1	59	2.69	< 0.02 *	two–sided
Q4	OR – OWR	Reading	IQ3, PR1	59	0.69	0.39	two–sided
Q5	DR – DRS	Reading	IQ3, PR1	59	3.03	< 0.01 **	one–sided

** = significant at $\alpha = 0.01$

* = significant at $\alpha = 0.05$

Table 6.5: *MANOVA tests for differences between reading conditions on the dependent variable 'transfer to writing'. The design includes all five indicators of the construct 'writing skill'.*

question:	conditions:	posttest:	covariates:	n:	F	signific.:	toetsing:
Q7	DR – OR	Writing	IQ2, IQ3, PR1	58	3.48	< 0.01 *	one–sided
Q8	DR – FR	Writing	IQ2, IQ3, PR1	59	3.02	< 0.02 *	one–sided
Q9	DR – DRS	Writing	IQ2, IQ3, PR1	59	3.56	< 0.01 *	one–sided

* = significant at $\alpha = 0.05$

Table 6.6: *Reading conditions: effects on learning–to–read and transfer–to–writing. Mean effect across indicators.*

Transfer as (ratio between posttest performance and control group performance x 100 %) – 100%

Learning to READ:

LEARNING EFFECT: (control group = DR)	R1 presentation	R2 soc.context	R3–1 text structure	R3–2 argu. simple	R4 argu. complex	mean effect:
DRS / DR:	58	22	28	21	42	34 %
OR / DR:	82	25	39	15	42	41 %
OWR / DR:	67	17	35	12	31	32 %
FR / DR:	36	18	31	24	32	26 %

Transfer to WRITING:

TRANSFER EFFECT: (control group = DW)	W1 presentation	W2 soc.context	W3–1 text structure	W3–2 argu. simple	W4 argu. complex	mean effect:
DR / DW:	–29	– 36	–32	–13	–39	–31 %
DRS / DW:	18	34	–10	5	19	13 %
OR / DW:	06	13	–19	1	10	02 %
OWR / DW:	16	29	–14	1	3	12 %
FR / DW:	02	12	–10	8	– 5	01 %

Table 6.7: *Writing conditions: effects on learning–to–write and transfer–to–reading. Mean effect across indicators. Transfer as (ratio between posttest performance and control group performance x 100 %) – 100%*

Learning to WRITE:

LEARNING W1 EFFECT: (control group = DW)		W2 *soc.context*	W3–1 *text structure*	W3–2 *argu. simple*	W4 *argu. complex*	Mean *presentation*	effect::
OW	/ DW:	64	24	28	36	115	53 %
OWR	/ DW:	43	17	21	22	37	28 %
FW	/ DW:	51	13	23	36	111	46 %

Transfer to READING:

TRANSFER R1 EFFECT: (control group = DR)		R2 *soc.context*	R3–1 *text structure*	R3–2 *argu. simple*	R4 *argu. complex*	Mean *presentation*	effect:
DW	/ DR:	–22	–27	–4	–48	–28	–26 %
OW	/ DR:	11	39	13	46	60	34 %
OWR	/ DR:	20	26	17	29	66	32 %
FW	/ DR:	20	46	20	49	62	39 %

Figure 6.3: Learning and Transfer effects of the instructional methods for Writing

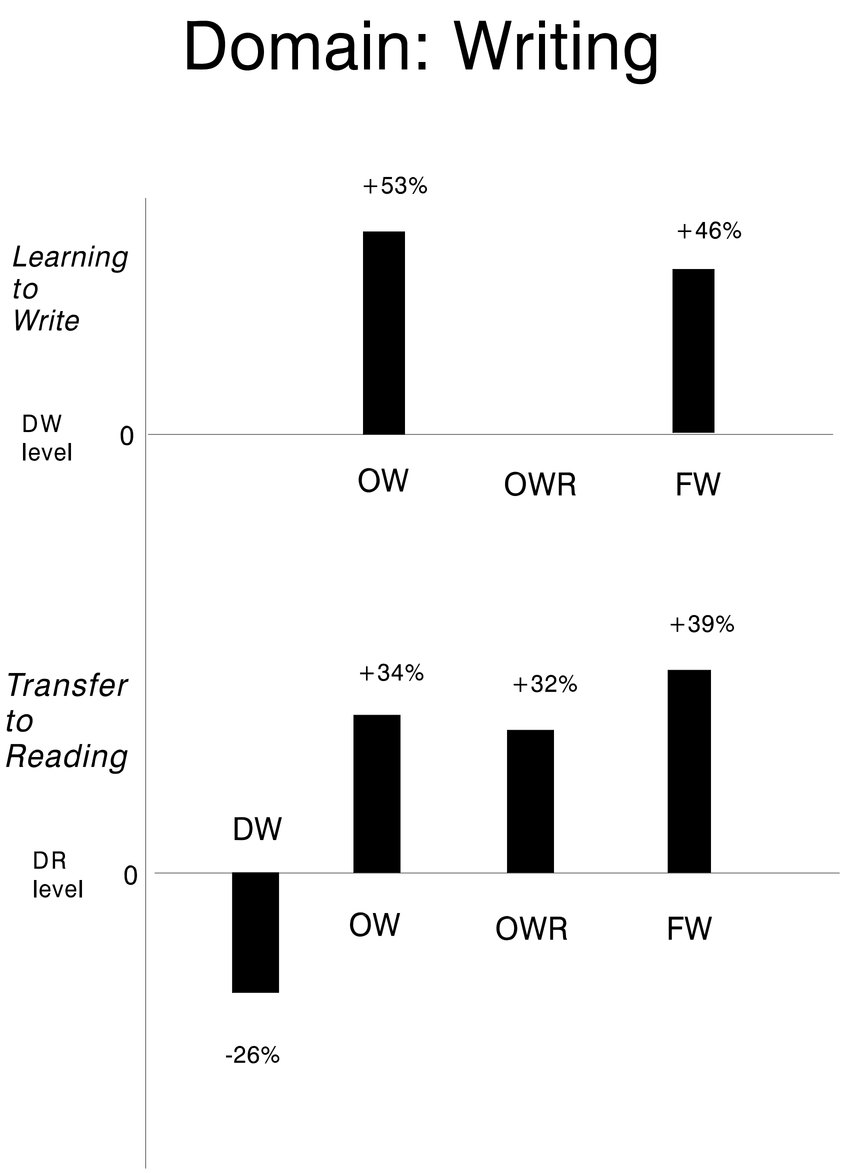
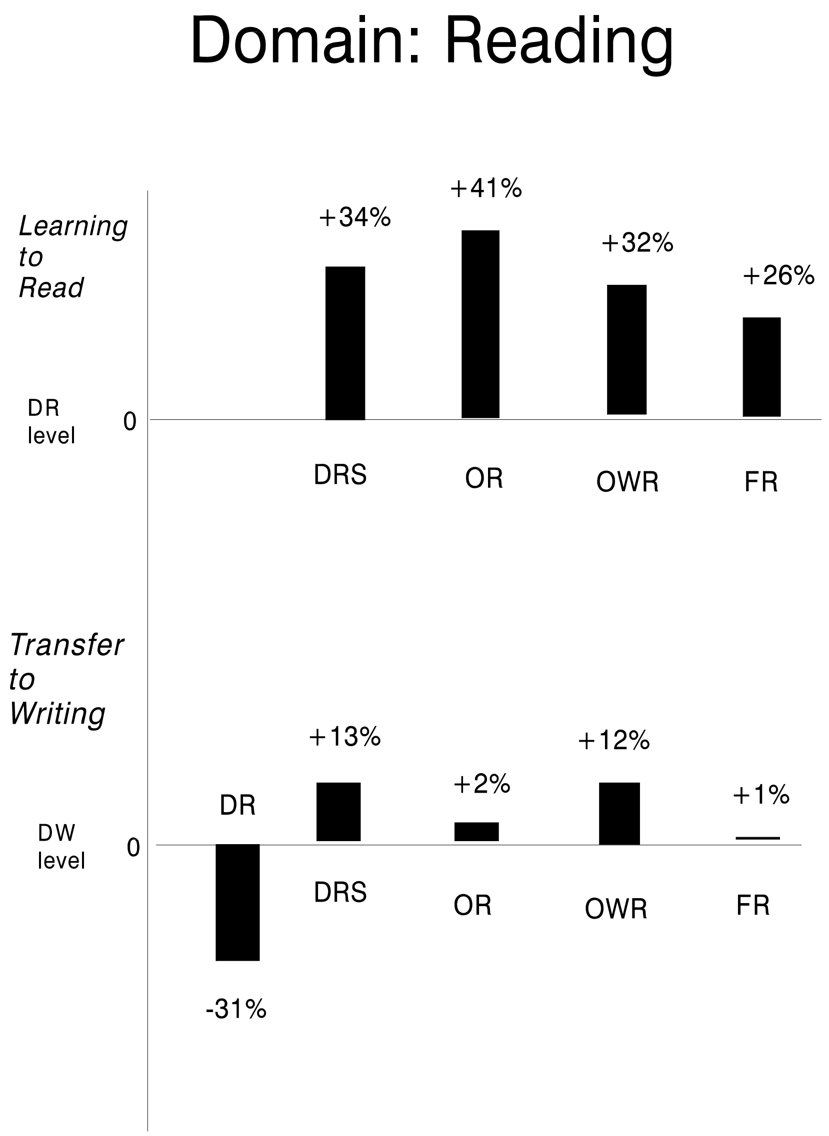


Figure 6.4: Learning and Transfer effects of the instructional methods for Reading



CHAPTER 7: Observations: Looking back and looking ahead

A summary of the theories underlying this study is followed by an overview of the answers to the research questions. These results are taken further in a discussion about the theoretical implications for research on learning, on transfer and on the domains of reading and writing. It is concluded that consecutive research should include process data. Finally, the relevance of this study for educational practice is discussed – and exemplified.

7.1 Promoting learning and transfer of reading and writing: The theory in a nutshell

This study has the development in view of more effective instruction for reading and writing at the secondary level. Two routes were chosen for this purpose: to enhance learning within each of these communicative modes, and to enhance transfer between these modes.

The need to enhance learning will need no special motivation. Virtually all efforts of teachers, learning psychologists and curriculum developers are aimed at designing instruction that is effective and efficient in achieving learning goals. As a starting point for this study, a simple, type of learning–by–doing was presented as an often–used method for reading and writing instruction. Several 'weak spots' or potential impediments of this method for the achievement of learning goals were sorted out (sections 1.2 and 3.3). The main criticism of learning–by–doing is that it focuses the learners' attention at the task or exercise, but not at the parallel learning task. Learners tend to complete aim for task completion, but not to invest in observing, evaluating and reflect on the cognitive activities they invoke, and which serve as input for the parallel learning process. Thus, even a successful task completion is in itself no guarantee that learning is optimal, or takes place at all.

The need to enhance transfer is not only motivated by the idea that transfer, as a 'side–effect' to learning, will save costly instruction time or effort, but that it will also contribute to the quality of learning. Transferrable knowledge is by definition more useful than non–transferrable knowledge. Transferrable knowledge is more abstract in that it has more handles for application than required by the few particular exercises during which it is constructed. It offers the learner the possibility to perceive and understand the connection between these exercises and other possible applications, to connect new domains of application, and thus to understand a bit more about how parts of his or her world are interrelated. Transfer–oriented reading and writing instruction should not just result in less instruction time, but in a deeper insight of the learners in why and how their knowledge is relevant to both reading and writing.

This statement with respect to the learner can also be made with respect to the researcher. By examining possibilities for transfer between language skill modi, the researcher can also acquire more insight into the nature of the reading–writing relationship. If, for instance, transfer is more likely to take place from writing to reading than vice versa, this is informative about the way in which writers' or readers' knowledge is acquired or stored. Thus, the results of this study may also be connected with the theory on unilateral or bilateral models of the reading–writing relationship, as presented in section 3.2.

The search for more effective instruction in this study was guided by considerations from three perspectives: a learning–theoretical perspective, a transfer–theoretical perspective, and a domain–oriented perspective.

First, in a learning setting that relies on the effect of doing exercises, the monitoring and evaluative activities determine to a large extent *whether* the students will learn, and *what* they will learn (section 3.3). If learners should gain knowledge from practice, they should at least have some idea about the task execution activities they invoke (or *can* invoke), and they should put effort in evaluating these activities. This enables construction of knowledge about an arsenal of activities, labeled as useful or avoidable, which they can use in future tasks. As said, the standard learning–by–doing instruction does not particularly stimulate self–monitoring or self–evaluative activities. In fact, learning–by–doing leaves it completely up to the learners if they want to only 'go through the motions' or want to take more advantage of the exercise. Such instruction will mainly benefit those students who are already keen and/or motivated enough to use the exercises to their cognitive advantage. For other learners, the learning process ought to be scaffolded. This is attempted here by directing their attention to what is assumed to be the most learning–oriented part of the task execution process: the monitoring and evaluation activities.

Second, to stimulate transfer between tasks or task domains, one should do two things. The tasks or domains should be described in such a way that specification of the 'identical elements' is possible. Then, one should develop instruction that will increase the chance that these elements, learned in a training setting, will be applied in other fields of application (section 2.4). As long as reading instruction and writing instruction are each focused on their own learning goals, transfer between these skills will remain a matter of 'incidental learning' or, more precise, 'incidental transfer' (Rijlaarsdam, Van den Bergh & Zwarts, 1992).

Third, in an attempt to describe the 'identical' or 'common' elements in the writing and reading domains, a communication–analytical perspective was chosen in which writing and reading are considered as counterparts in a joint social venture of information transfer. Writer and reader each take charge of a part of this information transfer, following certain guidelines for co–operation (cf. Grice's Cooperative Principle, 1975). Such guidelines assume that the writer uses coding rules to express meaning into verbal code, which may be transformed by readers using these rules in a reverse (decoding) order to reconstruct the writer's intended message. A writer may invent coding rules by transforming the decoding rules that he knows, assuming that his readers will know or be able to derive the correct decoding rules as well. This perspective opens the way for instruction enabling transfer between writing and reading tasks that share such related coding/decoding rules, and that stimulates the transformation of coding into decoding rules (or vice versa). The transformation requires abstraction from the specific mode, thus the instruction should stimulate the learner to make such (mindful) abstractions.

In sum, the quest in this study is for a type of instruction that a) focuses on monitoring and evaluation of task execution processes, b) can be applied to writing and reading tasks that rely on explicitized coding/decoding rules, and c) stimulates the abstraction of these rules in one mode to more general knowledge applicable to several modes.

The Sonnenschein & Whitehurst and the Schriver studies, which had been successful in fostering transfer, were taken as starting–points for the development of experimental instructional methods. The common factor in these studies, learning–by–observation of communication processes, is considered a possibly effective alternative to learning–by–doing. Two variants of this alternative are distinguished: observation–of–models, based on work by Bandura and Schunk, and Sonnenschein & Whitehurst; and observation–as–feedback, based on Schriver's Protocol–Aided Revision method.

Both experimental approaches may fit the above description of the sought–after instruction. It is the observer's task to monitor and evaluate the observed communication activities, not just to perform them. Moreover, the referential communication in which the writer and reader participate relies on shared coding–decoding rules, which are specially made explicit in the Sonnenschein & Whitehurst experiments ('difference rule'). And finally the abstraction of knowledge of mode–specific rules is obvious, as shown by the amount of transfer that is acquired due to observation.

7.2 An overview of the results

The research questions in this study are aimed at testing the effectivity of the experimental types of instruction. 'Effectivity' stands for 'relative effectivity', namely in relation to the learning output of a control condition based on learning–by–doing. I will present an overview of the main answers, using the numbers of the research questions, but re–ordering these questions by type of experimental instruction.

Learning–by–observation of models (one mode)

The observation of tasks performed by peers, with a view to discovering good and weak working methods, has been investigated in both experiments on argumentative text. In both cases learning–by–observation of models for one communicative role proved to be more effective than learning–by–doing–exercises (Q1). The advantage of this instructional method may be that it focuses the learner's attention at various ways in which a communicative task can be performed. The learners learn to discern effective from less effective task executions, and in looking for the difference between them they need to study the text and task demands in close detail.

The transfer effect resulting from this instructional method also exceeds transfer of learning–by–doing (Q7), with the annotation that the transfer effect from writing to reading is substantially larger than from reading to writing. It is important to understand why this transfer is asymmetrical. Is it due to a task factor 'domain' (e.g. if decoding rules would be inherently more difficult to abstract than coding rules)? Or is there some interaction with the particular 'type of instruction' (e.g. if such asymmetry occurs when learning–by–observation, and not when using other methods)? Let us explore the last possibility. An explanation for the assymetry may be that observers train at the same time the *observans* (that wat is observed: reading or writing activities) and the *observatio* (the act of observing). The *observatio* requires in itself co–reading and co–analyzing activities, both when observing reading activities as when observing writing activities. This may explain why observing writers contributes more to reading skill, than observing readers contributes to writing skill. For observers of writing, the act of observation may contribute to reading and analyzing skill; so these students learn to write and they learn to read. While for observers of reading, learning the act of observation will only yield 'more of the same', and not contribute anything to the productive mode. Obviously, further research is needed to determine what the actual causes of this asymmetry are.

Learning–by–observation of models (both modes)

The effectiveness of this method depends on the modus with which it is compared. In chapter 5 the data for this condition were compared to learning–by–doing writing exercises (DW) and to learning–by–observation of writers only (OW). No significant difference was found in the comparison to each of them (although DW and OW do differ). In chapter 6, however, observation of models (both modes) appeared to be more effective than learning–by–doing reading exercises (DR), and again no significant differenct with observation of readers only (OR). As said, the presumed extra effect of observation activities on reading skill may account for this difference.

Our expectations about the effectivity of OWR compared to DW or DR were neutral. It is important to note here that the OWR subjects observed only half the number of writing exercises as the OW or OR subjects did. The time for practice was the same, but the other half of the observed exercises were in the reading mode (they observed 1 writer + 1 reader for every exercise). So, a result in which OWR matches DW and DR should be called satisfactory. Still, the 'learning plus transfer effect' of OWR equalled DW, OW and OR, and even surpassed DR. This means that either the number of exercises has no influence on learning (which is unlikely) or that the learning effect in one mode is supported by observations in the complementary mode. Such explanations should be further investigated.

Subjects observing both writers and readers are able to follow and comment on a complete information transfer, in which they are personally not involved. This in contrast to the observation–as–feedback groups, who also follow the complete information track, and do participate in the observed communication. We will come back to these groups when discussion the difference between monitoring and evaluation activities.

Learning–by–observation as feedback

The only instructional method that has been put to the test in all three experiments is the observation of communicative partners as feedback. The feedback is aimed at evaluating a task performance that immediately precedes the feedback. Our expectations were positive, since the immediate feedback and the personal involvement made this theoretically a promising condition.

It was found in all three experiments that this feedback condition was more effective than the control group, regarding learning as well as regarding transfer. Schriver's (1992) learning theory is quite straightforward in explaining progress on the learning dimension: the writers (or readers) experience communication failures and successes and adapt their task behavior accordingly. There are some assumptions underlying this explanation. First, the observers must be able to detect the failures or successes and take them seriously; they should attribute them to their own task behavior (and not to the observed reader or writer); and they should identify elements in their task behavior (working methods, strategies) which they can label as 'successful' or 'unsuccessful' after observing their effect. This does not seem all too easy for unexperienced learners/observers, but apparently it can be done if the task and method are clear enough.

We can think of two mechanisms in which this instructional method fosters transfer: either rule–abstraction or direct imitation of the communicative partner. In any case writers, after having observed 'their' reader analyzing the text on the basis of the subject matter, are able to transform the productive rule into a receptive one. Likewise readers, after having observed how the text was conceived by the writer, are able to use such text construction activities themselves.

Learning–by–doing with and without reinforced self–reflection

We have operationalized 'reinforced self–reflection' in quite different ways. In chapter 4, we have stimulated self–reflection of writers by organizing a revision session in which they were prompted to generate comments on their first version of a text, and then were prompted to revise it on the basis of the reflection. This revision session did not contribute to the quality of the text, and neither to the quantity nor the quality of advices given in the generalisation task. It is remarkable that even purposeful and exclusive searches for flaws in one's own text do not result enhancing its quality. Apparently, writers in general do not have the tools at their disposal for effective error detection, let alone effective diagnosis and correction. This can be considered an argument for less individual, more social writing courses, in which peer or expert feedback is used as an input for revision.

In the same experiment, we tried to stimulate self–reflection with students who read and commented on a manual before they wrote one themselves. This did not contribute to the quality of the text either. Only when this 'strong orientation' was complemented with note–taking, the text quality and transfer amount would increase. It is tempting to consider the note–taking part as a *real* act of self–reflection. Since verbalizing one's experiences calls for selection, ordering and abstraction, it is not hard to assume that such activity prepares the subject for a better performance. However, it remains noteworthy that these subjects translated their *reading* experiences into working methods for *writing*.

In chapter 6, reinforced self–reflection was stimulated by having the subjects perform reading tasks while thinking aloud, as a means to supply feedback for writers, and write evaluative comments directly after. This condition is very similar to the subjects in chapter 4 who used the manual text while thinking aloud, and then wrote a summary of their comments. This type of reinforced self–reflection yielded strong effects regarding learning and transfer, so in this respect the experiments are in agreement.

Maybe it was not only the thinking–aloud activity that stimulated the reader's learning. We should not forget that this subject was in special circumstances: not just by him/herself doing a reading task aloud, but being observed by the writer of the text to be read (and by a test assistant). This can have resulted in extra high motivation, and consequently in better learning and transfer.

In sum: Learning–by–doing vs. learning–by–observation

Learning to read or write by means of peer observation seems to have the capacity to gain ground over the learning–by–doing paragdigm. In table 7.1 an overview is given over the most important research questions and the way in which they were answered in the previous three chapters. The main questions Q6 and Q10 were addressed in chapter 6, when learning and transfer effects in both modes (reading and writing) were compared. The only difference between these modes was that the OWR condition had a larger (mixed learning and transfer) effect on reading than on writing. The other comparisons between experimental and control conditions yielded the same results within the writing and the reading modes.

[hier tabel 7.1 invoegen]

Table 7.1: *Overview of main research questions and their answers.*

Other variants of the experimental instructional methods have been investigated as well, although they are not represented in this table with main research questions. In chapter 4, the effect of written comments as an addition to orientation or revision activities was striking. Reading as an orientation task alone was not effective (Q15 and Q16), however, writing down one's comments on a text after the reading and 'using' task did help to write better (Q13 and Q14) ; besides, receiving such reader's comments in addition to observing the reader helped subjects to write a better version (Q17 + Q18). Moreover, the authorship of a text did not seem to influence the effect of observing–as–feedback (Q11 + Q12): learners can just as well observe the reading of someone else's text, as a basis for revision.

7.3 Theoretical implications

What are the implications of this study's results for theory? As described in section 7.1, there are several theoretical perspectives at the basis of this study: learning–psychological theory, transfer theory, and theory on the reading and writing domains. I will consider some implications for each of them.

Learning

The learning–psychological perspective on the subject of this study was that a lack of self–monitoring and self–evaluative activities might hamper the acquisition of complex skills taught by a learning–by–doing method. Although it is impossible and undesirable for learners to be permanently occupied with self–monitoring and self–evaluation, such activities must be performed to the extent that they can

a) regulate the task execution and b) yield input for learning (cf. Ng & Bereiter, 1992; Elshout–Mohr, 1992; Zimmerman & Schunk, 1989). This supposed problem regarding learning has consequences for transfer as well: the occurrence and amount of transfer is directly dependent on the learning result.

Learning–by–observation of reading and writing activities was chosen as a possibly effective alternative to learning–by–doing. In this 'observational learning' approach (Bandura, 1977, 1986; Schunk & Zimmerman, 1994) it is assumed that observation of activities performed by others is informative in that it offers a preparation for one's own performance, and for self–observation and self–regulation of one's own behavior. Moreover, instructional methods based on observational learning have repeatedly shown to be more effective than other methods (reviewed in Schunk, 1991). To this we can add that the observational learning paradigm has slowly but steadily broadenend its interest from social learning to cogntitive, strategic and self–regulated learning (Schunk & Zimmerman, 1994). Research results acquired in the cognitive and strategic domains have also been promising.

Regarding the effectiveness of the experimental conditions in the present study, the message is evident: in almost all cases learning–by–observation offered a more effective alternative to learning–by–doing. To be precise: all comparisons between learning–by–doing on the one hand, and learning–by–observation–of–models (one mode) and observation–as–feedback on the other hand turned out to the advantage of the observation conditions. Observation–as–feedback yielded comparable results for writing instructive text as well as argumentative text (although under certain conditions; see below) which may be an indication of this effect's stability. It was also striking that observation–of–models (both modes) resulted in learning effects (probably mixed with transfer effects) which are as high as the combined effects of learning–by–doing in the reading and writing modes together. In this respect, the results confirm the effectiveness of observational learning, now applied to the cognitive domains of writing and reading. The results also confirm the value of imperfect or coping models in this educational application (see section 3.4).

However, there are still essential questions left to be answered in order to understand *why* and *under which conditions* observation activities may be more effective. The present study may yield favourable results for learning–by–observation and its variants, but this is in fact only a very global conclusion. Due to experimental control, the effects can be attributed to certain manipulated differences between the conditions. But the effects do not inform about the precise differences in psychological processes of observation, learning, or transfer that have induced the different amounts of learning and transfer. In chapters 5 and 6, I ventured to decompose the task execution processes (in tables 5.4 and 6.4) for each experimental type of practice, in order to determine to which extent the students' activities were (dis)similar. At present, the status of this decomposition is still hypothetical. No data was collected to support the view that the actual learning activities of the experimental subjects were different from the control subjects. For instance, if the experimental writing subjects really do apply knowledge about readers' needs and strategies in their writing; or if experimental readers consciously imitate the examples they have observed and positively evaluated.

Separate theoretical explanations were given for the effectivity of each type of learning–by–observation. We can see if they are in agreement with the present experiment. Sonnenschein & Whitehurst (1984) demonstrated that training in observing and commenting on other people's communicative performance would lead to mastery of the commenting task as well as the observed communicative task; while mastery of the communicative task does not guarantee mastery of the criticism task. Their explanation is related to Gagné's theory on hierarchical skills: if the superordinated skill (commenting) is trained until mastery, then the subordinated skill (performing) is automatically mastered as well. The same principle may be valid for observation of and commenting on both modes: mastering of the commenting task regarding communicative tasks in both modes leads to mastery of each separate mode. Obviously, these hierarchical rules can be applied to the learning results of the present experiments: practising by observing/criticizing a reading or writing task led to a mastery level at least equal to, and often more than, the level acquired by learning–by–doing.

Schriver (1991) offers an explanation for the effectiveness of observation–as–feedback. Writers who have the opportunity to study reader's thinking activities by protocols incorporate this knowledge into their writing process and activate it when useful, e.g. when anticipating on an audience or when revising the text with an audience in mind. This explanation is very plausible and would fit into a larger theoretical frame describing how writers and readers anticipate on their communicative partner's responses in order to enhance communication (cf. the FR and FW conditions in the argumentation experiments). Unfortunately, the evidence to confirm or refute this explanation was not collected in our study. To do so, we would need cognitive process data, such as thinking–aloud protocols, which tell us whether the experimental subjects use such 'audience orientation strategies' more than the control subjects.

The observation–as–feedback condition is an unusual extension of the 'observational learning' paradigm. Bandura (1977) makes a clear distinction between *enactive* and *vicarious* learning (section 3.4), but the observation–as–feedback method integrates these two types of learning. First, a student *performs* a (writing or reading) task with the intention to learn how to do this (the enactive part) and subsequently *observes* a complementary task (reading or writing) activity (the vicarious part). The aim of this observation is to evaluate one's own performance, and to learn about typical responses, strategies etc. of the communicative partner. Thus, the observations are not made for imitation, but to acquire knowledge about how successful communication 'works'. The distinction between enactive and vicarious learning is not as sharp as it may seem.

Theory on observational learning processes should not only describe the *psychological processes* of learning, but should also take the *conditions* in account under which such learning takes place. One such condition was investigated in chapter 4: the way in which the information acquired by observation is processed. Although all observation conditions outscored the plain learning–by–doing condition, a large extra effect was found if the observers were given a summary of comments written by the observed readers. Furthermore, the observed readers (who performed the reading task as orientation to a writing task) performed better *on the condition that* they had summarized their comments by writing them down. Not only did performance increase in these 'written comments' groups, also transfer to the 'letter of advice' task was much larger. We cannot attribute these extra effects to the single activity of writing comments, because these comments could never have been written without the prior observation/orientation task. It is the *combination* of learning activities that turned out to be effective.

This observation leads to my last point regarding learning theory. In previous research, the observational learning paradigm has shown to be useful regarding various domains: social behavior, cognitive behavior, motor–skills and self–regulation skills. It has also focused on the social and emotional factors that can account for its effectivity (Schunk, 1991). In addition, I think that structural research ought to be done into the cognitive activity 'observation' itself. What determines 'effective observation'? Which goals should be set for the observer? How should he/she process the information? Can something like 'observation skill' be developed, and how? Observation can probably be done in a good or effective way, and in a weak or ineffective way. For the development of theory on observational learning, as well as for acquiring knowledge about implementation, insight into such (in)effective factors is necessary.

Transfer

In chapter two, a *task* or *domain–oriented* approach and an *instruction–oriented* approach to transfer research were distinguished. We attempted to integrate both lines of research in this study.

The relationship between the domains of reading and writing was discussed in section 3.2. Three concurring models of this relationship were presented, with different implications for the possibility of transfer. By collecting evidence in support of each unilateral model, researchers have in fact jointly collected support for the rejection of unilaterality and opened the way for a bilateral or interactive model of the reading–writing relationship. The interactive model predicts the possibility of bilateral transfer. The results of the present study also yield support for the interactive model. If we grant the theoretical possibility of transfer from either mode to the other, then we may attribute the extent to which transfer is actually obtained to a) the choice of tasks and b) the instructional design. Other explanatory variables are also possible (subject factors come to mind) but are not addressed in this study.

For the experimental chapters 5 and 6, we selected reading and writing tasks that were related, in that their performance can be based on a common knowledge element about argumentative language. A task–oriented approach, such as advocated by Singley and Anderson (1989) requires stipulation of these 'common elements' between tasks or task domains. Describing cognitive task elements in such a way that the relation between different tasks becomes explicit, is far from easy. In this study I could only give a first impulse to describing such elements in terms of production or condition–action rules. There is no commonly accepted language to be used for this purpose, and procedures for testing the validity of the task analyses were not yet applied. Therefore, the validity of the analyses given here can be questioned. They have in any case been instrumental in modifying and formulating the reading and writing tasks that functioned as 'mirror images' in the experimental courses. The usability of production rules for the description of reading and writing tasks needs further study.

In chapter 4, not intermodal transfer was assessed, but an important condition to inter– or intramodal transfer: the extent to which declarative, non–task–specific knowledge was acquired (knowledge about criteria for 'writing good manuals'). Singley & Anderson's transfer theory (1989) explains transfer of cognitive skill by mental processes of abstraction and generalization, resulting in more versatile procductions. Therefore we assumed that the extent to which subjects could generalize across this particular writing experience might indicate the possibility of transfer. The validity of this assumption needs to be checked; preferably by realistic transfer tasks.

We also tried to promote transfer (of the specified 'common elements') by effective instructional design. Of the two 'roads' by which learners may transfer their knowledge (Salomon & Globerson, 1987; Salomon & Perkins, 1990), we decided to stimulate the 'high road', or the road of 'mindful abstraction'. We hypothesized that the experimental instructional methods (observation–of–models and observation–as–feedback) would enhance abstraction of the context–bound coding and decoding rules, in comparison with learning these rules by doing exercises. This choice was not obvious, since the 'observational learning' or 'social cognitive' theory has until now not paid much attention to transfer effects. Why would learning–by–observation enhance transfer?

First, we suggested that observation–of–models may yield a comparably large psychological distance to the observed activity (one of Bandura's distinctions between 'enactive' and 'vicarious' learning), resulting in a more abstract coding. In new task situations, it may be easier to identify other people's approach and to evaluate its result, then to identify one's own approach during execution. 'Active and mindful abstraction of the specific task situation', a condition that Salomon & Perkins (1990) put to high road transfer, may be fulfilled with less effort during observation from an 'outsider' perspective than during a personal involvement in task execution. A last reason why we chose for this type of instruction was the evidence that had been presented by Sonnenschein & Whitehurst (1984): their method of observation also led to high inter–modal transfer.

Second, it is easier to account for inter–modal transfer promoted by observation–as–feedback. Observers perform a task in one mode, and observe a related task in the complementary mode. The observations will most likely have some preparatory value for their performance in the other mode. A second explanation is, that the coding/decoding rule is abstracted when its use in the complementary mode is observed, even though in a reverse order. As an example, one can think of a learner who applies the following coding rule in an exercise: 'when I must write a paragraph, I must include the topic in the first sentence'. Consecutively a reader is observed who searches paragraph topics in the first sentences, because he has learned 'when I must understand the contents of a paragraph, I must first check for the topic in the first sentence. The writer may adapt his cognition of the rule to a more general, i.e. text–centered and mode–independent 'the first sentence of a paragraph should include the topic'.

The transfer effects observed in the present study were generally large. Several factors may account for this unexpectedly large transfer. First, the learning and transfer tasks were structurally highly related. Second, that these tasks were related was also evident by the fact that they were presented in one experimental setting (contextual cues). Both factors are known to have positive influence on transfer (Klausmeier, 1975; Cormier & Hagman, 1987; Simons & Verschaffel, 1992). The strucural similarity between the tasks was large because they were specially constructed for transfer to occur, and because they were simple and directed at clear goals (rather convergent than divergent tasks), so they could not differ on many distracting aspects. For instance, it is not difficult to perceive similarity between the tasks 1) 'write an argumentative text consisting of one standpoint and three multiple arguments', and 2) 'analyze the following text's argumentation structure' (containing one standpoint and three multiple arguments). As concerns the 'physics manual', the subjects will have easily recognized the relevance of the reading or writing task for the 'letter of advice'. However, we should not focus on the fact *that* transfer occurred, but on the differences in *how much* transfer occurred, due to systematic differences between the groups.

The learning–by–observation conditions have produced far more transfer than the learning–by–doing conditions.. The factors that explained this generally high transfer (task similarity and contextual cues) were equal for all participants, so these differences should be attributed to differences in instruction. For the experimental groups, there are two possible sources for these higher transfer effects: learning had also been higher in the experimental conditions (and transfer depends on prior learning) and the instructional method may be more supportive of transfer. It is not possible to define which part of the effect is due to which source. It is in any way evident that the type of instruction is responsible for the higher learning effects as well as for the higher transfer effects. It is another argument for the idea that instruction should be purposefully designed in order to promote transfer; selecting adequate tasks is not enough to acquire optimal transfer.

Finally, it is important to note the difference in operationalization of the concepts 'learning' and 'transfer' in chapter 4 and in chapters 5 and 6. In chapter 4, 'learning' is operationalized as the level up to which the student, without external help, can bring his performance by taking advantage of the learning environment, as seen in relation to a first trial. In chapters 5 and 6, learning is operationalized as the score on a set of posttests, which contain similar, but not identical tasks as the training tasks. The difference with respect to transfer operationalization is even stronger. In chapter 4, transfer is operationalized by the declarative knowledge that a student gains as a result of the learning activities. This is in line with the theory on skill acquisition and skill adaptation that Singley & Anderson (1989) offer: a declarative stage always precedes the compilation of new productions. The declarative knowledge was regarded a precondition to real transfer. In chapters 5 and 6 however, transfer is assessed in its most genuine form: by having the student perform a transfer task. This is not by 'showing' his knowledge, but by using it in adequate situations.

Reading and Writing

Since theorists and researchers have started to conceptualize reading as an active, knowledge–

constructing behavior, and to model writing as a recursive process including many re–reading and mentally representing activities, the similarity between reading and writing has received increasing attention (Wong, 1991). The interest focused on a number of questions, such as: how should we describe the similarity of reading and writing? how can we use this similarity in education? can we teach one skill by instructing the other?

Earlier in this section I have already mentioned an implication of this study for theory on the reading–writing relationship. This study offers more support for Shanahan & Lomax's interactive model, which opens the possibility for transfer effects from reading to writing and vice versa (Shanahan & Lomax, 1986). It is important to investigate the nature of this reciprocal relationship in order to find possibilities and limitations to taking advantage of itin instruction.

According to Shanahan & Lomax (1986), the variability of research findings on the nature of the relationship between reading and writing can be accounted for by the absence of detailed descriptions of which parts of the reading and writing process may influence each other. They posited a model (fig. 3.1 and 3.2) in which the writing and reading processes are divided in smaller functional parts, so that more detailed interrelationships could be specified and tested (section 3.2).

Researchers have often measured the effect of a particular aspect of reading on a very global measure of writing, or vice versa. Such effects (for instance, the influence of learning list with difficult words on essay composition) are usually small or absent, and cannot indicate the direct influential effect of two particular components of reading and writing. Too global measurement may undestimate, and even misrepresent the transfer potential. The fact that the transfer results of this study are comparably large can be considered a consequence of our close tuning of the reading and writing tasks. The more similar the reading and writing tasks, the nearer the transfer, and the easier to actually obtain it.

Besides, 'it is not clear whether reading's influence on writing is stronger than writing's influence on reading' (Gleason, 1995). The question Gleason puts about the preferred direction of transfer can be answered by comparing the transfer amounts regarding reading (caused by writing observation) and writing (caused by reading observation). In chapter 6, we found that transfer from writing to reading was the easier to obtain than transfer from reading to writing. Also, the effect of OWR (observing both modes) on reading skill is much higher. In all, writing practice or observation turns out to be more effective for reading skill, than reading practice is for wrting.

Thus the following three implications of the present study on reading and writing theory can be given. First, it is confirmed that reading and writing are connected skills. Second, the connection can be used to promote transfer, by means of selecting and explicitizing the reading and writing components and by tuning instruction toward the transfer–goal. Third, the relation between the skills is not balanced: writing practice influences reading more than reading practice influences writing.

Finally, many writing theories hold that revision is an important tool in the writing process, and in the writing–learning process (see for this distinction: Oostdam & Rijlaarsdam, 1995). In chapter 4, a quite extensive revision session was added to the learning–to–write–by–doing condition (condition I). In spite of prompts and requests for re–reading, this revision activity did not result in a higher quality of writing or a better knowledge, nor in better learning. The educational value of revision is apparently modified by other factors.

7.4 Future research

Judging from the results obtained in this study, learning–by–observation may be a promising method for instruction for reading and writing skills. The road from this point to implementation is, however, long and probably winding. I will give some possible directions for future research. Some of them concern new research questions based on this study, others concern remedial to validity or interpretation problems in the reported experiments.

What are the specifics of effective observation?

We need to know more about what distinguishes 'effective' from 'ineffective' observation. I see two sides to this question: a) what is the object of effective observation? and b) what are the mental activities during effective observation?

Objects for effective observation

Observation can be aimed at different objects: at task execution processes, or at the resulting products. What is needed is a more thorough examination of the process–product relationship within the framework of learning–by–observation. During observation, the observer's focus of attention ought to be at the process–product relationship: which processes (or working methods, activities, strategies) yield good products, which processes lead to failures? We found that it is very tempting for some students to neglect the process part, and only look at the resulting products. Their intention seemed to be to give a 'good answer' and then continue with the next 'observation'. Athough in our theoretical perspective, observation of the process should be indispensible (because it informs how to walk the road from task orientation to task completion, and about the time dimension in the workin method), it is of course useful to know how much of the experimental effect will remain if, for instance, only products, or only processes are observed.

Another type of 'objects' that may influence the effectivity of the observations, are the type and quality of the observed models or communicative partners. In the present study, the observed models were equal for most conditions, except for the 'FW' condition in chapter 5 (one proofreader was assigned to two writers). The quality of the feedback was therefore variable. It is theoretically interesting, and educationally relevant, to know how this quality influences learning of the observers. 'Good readers' may fill in communicative gaps in the texts that they read; for instance, they may not need connectives to understand the relation between two sentences. Their feedback would be less informative in this respect than feedback of readers who do need such connectives. On the other hand, weak readers may not be able to detect comprehension problems during reading, leaving the observer in the conviction that there are no such problems.

Mental activities for effective observation

Perhaps the most important factor determining the effectivity of observation is the type of thinking activities performed by the observer. For learning–by–observation–of–models, the thinking activities have been specified as process–observation and process– and product–evaluation. For learning–by–observation–as–feedback, they have been specified as process–observation, comparing and self–evaluation. Separating monitoring and evaluative activities (letting the subjects only monitor without explicit evaluation) would give insight in the relationships between these activities, and give information about their respecitive contribution to the overall–effectivity.

The experimental effects are attributed to learning activities that have been theoretically specified: monitoring, evaluation, orientation. It is a serious problem in this study that no backup can be provided for these attributions. Observational learning theory (Bandura, Schunk) as well as developmental theory (Sonnenschein & Whitehurst) and communication–psychological theory (Schriver) have provided for the hypotheses around which this study was built, but we cannot confirm yet that the learning or thinking processes took place in the way described by these theories. In this study, only pre– and posttest data was collected from the students. They cannot support the view that the effects were indeed caused by a more self–observing and self–evaluative behavior. Neither can we combine the results with data on students' thought processes while reading or writing texts. We consider this the major handicap of this study and at the same time the most interesting point of attention for more profound follow–up studies.

Which students apply for learning–by–observation?

A factor concerning the external validity of this study is the age–group of the subjects. The ability to perform the type of thinking activities during observation (orientation, monitoring, evaluation, reflection) may be related to age. The same can be said about the ability to perform cognitive tasks while thinking aloud. Younger students – or generally students of lower developmental level – may not be able to use the observations to their advantage because they fail to see the instructive quality of their observations. On the other hand, older students may profit even more from observation if the quality of their monitoring, evaluating etc. is more instructive to them.

Personal factors, such as motivation or learning style, may also influence the effectivity of learning–by–observation methods. For some students, receiving 'live' feedback on their performance may not be very inviting. For others, the act of thinking–aloud may be scary. It would be educationally relevant to know which personality traits interact with the instructional methods.

Which tasks or skills can be acquired by learning–by–observation?

Two different text types were used in the three experiments reported here. 'Instructive texts' were chosen because they provoke visible responses from the reader, who thought–aloud. The observing writers could interprete these responses comparably easily as 'correct' or 'incorrect'; thus they informing them about the comprehensibility of the text. 'Argumentative texts' provoked less visible responses, in that no tools had to be manipulated or experiments to be performed. However, by asking the reader to not only read the text, but to perform some kind of analysis task while thinking–aloud, sufficient information about the readers thinking activities could be given to the observer.

To validate and extend the learning–by–observation method, it should be examined with many more text types: formal letters, explanations, and in a later stadium maybe also short stories and essays. It seems important that the text is well–structured. The method may also be applied to verbal tasks: observing peers performing speaking tasks.

What will the long–term effect of learning–by–observation be?

A possible threat to the validity of this research is the comparable novelty of the experimental interventions, compared to the control condition. This novelty offers an important alternative explanation, since it is shared by all experimental conditions which show an advantage to the learning–by–doing conditions. Longer lasting studies, during which the subjects will get used to the observation activities, may answer the question to which extent the novelty threatens the validity of the results.

7.5 Practical implications – and an extra

The scientific study of education has a practical goal; maybe not in the first instance, but it is driven by the motivation to understand and enhance the relation between teaching and learning.

The present study is aimed at enhancing learning and transfer of reading and writing education. If there is yet a practical implication to be derived, it is the recommendation to start experimenting with observation activities in the language skill curriculum. However, due to the early stage of this research into learning–by–observation, there are many reservations to such a recommendation. They concern the curriculum structure, the nature of the observation activities, the selection of tasks, and the applicability for student audiences.

An important implication of the present study for education is, that the curricula for reading and writing would have to become more co–ordinated. As was stated in chapter 1, the whole of reading and writing education is hardly ever more than the sum of its parts: namely reading education and writing education. When texts and task types for reading and writing are placed more in line, and are explained and practiced at the same time, even automatic transfer would become more likely. There are many curricular elements for reading and writing that are related; it would be a missed opportunity to not connect them.

It is important to note that this experimental method works best if the reading and writing tasks are clear and – certainly in the beginning – the number of possible answers is limited. The students should experience (vicariously) how a certain coding or decoding rule is used by others, and then apply this experience to themselves. If the task is not very demanding, the students can do more trials on the cognitive element within the same time. This may be desirable, depending on the level of the students.

Another implication is that the succesful learning–by–observation (models or feedback) conditions call for some logistic interventions in the classroom. It is certainly not necessary to reconstruct the experimental setting used in the reported experiments. The teacher can, as an expert–

model, demonstrate reading or writing processes in front of the classroom. Or several students (coping models) ca take his place. The type of instruction should be designed such that the students are mainly occupied with alternately reading or writing, and observation and evaluation. Some examples are given below.

It is not the aim of this study to disqualify learning–by–doing as invalid. This method deserves a prominent place in every kind of skill instruction. Nevertheless, learning–by–doing and other methods should maintain a better balance in complex, long–term curricula. Reading and writing instruction should not rely solely on a learning–by–doing approach to instruction, but should be supported by other, for instance observational learning activities.

In all, this study offers a few steps on the road towards development, validation and implementation of a new instructional method for the reading and writing domain. There are many more questions left to be answered about learning–by–observation. The most important concern the specific characteristics of effective observation, the type of reading and writing tasks to which the method can be applied, and the type of students that would benefit from the learning–by–observation method.

Fortunately, teachers are impatient and cannot wait to see these questions answered. Therefore, I include two samples of lessons which might be used as a first step towards observational learning in the classroom. Also without a videoplayer in the classroom, teachers and their students can start to work on 'observation of models' or 'observation as feedback' in their language arts program.

Lesson 1: Instructive Texts

Goal Learning to write instructive texts (manuals, guides for instruction, assignments, recipes etc.)

Method A

preparation

Invent two writing assignments for an instructive text. Examples: guide to the train schedule, guide to the thematic catalogue in the school library. The assignment should not be too easy.

1st lesson

The class is halved: each half makes one of the assignments. In this way, each student writes a 'first version' of the text.

2nd lesson

Students are matched to a partner: a 'train'–student gives his text to a 'catalogue'–student and vice versa.

One of the two students follows the instructions in the text of the other student. While doing the task, he thinks aloud about the thinking steps he makes and about what he does and does not understand in the text. (You should demonstrate this in the class first.) The writer of the text takes notes while observing the reader, but is not allowed to give comments or explanations.

Next, the students change roles.

Homework is: a) making a list of proposed improvements for the first version, and b) revising the first version.

3rd lesson

The 'train'–students work in groups and make an inventory of the improvements that are proposed. They use their homework lists for this. The 'catalogue'–students do the same for their text.

The results are discussed in the class.

In this way, students acquire for each writing assignment a list of 'experimental points of attention'; apart from that, each student has once written, once observed, once revised, and once reflected.

Method B

If you consider the thinking–aloud by the students as too risky, you can also act yourself as the thinking–aloud reader/user of the instructive texts. You do this in the classroom, using e.g. two students' texts for each assignment. You deliberately make mistakes if unclear instructions allow you to. This method requires that you are very well aware of possible and imaginable 'traps' in the texts.

After the students have listened to you in you reader's role, they try to improve their texts and make a list of the improvements made. Alternatively, you can invite some students in front of the classroom to read the texts aloud, use them for what they are meant and give as many comments as possible.

Requirements

3 lessons

photocopied material of both writing assignments

copied matierial from a train schedule

access to the school library

Lesson 2: Argumentative Texts

Goals Learning to write argumentative texts in a 'reader oriented' way; learning to read argumentative texts in a 'writer oriented' way.

Method A

Remark: The methods presented here are not specific for argumentative texts: every text type can be practiced in this way. The essence is to observe other students during their execution of a writing or reading task; the observing students must concentrate on the evaluation of this execution. It is very important that the criteria he or she uses are clearly defined and recognizable.

Preparation

Subject–matter about the construction of argumentative texts must be explained. (It is preferable to use a text scheme such as the one we used in our argumentative courses: see lesson 2 in table 5.3.) You invent two writing assignments in which the students must use the particular structure. Examples: an argumentative essay about 'motorized traffic in the city', and one about 'school uniforms: history or future?'. The assignment should clearly state that all parts of the text structure must be present and recognizable.

The writing–lesson

The class is halved again. Each half makes one of the assignments. In this way, each student writes a 'first version'.

The feedback–lesson

Students are matched: a 'traffic'–student gives his text to a 'uniform' student and vice versa. Half of the students are now instructed to analyze the received text while thinking–aloud, preferably with the use of the text scheme (you must decide whether the students can have the scheme at hand or not). While working on this assignment, the analyzing student tells the writer which thinking steps he makes and which aspects of the text are difficult. The writer takes notes while observing the reader, but is not allowed to give comments or explanations. Next, the students change roles.

For homework and the 3rd lesson: see lesson 1

Required 3 lessons; photocopies of both writing assignments, textual schema

Method B

Research seems to indicate that observing writers is more efficient than observing readers, because the transfer from writing to reading is higher. In a class that is used to work in groups, you can experiment with the next idea:

You prepare three short writing assignments, for which the students must use subject–matter that has just been treated. Along with the assignments, you supply a list with explicit criteria (e.g. in the form of questions).

The class is divided in small groups of three students. Taking turns, the group members perform a writing task, while thinking aloud. The other s must observe and take notes (evaluative comments); they can use your list of criteria. Only when all three students have 'written–aloud', the writers compare their comments; first within the group, and later in the class. (See evaluative criteria at the bottom of the next page.)

Requirements

– (depending on the size of the writing assignments:)

1 hour writing – 1 hour commenting and revision

Sample writing assignments

Assignment 1:
Write a short argumentative text in which you defend a positive standpoint about the question whether it is better to be a student in a large school than in a small school.
You are writing for an audience of opponents, who think a small school is more sociable.
Use the arguments you find below, make sure a reader would have little trouble recognizing the given structure in your text. Pay attention to your opponent's argumentation.

(1)
it is better to be student in a large school than in a small school.

(1.1)
more choice in friends

(1.2)
better accomodations

(1.3)
chance of having to leave the school is comparably small

(1.2.1)
school–library has more books

(1.2.2)
more and larger sports–fields

(1.3.1)
you can stay for lower, general and higher secondary education

Assignment 2:
Write a short essay in which you defend a positive standpoint regarding the issue if Suzan should come along for holiday.
Your friend does not want to take Suzan along. She thinks that your group of three is alread big enough.
Use the arguments that you find below. Make sure a reader would have little trouble recognizing the given structure in your text. Pay attention to the argumentation of your opponents.

(1)
Ask Suzan to join in for the holiday

(1.1)
knows a lot about France

(1.2)
is nice company

(1.3)
hasn't been on holiday for a long time

(1.1.1)
has worked as au–pair in Bordeaux

(1.1.2)
has often been to France on a holiday.

(1.2.1)
going out with her is a whole night of fun.

Possible criteria for the observers to comment on:

a) does the writer, before he starts arguing, make clear in an introduction what the discussion is about?
b) does the writer state his own standpoint in a clear way?
c) does the writer link the argumentation to the standpoint using the correct connectives? (because, and, also, firstly, secondly, etc.)
d) are 'branches' of argumentation kept together in paragraphs?
e) does the writer reply adequately to arguments of opponents?
f) does the text end with a conclusion?

Table 7.1: *Overview of main research questions and their answers.*

INSTRUCTIONAL METHOD:	LEARNING GOAL: (CH. 4) learning to write instructive text	(CH. 5) learning to write argumentative text	(CH. 6) learning to read argumentative text
Q5: learning–by–doing with prompted self–reflection	_		+
Q9: transfer after learning–by–doing with prompted self–reflection	+/–		+
Q2: learning–by–observation as feedback	+	+	+
Q8: transfer after observation–as–feedback	+/–	+	+
Q1: learning–by–observation of models (one mode)		+	+
Q7: transfer after observation of models (one model)		+	+
Q3: learning–by–observation of models (both modes)		_	+
Q4: observation (both modes) vs. (one mode)		_	_

\+ = significant difference found, when compared to learning–by–doing only

_ = no significant difference found when. compared to learning–by–doing only

+/– = only if supported by written comments after the observation

Caveat lector: + and – should not be confused with positive and negative effects, or with positive and negative transfer.

References:

Ali, K. (1990)
Instructiestrategieën voor het activeren van preconcepties. [Instructional strategies for the activation of preconceptions.] Dissertation. Tilburg: Katholieke Universiteit Brabant.

Anderson, J.R. (1983)
The architecture of cognition. Cognitive Science Series, 5. Cambridge: Harvard University Press.

Anderson, J.R. (1987)
Skill acquisition. Compilation of weak–method problem solutions. Psychological Review, 94, 192–210

Anderson J.R. (1990)
Cognitive psychology and its implications. (3rd ed.) New York: Freeman.

Anderson, J.R. & R. Thompson (1990)
Use of analogy in a production system architecture. In A. Ortony et al. (eds.), Similarity and Analogy. Cambridge: Cambridge University Press.

Angell, J.R. (1908)
The doctrine of formal discipline in the light of the principles of general psychology. Educational Review, 36, 1–14.

Austin, J.L. (1976)
How to do things with words. 2nd edition. London: Oxford U.P.

Ausubel D.P. (1968)
Educational Psycholoy: A Cognitive View. New York: Holt, Rinehart & Winston.

Baldwin, T.T. & J.K. Ford (1988)
Transfer of training: A review and directions for future research. Personnel psychology, 41, 63–105.

Bandura, A. (1977)
Social Learning Theory. Englewood Cliffs, NJ: Prentice Hall.

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

Bartlett, (1982)
Learning to revise: Some component processes. In: M. Nystrand (ed.) What writers know: The language, process and structure of written discourse. New York: Academic Press.

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

Bonset, H., De Boer, M. & Ekens, T. (1992)
Nederlands in de basisvorming. Een praktische didactiek. [Dutch in basic secondary education. A practical pedagogy]. Muiderberg: Coutinho.

Braet, A. & R. Schoonen (1994)
Adviezen van leerlingen over het schrijven van opstellen. [Students giving advice on writing essays] Tijdschrift voor Taalbeheersing, 15, 1–23.

Broad, M.L. & J.W. Newstrom (1992)
Transfer of training: Action–packed strategies to ensure high payoff from training investments. Reading: Addison–Wesley.

Brooks, L.W. & D.F. Dansereau (1987)
Transfer of information: An instructional perspective. In: S.M. Cormier & J.D. Hagman (Ed.), Transfer of Learning. Contemporary Research and Applications. The Educational Technology Series. Academic Press, San Diego (etc.) 1987.

Bruffee, K.A. (1985)
A short course in writing. (3rd ed.), Bosten: Little–Brown and Co.

Butterfield, E.C., T.A. Slocum and G.D. Nelson (1993)
Transfer as Learning. In: Detterman & Sternberg (eds.) Transfer on Trial Norwood N.J.: Ablex.

Carroll, W.R. & Bandura, A. (1982)
The role of visual monitoring in observational learning of action patterns: Making the unobservable observable. Journal of Motor Behavior, 14, 153–167.

Cecco, J.P. & W.R. Crawford (1974)
The psychology of learning and instruction. Educational Psychology. Second Edition. Englewood Cliffs, N.J.: Prentice–Hall.

CITO (1991)
Reultaten van de prestatiepeiling taalvaardigheid. [Results of a performance assessment of language skills.] Arnhem: CITO.

Clark, R.E. & A. Voogel (1985)
Transfer of training principles for instructional design. Educational Communication and Technology Journal, 33, 113–123.

Combs, W.E. (1975)
Some further effects and implications of sentence–combining exercises for the secondary language arts curriculum. Doctoral Dissertation. University of Minnesota.

Cormier, S.M. & J.D. Hagman (1987)
Introduction. In: S.M. Cormier & J.D. Hagman (Ed.), Transfer of Learning. Contemporary Research and Applications. Educational Technology Series. San Diego: Academic Press.

CVEN (1992)
Definitieve voorstellen Commissie Voorbereiding Eindexamenvernieuwing Nederlands. Den Haag: SDU.

DAT (1984)
Differentiële Aanleg Test. Lisse: Swets & Zeitlinger.

De Corte, E. (1987)
Introduction. International Journal of Educational Research 11, 603–607.

De Corte, E., L. Verschaffel & H. Schrooten (1990)
Cognitieve effecten van het systematisch leren programmeren in LOGO: een onderwijsexperiment bij zesdeklassers. [Cognitive effects of systematically learning to program in LOGO: An instruction experiment with sixth graders.] Pedagogische Studieën, 67(7), 293–305.

De Glopper, K. (1988)
The results of the international scoring sessions. In: T.P. Gorman, A.C. Purves & R.E. Degenhart (eds.) The IEA Study of Written Composition I: The International Writing Tasks and Scoring Scales. Oxford: Pergamon Press.

De Jong, F.P.C.M. (1992)
Zelfstandig leren. Regulatie van het leerproces en leren reguleren: een procesbenadering. [Independent learning. Regulation of the learning process and learning to regulate: a process approach.] Published dissertation, Tilburg University, The Netherlands. Tilburg: De Jong.

De Jong, F.P.C.M. & Van Hout–Wolters, B.H.A.M (1993)
Knowledge construction, self–regulation and process–oriented instruction. In: B. Van Hout–Wolters & F. De Jong (Eds.), Process–oriented instruction and learning from text. (...)

Detterman, D.K. & R.J. Sternberg (ed.) (1993)
Transfer on Trial: Intelligence, Cognition and Instruction Norwood N.J.: Ablex.

Detterman, D.K. (1993)
The Case for the Prosecution: Transfer as an Epiphenomenon. In: Transfer on Trial: Intelligence, Cognition and Instruction Norwood N.J.: Ablex.

Edwards, A.L. (1985)
Multiple regression and the analysis of variance and covariance. New York: Freeman & Co.

Ellis, H.C. (1965)
The transfer of learning. New York: Macmillan.

Elshout, J.J. (1966)
Conclusies III. [Conclusions III.] Amsterdam: Psychological Laboratory. University of Amsterdam.

Elshout, J.J. (1976)
Karakteristieke moeilijkheden in het denken. (Characteristic Difficulties in Thinking). Dissertation. Amsterdam: University of Amsterdam.

Elshout, J.J & Veenman, M.V.J. (1987)
The relation between intellectual ability and working method as predictors of learning. Journal of Educational Research. 85, 134–143.

Elshout–Mohr, M. (1992)
Metacognitie van lerenden in onderwijsleerprocessen. [Metacognition of learners in teaching–learning situations.] Tijdschrift voor Onderwijsresearch, 17, 273–289.

Faigley, L., R.D. Cherry, D.A. Joliffe and A.M. Skinner (1989)
Assessing writer's knowledge and processes of composing. Norwood, NJ: Ablex.

Fish M.C. & Pervan, R. (1985)
Self–instruction training: A potential tool for school psychologists. Psychology in the schools, 22, 83–92.

Flavell, J.H. (1985)
Cognitive Development. (2nd ed.) Englewood Cliffs, NJ: Prentice Hall.

Fleishman, E.A. (1987)
Foreword. In: S.M. Cormier & J.D. Hagman (Eds.), Transfer of Learning. Contemporary Research and Applications. Educational Technology Series. San Diego: Academic Press.

Flower, L., Hayes, J.R., Schriver, K.A., Carey, L. & Stratman, J. (1986)
Detection, diagnosis, and the strategies of revision. College Composition and Communication, 37, 16–55.

Gagné R.M. (1973)
The conditions of learning. Second Edition. London: Holt, Rinehart & Winston

Gagné, R.M., Briggs, L.J. & Wagner, W.W. (1988)
Principles of instructional design. Third Edition. New York: Holt, Rinehart & Winston.

Gagné, R.M, Forster, H. & Crowley, M.E. (1948)
The measurement of transfer of training. Psychological Bulletin, 45, 97–130.

Gagné, R.M, Paradise, N.E. (1961)
Abilities and learning sets in knowledge acquisition. Psychological Monographs, 75 (14).

Gick, M. & K. Holyoak (1987)
The cognitive basis of knowledge transfer. In: S.M. Cormier & J.D. Hagman (Ed.), Transfer of learning. Contemporary research and applications. Educational Technology Series. San Diego: Academic Press.

Glaser, R. (1990)
Toward new models for assessment. International Journal of Educational Research, 14(5), 475–483.

Gleason, M.M. (1995)
Using direct instruction to integrate reading and writing for students with learning disabilities. Reading & Writing Quartely: Overcoming learning difficulties. 11, 1, 91–108.

Glopper, K. de (1986)
Onderzoeksdossier Project Internationale Vergelijking van het Stelvaardigheidsonderwijs. [Research Dossier Project International Comparison of Writing Education.] Amsterdam: SCO.

Grice, P. (1975)
Logic and Conversation. In: Cole, P., & Morgan, J.L. (eds.) Syntax and semantics 3: Speech acts. New York: Academic Press.

Hall, M., Moretz, S.A. & Statom, J. (1976)
Writing before grade one – a study of early writers. Reading Teacher, 31, 582–585.

Hayes J.R. & Flower, L. (1980)
Identifying the organization of writing processes. In L.W. Gregg & E.R. Steinberg, Cognitive Processes in Writing,. (pp. 3–30) Hillsdale, NJ: Erlbaum.

Hayes, J.R., L.S. Flower, K. Schriver, J. Stratman and L. Carey, (1987)
Cognitive processes in revision. In: S. Rosenberg (ed.) Advances in applied psycholinguistics, Volume II: Reading, writing and language processing, pp. 176–240. Cambridge, England: Cambridge University Press.
Hayes, J.R. (1992)
Naar een nieuw model voor teksproduktie. [Towards a new model for text production.] Tijdschrift voor Taalbeheersing, 14,

Hillocks (1986)
Research in written composition: New directions for teaching. Urbana, IL: ERIC Clearinghouse on Reading and Communication skills and NCRE.

Inspectierapport PABO (1988)
Nederlandse Taal en Letterkunde op de PABO. Inspectierapport Hoger Onderwijs. Zoetermeer: Ministerie van Onderwijs.

Jelsma, O. (1989)
Instructional Control of Transfer Dissertation Twente University. Enschede: Bijlsma & Van Merriënboer.

Judd, C.H. (1908)
The relation of special training and general intelligence. Educational Review, 36, 28–42.

Just, M.A. & Carpenter, P.A. (1987)
The psychology of reading and language comprehension. Newton: Allyn & Bacon.

Katona, G. (1940)
Organizing and Memorizing. New York: Columbia University Press.

Klausmeier, H.J. (1975)
Learning and human abilities. Educational Psychology. 4th Edition. New York: Harper & Row.

Koetsenruijter, W. & Slot, P. (1990)
Het schrijven van betogen: Handleiding voor het opstellen van argumentatieve teksten. [Writing argumentative texts. A guide for the composition of argumentative texts.] Leiden: Martinus Nijhoff.

Kucer, S.L. (1985)
The making of meaning: Reading and writing as parallel processes. Written Communication, 2, 317–336.

Kuhl, J. & Kraska, K. (1989)
Self–regulation and meta–motivation: Computational mechanisms, development, and assessment. In R. Kanfer, P.L. Ackerman & R. Cudek (eds.) Abilities, motivation, and methodology. Hillsdale, NJ: Lawrence Erlbaum.

Lannon, J.M. (1985)
Technical writing. (3rd ed.) Boston, MA: Little–Brown and Co.

Larkin, J.H. (1989)
What kind of knowledge transfers? In: L.B. Resnick (Ed.), Knowledge, learning and instruction. Essays in honor of Robert Glaser. Hillsdale N.J.: Lawrence Erlbaum.

Levelt, W. (1989)
Speaking; From intention to articulation. Cambridge, MA: MIT Press.

Mayer, R. & J. Greeno (1972)
Structural differences between learning outcomes produced by different instructional methods. Journal of Educational Psychology, 67, 331–350.

Mayer, R.E. (1983)
Thinking, problem solving, cognition. New York: Freeman

Mayer, R. (1987)
Educational psychology: A cognitive approach. S.L.: Harper Collins.

McKeachie, W.J. (1987)
Cognitive skills and their transfer: A discussion. International Journal of Educational Research, 11, 707–712.

McNeil, J.D. (1987)
Reading comprehension: New directions for classroom practice. (2nd ed.) Glenvieww, IL: Scott, Foresman & Company.

Meichenbaum, D. & Goodman, J. (1971)
Training impulsive children to talk to themselves: A means of developing self–control. Journal of Abnormal Psychology, 77, 115–126.

Meuffels, H.L.M. (1982)
Studies in taalvaardigheid. [Studies in language skill.] Dissertation University of Amsterdam, Amsterdam.

Miller, B.D. & Ney, J.W. (1968)
The effect of systematic oral exercises on the writing of fourth–grade students. Research in the Teaching of English, 2, 44–61.

Newell, A. & H. Simon (1972)
Human problem solving. Englewood Cliffs, N.J.: Prentice–Hall.

Ney, J.W. (1966)
Applied linguistics in the seventh grade. English Journal, 55, 895–897.

Ng, E. & Bereiter, C. (1992)
Three levels of goal–orientation in learning. The Journal of Learning Sciences, 1, 243–273.

Nijhoff, W.J. & J.L.M. Remmers (1990)
Basisvaardigheden en het beroepsonderwijs ('Transferable skills: toward a definition and operationalisation'). Pedagogische Studieën, 67, 9, 389–402.

Odlin, T. (1987)
Language transfer. Cross–linguistic influence in language learning. The Cambridge Applied Linguistics Series. Cambridge: Cambridge University Press.

Oostdam, R.J. & M.J. Couzijn (1989)
Instrumenten voor het peilen van receptieve argumentatievaardigheden. [Instruments for the assessment of receptive argumentation skills.] In: Tijdschrift voor Taalbeheersing, 10, 4

Oostdam, R.J. (1991)
Argumentatie in de peiling. [Assessing Argumentation.] Amsterdam: SCO. Diss. UvA.

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

Osgood, C.E. (1949)
The similarity paradox in human learning: A resolution. Psychological Review, 56, 132–143.
Otter, M.E. (1987)
Tijdbesteding aan moedertaalonderwijs op vier manieren gemeten. [Time spent on mothertongue education measured in four ways.] Universiteit van Amsterdam: Stichting Centrum voor Onderwijsonderzoek. Rapport 131.

Palincsar, A.S. & Brown, A.L. (1984)
Reciprocal teaching of comprehension–fostering and comprehension–monitoring activities. Cognition and instruction, 1, 117–175.

Paris, S.G. & Byrne, J.P. (1989)
The constructivist approach to self–regulation and learning in the classroom. In B.J. Zimmerman & D.H. Schunk (eds.), Self–regulated learning and academic achievement: Theory, research and practice. New York: Springer.

Parreren, C.F. van (1978)
Psychologie van het leren 1. [Psychology of learning 1.] Psychologische Monografieën. Deventer: Van Loghum Slaterus.

Pea, R.D. (1987)
Socializing the knowledge transfer problem. International Journal of Educational Research, 11, 639–663.

Pearson, P.D. & Tierney, R.J. (1984)
On becoming a thoughtful reader: Learning to read like a writer. In A.C. Purves and O. Niles (eds.) Becoming readers in a complex society. Chicago: University of Chicago Press, 1984.

Popper, K.R. (1963)
Conjectures and refutations. The growth of scientific knowledge. London: Routledge and Kegan Paul.

Prawat, R.S. (1989)
Promoting access to knowledge, strategy, and disposition in students: A research synthesis. In: Review of Educational Research, 59, 1, 1–41.

Pressley, M. e.a. (1987)
In S.M. Cormier & J.D. Hagman (Ed.), Transfer of Learning. Contemporary Research and Applications. San Diego: Academic Press.

Raub, D.K. (1966)
The audio–lingual drill technique. An approach to teaching composition. Master's thesis. Nashville: Peabody College for Teachers.

Rayner, K. & Pollatsek, K. (1989)
The psychology of reading. Englewood Cliffs, NJ: Prentice Hall.

Resnick, L.B. (1976)
Task analysis in instructional design: some cases from mathematics. In: D. Klahr (ed.), Cognition and instruction. Hillsdale N.J.: Erlbaum Associates.

Reynolds, J.H. (1966)
Cognitive transfer in verbal learning. Journal of Educational Psychology, 57(6), 382–388.

Reynolds, J.H. (1968)
Cognitive transfer in verbal learning. Journal of Educational Psychology, 59(2), 133–138.

Rijlaarsdam, G.C.W. (1986)
Effecten van leerlingrespons op aspecten van schrijfvaardigheid. [Effects of peer response on aspects of writing ability] Dissertation University of Amsterdam. Amsterdam: SCO–rapport 88.

Rijlaarsdam, G.C.W. (1987)
Effects of peer evaluation on writing performance, writing processes, and psychological variables 38th Annual meeting of the Conference on College Composition and Communication, March 19–21, 1987. ERIC Document ED 284288.

Rijlaarsdam, G.C.W. (1989)
Basisvorming: Nederlands. [Basic Secondary Education: Dutch] Leiden: Martinus Nijhoff.

Rijlaarsdam, G.C.W. (1993)
Learning, whose learning? Autonomous learning in mother tongue education. 5th Int. Convention on Language and Education. March 22–26th 1993. Norwich: School of Education: University of East Anglia

Rijlaarsdam, G, Eiting, M. & Schoonen, R. (1989)
Wat werkt in het stelonderwijs? Een exploratie naar de effectiviteit van omgevingsvariabelen en vakinhoudelijke kenmerken van een stelvaardigheidscurriculum. [What works in composition instruction? An exploration into the effectiveness of environment variables and subject–dependent characteristics of a composition curriculum.] Tijdschrift voor Taalbeheersing, 2, 81–91.

Rijlaarsdam, G., Van den Bergh, H. & Zwarts, M.
Incidentele transfer bij produktieve taalopdrachten: een aanzet tot een base–line. [Incidental transfer with productive language assignments: towards a base–line.] Tijdschrift voor Onderwijsresearch, 17, 55–66.

Ripple, R.E. & D.J. Drinkwater (1982)
Transfer of Learning. In: H.E. Mitzel (Ed.), Encyclopedia of Educational Research. Heffers, Cambridge. Fifth edition, 1982.

Salomon, G. & T. Globerson (1987)
Skill may not be enough: The role of mindfulness in learning and transfer. International Journal of Educational Research, 11, 623–637.

Salomon, G. & Perkins, D.N. (1990)
Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. <u>Educational Psychologist</u>, <u>24</u>, 113–142.

Schriver, K.A. (1987)
<u>Teaching writers to anticipate readers' needs: Empirically based instruction.</u> Dissertation Carnegie–Mellon University. Pittsburgh (PA).

Schriver, K.A. (1989)
Evaluating text quality: the continuum from text–focused to reader–focused methods. <u>IEEE Transactions on professional communication</u>. <u>32</u>, 238–55.

Schriver, K.A. (1991)
Plain language through protocol–aided revision. In: E.R. Steinberg (ed.) <u>Plain Language: Principles and Practice.</u> Detroit, Michigan: Wayne State University Press, 148–172.

Schriver, K.A. (1992)
Teaching writers to anticipate readers' needs: A classroom–evaluated pedagogy. <u>Written Communication,</u>

Schultz von Thun, O. (1977)
<u>Psychologische Vorgange in de zwischenmenschliche Kommunikation</u>. Braunschweig: Westermann.

Schunk, D.H. (1991)
<u>Learning theories: An educational perspective</u>. New York: Merrill.

Schunk, D.H. en Hanson, A.R. (1985)
Peer models: Influence on children's self–efficacy and achievement. <u>Journal of Educaonal Psychology</u>, <u>77</u>, 313–322.

Schunk, D.H., Hanson, A.R. en Cox, P.D. (1987)
Peer–model attributes and children's achievement behaviors. <u>Journal of Educational Psychology</u>, <u>79</u>, 54–61.

Schunk, D.H. & Zimmerman (eds.) (1994)
<u>Self–regulation of learning and performance. Issues and educational applications.</u> Hillsdale, NJ: Lawrence Erlbaum.

Searle, J.R. (1979)
<u>Expression and meaning; studies in the theory of speech acts.</u> Cambridge: Cambridge University Press.

Selz, O. (1935)
Versuche zur Hebung des Intelligenzniveaus. <u>Zeitschrift für Psychologie</u>, <u>134</u>, 263–302. (Translated into English: <u>On Problem Solving</u>, Psychological Monographs, 58)

Shanahan, T. & Lomax, R.G. (1986)
An analysis and comparison of theoretical models of the reading–writing relationship. <u>Journal of Educational Psychology</u>, <u>78</u>, 116–123.

Shanahan, T. 1987)
The shared knowledge of reading and writing. <u>Reading Psychology</u>, <u>8</u>, 93–102.

Simons, P.R.J. (1990)
<u>Transfervermogen</u>. [Transfer capacity.] Inaugurate speech. Nijmegen: Quick Print.

Simons, P.R.J. & Beukhof, G. (1987)
<u>Regulation of Learning.</u> Den Haag: Instituut voor Onderzoek van het Onderwijs: Selecta reeks.

Simons, P.R.J. & Verschaffel, L. (1992)
Transfer: Onderzoek en onderwijs. [Transfer: Research and education.] Tijdschrift voor Onderwijsresearch, 17, 3–16.

Simonsen, S. (1988)
The generalizability of metacognition from the writing to the reading mode. Dissertation. Los Angeles: University of California Riverside.

Singley, M.K. & Anderson, J.R (1989)
The transfer of cognitive skill. Cognitive Science Series. Cambridge: Harvard University Press.

Smith, L. (1993)
Necessary knowledge: Piagetian perspectives on constructivism. Hove: Erlbaum.

Snow, R.E. (1990)
New approaches to cognitive and conative assessment in education. International Journal of Educational Research, 14(5), 455–473.

Sonnenschein, S. & G.J. Whitehurst (1983)
Training referential communication skills: The limits of success. Journal of Experimental Child Psychology, 35, 426–36.

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

Sternberg, R.J. (1985)
Beyond IQ: A triarchic theory of human intelligence. New York: Cambridge University Press.

Sternberg, R.J. & P.A.French (1993)
Mechanisms of transfer. In: Detterman & Sternberg (1993) The Case for the Prosecution: Transfer as an Epiphenomenon. In: Transfer on Trial: Intelligence, Cognition and Instruction Norwood N.J.: Ablex.

Stotsky, S. (1983)
Research on reading/writing relationships: A synthesis and suggested directions. Language Arts, 60, 627–42.

Straw, S.B. & Schreiner, R. (1982)
The effect of sentence manipulation on subsequent measures of reading and listening comprehension. Reading Reseach Quarterly, 17, 339–351.

Ten Brinke, J.S. (1976)
The complete mother–tongue curriculum. A tentative survey of all the relevant ways of teaching the mother–tongue in secondary education. Groningen: Wolters–Noordhoff.

Thelen, M.H., Fry, R.A., Feherenback, P.A. & Frautschi, N.M. (1979)
Therapeutic videotape and film modeling. Psychological Bulletin, 86, 701–726,

Thorndike, E.L. and R.S. Woodworth (1901)
The influence of improvement in one mental function upon the efficiency of other functions. Psychological Review, 8, 247–261.

Thorndike, E.L. (1906)
Principles of teaching. New York: A. G. Seiler

Van den Bergh, H. (1989)
De correlationele structuur van enkele schriftelijke taalvaardigheden. [The correlational structure of several written language skills.] Tijdschrift voor Taalbeheersing, 11, 38–59.

Van Dijk, T. & Kintsch, W. (1983)
Strategies of discourse comprehension. New York: Academic Press.

Van Eemeren, F.H. (1980)
Object en doelstelling op herhaling. [Object and goals revisited.] In A. Braet, Taalbeheersing als nieuwe retorica. Groningen: Wolters–Noordhoff.

Van Eemeren, F.H. & R. Grootendorst (1983)
Argumentatieleer 1. Het analyseren van een betoog. [A course in argumentation, 1: Analyzing argumentative text.] Groningen: Wolters–Noordhoff.

Van Eemeren, F.H. & R. Grootendorst (1992)
Argumentation, communication, and fallacies. Hillsdale, N.J.: Lawrence Erlbaum.

Van Gelderen (1987)
Relaties tussen spreek–, schrijf–, lees– en luisterprestaties. [Relations between speaking, writing, reading and listening performance.] Amsterdam: SCO.

Van der Horst, J.M. (1986)
Historische grammatica en taaltekens. [Historical grammar and languistic signs.] Dissertation. Amsterdam: University of Amsterdam.

Van Hout–Wolters, B.H.A.M. (1992)
Cognitieve strategieën als onderwijsdoel. [Cognitive strategies as educational goal.] Inaugurate speech. Groningen: Wolters Noordhoff.

Van der Pool, E. & Van Wijk, C. (1995)
Proces en strategie in een psycholinguïstisch model van schrijven en lezen. [Process and strategy in a psycholinguistic model of writing and reading.] Tijdschrift voor Onderwijsresearch, 20, 200–214.

Veenman, M. (1993)
Intellectual ability and metacognitive skill. Determinants of discovery learning in computerized learning environments. Dissertation University of Amsterdam. Amsterdam: University of Amsterdam.

Verbeek, J. (1993)
Een leerplan Nederlands. [A curriculum for Dutch.] Bouwstenen voor de Basisvorming. Groningen: Wolters–Noordhoff.

Vermunt, J.D.H.M. (1992)
Leerstijlen en sturen van leerprocessen in het hoger onderwijs. [Learning styles and regulation of learning processes in Higher Education.] Lisse: Swets & Zeitlinger.

Verschaffel, L., E. de Corte & H. Schrooten (1992)
Transfer van LOGO–kennis en programmeervaardigheden. (Transfer of LOGO knowledge and programming skills). Tijdschrift voor Onderwijsresearch, 17, 1.

Voss, J.F. (1987)
Learning and transfer in subject–matter learning: A problem–solving model. International Journal of Educational Research, 11, 607–622.

Wertheimer, M. (1945)
Productive Thinking. New York: Harper & Row.

Winer, B.J. (1971)
Statistical principles in experimental design. (2nd ed.) Kogakusha, Tokyo: McGraw–Hill.

Wong, B.Y.L. (1991)
Three conceptual perspectives on the connections between reading and writing processes. In L. Lupart &

A. McKeogh (Eds.) Toward the practice of theory–based instruction: Current cognitive theories and their educational promise. (p. 66–91). Hillsdale, N.J. Erlbaum.

Woodworth, R.S & H. Schlosberg (1955)
Experimental Psychology. Third, revised edition. London: Holt, Rinehard & Winston.

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