Discovery of Ideas and Ideas about Discovery : The influence of prior knowledge on scientific discovery learning in computer-based simulations.

Casper Hulshof

Discovery of Ideas and Ideas about Discovery

Casper Hulshof

2001

Ph.D. thesis
University of Twente

Twente University **Press**

Also available in print:
www.tup.utwente.nl/uk/catalogue/educational/discovery

Casper Hulshof

Discovery of Ideas and Ideas about Discovery

NWO

The research reported in this thesis was funded
by The Netherlands Organisation for Scientific Research
(grant number 575-22-001).

ico

The research was carried out in the context of
the Interuniversity Centre for Educational Research.

Twente University **Press**

Publisher: Twente University Press,
P.O. Box 217, 7500 AE Enschede, the Netherlands, www.tup.utwente.nl

Cover design: Jo Molenaar, [deel 4] ontwerpers, Enschede
Print: Grafisch Centrum Twente, Enschede

ISBN 9036516005

DISCOVERY OF IDEAS AND IDEAS ABOUT DISCOVERY

THE INFLUENCE OF PRIOR KNOWLEDGE ON SCIENTIFIC DISCOVERY LEARNING IN COMPUTER-BASED SIMULATIONS

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof.dr. F.A. van Vught,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op donderdag 7 juni 2001 te 16.45 uur

door

Casper Daniël Hulshof
geboren op 18 juni 1973
te Heiloo

Dit proefschrift is goedgekeurd door de promotoren

prof.dr. A.J.M. de Jong
prof.dr. J.M. Pieters

Contents

Voorwoord

Toen Mozart een jaar of zeventien oud was, componeerde hij zijn vijfentwintigste symfonie. De musicoloog Hans Keller heeft hierover geschreven: "…For the first time, darkly aware of his genius, Mozart dares. And for this very reason, he partly fails where on some previous occasions his undaring talent had all too easily, and not very significantly, succeeded. Only bad composers write chronically good music. An uninterrupted flow of immaculacy is the prerogative of mediocrity: if you don't say anything, there is not much risk of spoiling anything." Het afronden van een proefschrift geeft de auteur onvermijdelijk het gevoel gefaald te hebben: de theorie rammelt aan alle kanten, de experimenten hadden veel beter kunnen worden uitgevoerd, en de getrokken conclusies lijken kant noch wal te raken. Ook bij mij is dat een beetje het geval; maar als ik gefaald heb, dan hoop ik dat het is op Mozart's manier: dan valt het achteraf allemaal wel mee. Het is mijn stijl om de dingen vanuit een dergelijk negatief perspectief te bekijken, iets wat hier en daar in mijn teksten nog wel eens naar voren komt. Als de lezer het maar met een korreltje zout neemt dan komt het wel goed.

De totstandkoming van een proefschrift is geen eenmanstaak: er zit een heel team van knappe koppen achter. Een aantal van hen wil ik hier bedanken.

Allereerst wil ik mijn dagelijks begeleider en promotor Ton de Jong van harte danken voor zijn steun gedurende de afgelopen jaren. Ik weet dat hij het niet altijd even gemakkelijk heeft gehad met mij, en ik bewonder het grote mentale uithoudingsvermogen dat hij aan de dag heeft gelegd. Niet iedereen kan zoals hij ambitie en systematiek combineren met humor en een gezonde dosis relativeringsvermogen, iets wat ik als een voorbeeld voor mijzelf zal blijven zien. Ook mijn van eerste tot tweede 'gedegradeerde' promotor Jules Pieters wil ik bedanken, voor de steun tijdens de laatste fase in de afronding van dit proefschrift, en voor de gastvrijheid tijdens de bijna-traditionele sinterklaasviering.

Veel dank ben ik verschuldigd aan mijn collega's in het Inductief Lerenproject. Allereerst de OiO's: Frans Prins, Hedderik van Rijn, en Pascal Wilhelm. Het contact tijdens de 'OiO-meetings' en de informele omgang daaromheen was zeer vruchtbaar, en ik ben blij dat we veel voor elkaar hebben kunnen betekenen. Ook van de overige projectleden, Jos Beishuizen, Bert Bredeweg, Wouter van Joolingen, Maarten van Someren, Marcel Veenman, Jan Wielemaker, en Bob Wielinga heb ik tijdens mijn onderzoek veel steun ontvangen, zowel tijdens projectvergaderingen als tussendoor.

De collega's die gedurende de afgelopen jaren altijd nabij zijn geweest verdienen veel dank. Om te beginnen met mijn kamergenoten: Ilias heeft mij als 'cognitivistic pig' getolereerd, en mij op het spoor gezet van zeer interessante ontwikkelingen die zich in de marge van de psychologie voltrekken. Renate was daarna een heel wat rustiger kamergenoot, maar we hebben ons uitstekend vermaakt. Bregje heeft tenslotte de afronding van het werk kranig ondergaan, en onze gemeenschappelijke voorkeur voor snoepgoed, chips en muziek zorgde voor een aangename sfeer.

Wat betreft de andere collega's wil ik allen danken die aan de ProIST-bijeenkomsten hebben deelgenomen: het informele karakter ervan zorgde samen met de diepe discussies voor een altijd leerzame kijk op andermans (en zo nu en dan eigen) werk. In het bijzonder wil ik ook Koen bedanken voor de mentale stimulatie in de vorm van psychologie-artikelen en Colombiaanse koffie.

De experimenten die in dit proefschrift worden beschreven zouden niet kunnen zijn uitgevoerd zonder de medewerking van verschillende scholen. Ik dank daarom de leraren en leerlingen van scholengemeenschap De Grundel te Hengelo, het Marianum College te Groenlo, ROC Zadkine te Rotterdam, en het ROC afdeling techniek te Hengelo en Enschede. In het bijzonder wil ik noemen de heren Altena, Jetses, en Vroling voor hun steun voorafgaand en tijdens het eerste experiment, de heer van Woerkom voor hulp bij de organisatie van het tweede experiment, en de heren Buitink en Nijmijer voor ondersteuning bij het derde experiment.

De citaten die her en der het werk enigszins opleuken zijn voor een belangrijk deel afkomstig uit de onvolprezen 'quotefile' van Hans van Dok. Hiervoor mijn dank.

Meer dan vier jaar heb ik op de campus gewoond in huize .soixante-neuf (IP-bereik 130.89.226.209 tot en met 130.89.226.226), en ik dank alle bewoners en ex-bewoners die mij gedurende die periode hebben aangespoord om een echte baan te zoeken, hebben uitgelachen om mijn salaris, en toch indirect enorm gesteund hebben. Grolsch, koken en campusnet hebben voor leerzame ervaringen gezorgd.

Tenslotte dank ik mijn ouders voor hun steun en het vertrouwen, en Petra voor haar aanwezigheid en morele steun tijdens de laatste jaren.

Casper Hulshof Enschede, Mei 2001

1

Preliminaries: putting scientific discovery learning in context

> "I am very odd. That is to say, I am methodical, orderly, and logical—and I do not like distorting facts to support a theory—that, I find—*is* unusual!"
>
> *Agatha Christie, One, Two, Buckle my Shoe (1940)*

1.1 Introduction

In colloquial language, the term discovery denotes an event that contains an element of surprise: something that you did not expect to happen, actually does. In this sense, discovery possesses on the surface the characteristics of a ballistic movement: it entails a qualitative change of knowledge after which it can be difficult to return to the previous state. Discoveries are not unique events: they are made at all times at many levels. The highest level at which discoveries are made is in science. Scientific discoveries mark those points in research that involve a breakthrough in knowledge about a particular subject. Most of these breakthroughs are relatively minor, but every once in a while a major discovery is made. What is important to note, is that the different levels at which discoveries are made – from everyday life to the scientific endeavor – share similar characteristics: a novel, usually unexpected result, and a qualitative change in knowledge (which, in science, can result in a so-called 'paradigm shift'). However, scientific discovery also has an air of mysticism surrounding it. The scientific literature is awash with anecdotes about the mysterious circumstances that have accompanied many important scientific breakthroughs throughout history (e.g., Archimedes' bath, Newton's apple, Kekule's dream; also see Gigerenzer, 1991).[1] Modern science has tried to undo scientific discovery of its mystical wrapping, and to get an understanding the circumstances under which discoveries take place. This no-nonsense approach has been very fruitful. Knowledge has been gained about the logical steps that are necessary to perform 'good' science (e.g., De

[1] For an interesting comparison to the argument that is put forward here, consider the statement from James Joyce's *Ulysses* (1922): 'A man of genius makes no mistakes. His errors are volitional and are the portals of discovery.'

Groot's (1969) *empirical cycle*). Also, processes that foster the development of new knowledge (so-called *inductive* processes) have been studied and described in detail (e.g., Holland, Holyoak, Nisbett, & Thagard, 1986). Viewing discovery from a scientific perspective not only has had a profound impact on understanding of the scientific enterprise itself, but also on ideas about educational effectiveness. The assumption is that the characteristic of qualitative knowledge change – that is the hallmark of discovery – can be used as an educational tool in itself. The educational field has recently experienced many changes, which have fostered interest in the use of new instructional methods, such as learning by discovery. Of course, when discovery is employed in the classroom, a different approach to teaching and learning has to be taken than when, for example, a lecture is delivered. Fortunately, the technology of today allows for the realization of completely new styles of pedagogy. In education, there is now large and broad support for the creation and use in education of computer based learning environments (e.g., Hawkridge, 1990). The type of computer based learning environment that is typically associated with learning for discovery is *computer simulation* (De Jong & Van Joolingen, 1998). There is a large body of research on the effectiveness of using simulations and the particular learning processes that they elicit. The present thesis is concerned with the study of processes that occur during discovery learning with a computer simulation.[2] More specifically, the influence of different types of prior knowledge on these processes is examined. The research was carried out in the larger context of the 'Inductive Learning' project.

Context of this research project. The empirical studies that were carried out for this thesis were performed in the context of the project 'Inductive Learning', which was funded by the Dutch Organization for Scientific Research (NWO). The inductive learning project aimed at 'clarifying the major constraints of inductive knowledge development' (as stated in the research proposal). The constraints of inductive knowledge that were specifically studied in the project were prior knowledge, discovery skills, and meta-cognitive skills. These were separately explored in three research projects. The goal of the fourth project is to build a computational model of inductive learning, in which findings from the other three projects are incorporated. The studies

[2] In this thesis, the term 'discovery learning' is used consistently to refer to the educational activity of investigating, hypothesizing, and evaluating. Some researchers (e.g., Kuhn, Black, Keselman, & Kaplan, 2000) prefer to use the term *inquiry learning* instead, but there appears to exist no defining characteristic that separates between discovery and inquiry.

that are reported here are concerned with the prior knowledge constraint, although a measure for discovery skills was later included in the study (see Chapters 6 and 7).[3]

Structure of this chapter. In the following sections, discovery as it is used in education is described from both a practical and a theoretical point of view. The practical perspective (Section 1.2.1) concerns the influence of technology on classroom practice. The theoretical perspective (Section 1.2.2) concerns the evolution of the shift in perspective that has taken place in history from a learner-as-recipient to a learner-as-active-participant point of view. The discussion focuses (in Section 1.3) on the use of computer simulations to elicit scientific discovery learning processes. Learning processes that are involved in scientific discovery learning, and types of studies that have been carried out to study these processes are discussed. This leads us (in Section 1.4) to a general research question with respect to the influence of prior knowledge on these processes. Finally, in Section 1.5, a chapter-by-chapter overview of this dissertation is given.

1.2 Historical changes in education

Two external forces, practical and theoretical, have had a large influence on educational practice: in a practical sense the increasing use of technological tools, and in a theoretical sense the development of a variety of ideas that can be grouped under the heading of *constructivism*. In this section, these factors are separately discussed.

1.2.1 Technology in the classroom

"For the foreseeable future, computing will play an increasingly important role in human learning". With this prediction, an overview of the role of the

[3] The research project on discovery skills was carried out by Pascal Wilhelm and Jos Beishuizen, the project on meta-cognitive skills by Frans Prins and Marcel Veenman (both these projects at the university of Leiden), and the modeling project by Hedderik van Rijn, Maarten van Someren, and Bob Wielinga (university of Amsterdam). Other members of the research programme were Wouter van Joolingen, Bert Bredeweg, and Jan Wielemaker.

computer in education by Taylor (1980) begins. Twenty years later, the large-scale use of computers in everyday school life is finally being realized. This is the result of the ever-developing state of electronics, which, paradoxically, offers improving technology at a decreasing cost, and of the growth in popularity of hitherto unexplored areas in educational territory, such as (possibly) the internet. Computers have gained a strong foothold in the educational field, and their use has grown exponentially in recent years.[4] Amarel stated in 1983 that there are at least two ways in which computers can contribute to education: as learning aids in classroom instruction, and as tools in research on human cognition (Amarel, 1983). Most subsequent research has centered on the effectiveness of using computers in education versus using other educational methods. Early texts such as Taylor (1980) had shown unreserved optimism about the advantageous effects of using computers in education. However, whether the computer is used as a tutor, tool, or tutee (Taylor, 1980), the lack of beneficiary effects of computer use, as measured by controlled studies, has led to some controversy in the educational field (see, for example, the debate between Kozma, 1994, and Clark, 1994, about the influence of media on learning). The justification for computer use in education is not always given, which may lead to what Salomon and Perkins (1996) refer to as 'The Mount Everest rationale".[5] It has been difficult to show unequivocal positive educational effects of using computers in education. An early meta-analysis of 42 controlled studies (Bangert-Drowns, Kulik, & Kulik, 1985), showed only small positive effects of computer-based secondary school teaching. Some types of computer-based teaching were more effective than others, with computer-assisted instruction being more effective than computer-enriched instruction. The lack of substantial positive findings of the use of computers in education may be due to a difference of opinion about the role the computer should play in educational settings.

Jonassen and Reeves (1996) argue that the use of educational software can turn the computer into a *cognitive tool*. Cognitive tools refer to technologies that enhance the cognitive powers of human beings during thinking, problem solving, and learning. The use of cognitive tools has a large impact on

---

[4] The statement that computers are nowadays ubiquitous has become trivial to such an extent, that we would estimate that about one in five relevant articles begins with a version of it.

[5] This refers to the reply given by the famous mountain climber George Mallory (1866-1924) to the question why he would want to climb the largest mountain in the world: "Because it is there". (New York Times, March 18, 1923)

learning. As Säljo (1996) puts it, "[They] extend, or rather transform, the capacities for physical and intellectual action that were bestowed upon human beings by nature, as it were". It is argued by Jonassen and Reeves that skepticism about the effectiveness of different types of media as cognitive tools is mainly elicited by critics who use a narrow definition of the term media: a conveyor of information from teacher to student. They claim that media should not be used to convey information, but instead to facilitate 'knowledge construction' (also see Jonassen, Campbell, & Davidson, 1994). The argument they give for this claim is that during learning, more than anything else, students actively construct knowledge. The active construction of knowledge is defined by Duffy and Cunningham (1997) as one of the cornerstones of *constructivism*. Other critics of the 'narrow' view on the computer as a cognitive tool stress the way computer tools enable *augmentation* of learning, which is not expressed as a set of mental representational structures, but as a disposition to engage in appropriate scientific conversation (Pea, 1992; Tikhomirov, 1974). Papert (1987) has even questioned the whole scientific methodology that surrounds studies on the effectiveness of computers in education, going as far as to state:

> "It is a self-defeating parody of scientism to suppose that one could keep everything else, including the culture, constant while adding a serious computer presence to a learning environment . . . The treatment methodology leads to a danger that all experiments with computers and learning will be seen as failures: either they are trivial because very little happened, or they are 'unscientific' because something real *did* happen and too many factors changed at once. (p. 26)"

Salomon (1992a) partly corroborates Papert's suggestion. He compares the advent of the computer in the classroom with a Trojan horse. Introducing a computer affects the whole classroom environment. "It is the whole culture of a learning environment, with or without computers, that can affect learning in important ways. (Salomon, 1996, p. 113)". The conclusions that Salomon draws implicate that experimental effectiveness studies are more likely than not to show negative results. Salomon (1992b) states that, only when computers are used in such a way that an 'intellectual partnership' is created, there can be a positive learning effect.

It should become clear from this discussion, that computers have a number of potential benefits to offer in the sense that they can amplify the learning

experience.[6] However, to get a better appreciation of the possibilities that are offered by computers and the pedagogical vistas they open up, a change in theoretical perspective is needed. Technological innovations in education should be based on a theory, and not be driven by technological advances only (Koschmann, Kelson, Feltovich, & Barrows, 1996). The constructivistic approach, which was already briefly mentioned, is based on a new perspective on teaching and learning. Before moving on to a discussion of the type of computer tools that are used in education, we will describe this approach.

1.2.2 Constructivism

It was argued in the previous section that computers can change the learning experience by aiding cognition and by amplifying the learning process. Interest in tools that allow learners to be active and participate in their own learning has grown in recent years. This is the result of the development of the constructivistic perspective. In this section, the impact of constructivism on discovery learning is discussed.

Traditionally, teaching tools were mainly used to facilitate the direct transmission of all the relevant information to learners. The learners were invited (or even forced) to absorb and assimilate this information; no active participation, such as interaction with the learning material, was needed on their part. The teaching tools that are used in education determine at least partially how learners are being taught: passively absorbing information or actively engaged in the learning process. Bruner (1961) was one of the first to distinguish between a passive and active 'mode' of teaching, and referred to them as an expository versus a hypothetical teaching mode. The expository mode of teaching follows the traditional approach as described. Interestingly, Bruner argued that the hypothetical teaching mode stresses the point of discovery in learning: learners are led to seek for themselves for regularity and relatedness in the information that is offered to them, thus becoming 'constructionists'.

Ausubel (1963) endorsed Bruner's distinction between two modes of teaching. Furthermore, he argued that the main difference between these modes can be traced to differences in the learners' elicited behavior. Therefore, instead of speaking of *teaching* modes, Ausubel called this distinction a difference in *learning* modes. These learning modes were respectively referred to

[6] See Karasavvidis (1999) for a more detailed discussion on the concept of amplification.

as *reception learning* and *discovery learning*. Finally, Ausubel argued that there is another dimension to learning: the distinction between rote learning and meaningful learning. In rote learning knowledge is arbitrarily incorporated into the cognitive structure; this in contrast to meaningful learning, where new knowledge is embedded into specifically relevant structures (cf. Tikhomirov, 1974; Ausubel, Novak, & Hanesian, 1978; Novak, 1980). These structures reflect the background knowledge a person has about a subject, prior to a learning task. By making this difference between rote and meaningful learning, Ausubel recognized the influence of prior (background) knowledge on the way learning takes place in various learning tasks. Because of the influence of background knowledge, the meaningfulness of material to-be-learned differs from person to person. Although Ausubel admitted that learning by discovery is a good way to make material meaningful to students, he was skeptical about its usefulness as a general method for learning and instruction. He considered the benefits of discovery learning not to weigh up to the disadvantages: the procedure simply was too troublesome and time-consuming.

The constructivistic perspective changes the view on the nature of knowledge. Knowledge is not seen anymore as being something completely objective. Rather, what is learned can be different from person to person. As we have explained before, learning is not seen anymore as a process in which knowledge is transferred as a whole from teacher to learner; instead, knowledge is constructed in an active learning process, and thus becomes subjective by its nature (De Jong, 1999). This view on learning has been adopted by the scientific community, and it is one of the cornerstones of the *constructivistic perspective* (Duffy & Cunningham, 1996; Grabinger, 1996; Jonassen, 1991; De Jong & Van Joolingen, 1998).[7] There is now considerable consensus that learners cannot be compared with sponges: they are not just passively absorbing knowledge. On the contrary: learners play an active role in knowledge assimilation (De Jong, 1991). According to the constructivist view, learning is a cumulative process, in which learners construct a new mental representation by linking knowledge to the knowledge they already have (Ausubel, 1963; Biemans, 1997).

[7] This change in perspective can be seen as representing a paradigm shift, in this case from an 'objectivist' theory of learning to a 'constructivist' theory (Jonassen, 1991).

The importance that constructivism adheres to viewing learning as an active, instead of passive process, can be seen as analogous to Ausubel's idea of meaningful learning. Meaningful learning is seen as more closely resembling the way learning takes place in the real world than rote learning. When coupled with the large advances in technology that have been made in recent years, the theoretical perspective of active learning can finally become fruitful. Adopting a constructivistic point of view has an important implication for education. Because the process of knowledge acquisition is seen as an active, constructive process rather than a passive one, instruction has to aim at supporting and facilitating this constructive process, instead of relying on passive absorption of information. For instruction to assume this new role, it is necessary to create an atmosphere that evokes the processes that were discussed in this section. In a constructivistic discovery learning environment, students are invited to learn at their own pace, to actively seek relevant information and to attempt to make sense of it. They have to be challenged to become actively engaged in a subject. A computer simulation can offer such a constructivistic learning environment. In Chapter 2, the process of scientific discovery learning as it occurs while a learner works with a computer simulation is described in detail. The present discussion continues with a look on reasons for using computer simulations as tools for discovery learning, and on types of simulations that have been studied.

1.3 Discovery learning with computer simulations

As was argued by Jonassen and Reeves (1996), new types of media use appear to be eminently suited to perform the task of providing a constructivistic learning environment. The use of interactive multimedia allows learners to work at their own pace, and to assert (to some extent) control over the information that is delivered. Furthermore, recently positive relations between the use of computers in instruction and learning outcome have been found. Liao (1998), in a recent meta-analysis of 36 studies that compared effects of traditional instruction with (what Liao calls) 'hypermedia' instruction, found an overall positive effect for school classes that used computer-based interactive videodiscs, computer simulations, or interactive multimedia as instructional tools. Liao also made note of the fact that studies in which computer simulations (which are referred to as 'simulators' in that article) were employed showed better learning gains than studies that employed other multimedia systems for instructional delivery.

Interest in the use of computer simulation for education marks a logical step in an ongoing process of educational change. In the previous section both

the practical and theoretical advances that have led to this step were discussed. Simulations are tools that can fully exploit the various strengths that computers have, such as interactivity and the possibility to display objects in a multitude of ways (Riever & Parmley, 1995). Indeed, the characteristics of computers are such that they allow learning to proceed at a high level of interactivity (depending on the characteristics of the software that is used). The high level of interactivity that a simulation allows for can also motivate students to perform better than a different type of instruction might do (Shute & Gawlick-Grendell, 1994). Computer based learning environments enable students to move through the subject matter at their own pace, and meet their own interests and needs. When it is used as a learning environment, a computer simulation basically allows learners to change the values or properties of one or more input variables, and observe the effect of this on one or more output variables. The effect on the output variables is determined by the underlying rules of the simulation, which is also called the simulation's *model*. A learning environment uses some kind of interface that allows for the actual learner-simulation interaction. An interface can be strictly numerical, but can also contain pictures, charts, and other graphical arrangements that create a more real-world appearance to the simulation. In addition, the simulation's model can be either a simplified representation of reality or a faithful one. All this is decided during the design of the computer simulation. These two characteristics, the presentation to the learner and the faithfulness of the represented model, determine the *fidelity* of the simulation. Whether a low or high fidelity simulation is more effective in learning, is dependent on the simulated domain. Some domains are not faithfully represented on purpose, such as training simulations for dangerous situations and simulations in which time plays an important role. What is clear, however, is that the direct interaction that a computer simulation allows makes learners' behavior different from working with a more traditional type of interface (De Jong, De Hoog, & De Vries, 1993).

A large number of studies have focused on discovery learning in the context of computer simulations. In Chapter 3, the Optics computer simulation is described which is used as a tool for studying discovery learning in the empirical studies that have been carried out for this research project. In the present section, firstly the type of simulations that have been used to study scientific discovery learning is discussed. Secondly, learning processes that a computer simulation can elicit from learners are described. Thirdly, two

possible types of studies with computer simulations, process studies and effectiveness studies, are discussed.

1.3.1 Types of computer simulations

Many computer programs that have been used in the scientific literature can be listed under the header of computer simulation, although they have not always been referred to in that way. In this section, an overview of studies in which a computer tool was used to examine (discovery) learning processes is given.

The most general use of tools that allow for interactive learning is made by Grabinger (1996), who refers to these as REALs (Rich Environment for Active Learning). According to Grabinger, a REAL is very different from a computer simulation, because REALs are "much more comprehensive and holistic than individual computer applications (p. 668)". Grabinger uses the concept of *microworld* to refer to systems that play a role in a REAL through the delivery of information, practice, finding and presenting information, stimulation of high-level thought processes, promotion of collaboration, or exploration. Use of the term microworld can be traced back to Papert (1980), who used the word to refer to the self-contained world that the instructional tool LOGO represented. Papert (1987) later made a distinction between microworlds and simulations. According to Papert, simulations refer to environments that represent elements from the natural world, in contrast to a microworld in which the rules are purely mathematical constructs. The latter type of discovery learning environment is also called an 'intrinsic model' (Edwards, 1993; Dugdale & Kibbey, 1990). Van Joolingen and De Jong (1997) also make a distinction between computer-based simulation and microworlds. According to their definition, however, computer simulations are exploratory environments, and microworlds (such as LOGO) are modelling environments. Such a sharp distinction is not made in all studies. For example, a study by Miller, Lehman, and Koedinger (1999) uses a computer simulation (called 'Electric field hockey') to instruct students on the principles of electrical forces, but is referred to as an 'interactive microworld of electrical interaction' (p. 307).

In addition to microworlds, computer simulations can be compared to *games*. In a number of studies, the game-like nature of discovery learning

with computer simulations has been emphasized (e.g., Leutner, 1993).[8] Basically, the act of uncovering new facts can be likened to solving a puzzle. Also, a game-like environment can be said to allow learners to show creative behavior. Differences between simulations and games are the absence of a domain-related goal (only general goals are used), and the absence of the element of competition in simulations (Leemkuil, De Jong, & Ootes, 2000). The non-use of specific goals (e.g., 'winning the next round') also implies that behavior in a simulation is less bound to constraints.

A number of studies refer to a computer simulation as a *laboratory* (e.g., Friedler, Nachmias, & Linn, 1990; Schauble, Glaser, Raghavan, & Reiner, 1991; Veenman, Elshout, & Meijer, 1997). A laboratory can be seen as presenting the type of place where important scientific discoveries are made.

A characteristic that all computer simulations share is their high level of interactivity. Shute (1991) made use of an i*nteractive environment* in a comparison of two types of tutoring system (one that required rule application and another that required rule induction). Also, many computer simulation are designed to elicit discovery learning processes. This means that a computer simulation is sometimes referred to as a *discovery environment* (Shute & Glaser, 1990; Reimann, 1991; Swaak & De Jong, 1996).

Researchers appear to have their own preference in referring to simulations in a particular way. For this dissertation, the terms computer simulation and (scientific) discovery learning environment where chosen to refer to the simulation-based learning environments that were used in the empirical studies.

1.3.2 Discovery learning processes

In the present section, an overview of processes that take place during scientific discovery learning are described. A more detailed description of these processes is given in Chapter 2.

[8] Some studies even study the educational effect of computer games themselves; e.g., Betz (1995), in which the educational benefits of playing the game Sim City 2000 are examined.

In an analysis of studies of scientific discovery, Klahr and Simon (1999) have given five reasons for taking an interest in studying scientific discovery at all (also see Klahr, 2000, for a similar overview). The reasons that were given by Klahr et al., can be summarized as follows:

1) *Human value*. The enormous advances that have been made in all fields of science in the last centuries, invoke a curiosity in the process behind scientific discoveries.
2) *Mythology*. Many important discoveries in the history of science are surrounded by vague, sometimes mystic stories of how they came about. Thorough study of such 'magical' discoveries can uncover interesting information about the scientific process.
3) *The boundaries of cognition*. The products of scientific thinking lie close to the border of human cognitive ability. It is interesting to see whether the same cognitive operations are performed in normal thinking, as well as in thinking at the edge of performance.
4) *Developmental processes*. There is a curious parallel in children's thinking and scientific reasoning. However, many studies report developmental differences in scientific reasoning ability. The discussion that surrounds the question whether or not young children can be related to scientists has important ramifications for science education.
5) *Modeling*. More insight in scientific discovery processes can lead to better computer models of scientific reasoning.

The variety in reasons for stuying scientific discovery learning explains why there has been interest in formulating a theory for discovery learning processes. A framework for describing discovery learning is outlined in Chapter 2. Here, an overview is given of the tasks that are set by a discovery learning environment.

To get more insight into the theoretical background to discovery learning, it is important to understand the type of activities that are related to it, and that are part of it. One of the main goals of the scientific enterprise, is the uncovering of new truths (Simon, Langley, & Bradshaw, 1981).[9] Scientific actions sometimes take place in the face of uncertainty: the scientist does not (yet) know what will be found, if anything at all. What lies behind the discovery of facts, is a process of induction. Induction refers to all inferential processes that expand knowledge in the face of uncertainty (Holland,

[9] The question, 'what is truth?' is not discussed in this thesis. However, an effect that discovery learning can have, is to show how knowledge that is gathered by induction is of a probabilistic nature, and that scientific truths are often relative.

Holyoak, Nisbett, & Thagard, 1986). Learning the structure and rules of a particular domain in the absence of direct instruction can be referred to as inductive learning (Anderson, 1995). In successful discovery learning, learners are encouraged to induce the underlying relations in a domain. This process has certain advantages over other types of learning (such as classroom lectures).

As was discussed in the previous section, a computer simulation models a system (De Jong & Van Joolingen, 1998). The system that is represented in the simulation can be either fictitious or real. Making sense of the simulation (by means of discovery of the underlying rules) means making sense of the represented system. This type of learning, in which discovery plays a pivotal role, is more generally known as *learning by discovery*. An essential feature of computer simulated discovery learning is that the represented model is not disclosed directly to learners. Instead, its properties have to be inferred by experimentation with the simulation. To learners, the underlying model is, at least initially, not transparent, because of the interface that is built on top of it (Edwards, 1993; Swaak & De Jong, 1996). These characteristics make for a very individual learning experience, because not only can learners work at their own pace, they can also choose their own experimentation strategy and make their own inferences about the underlying model.

The complete learning process of performing experiments, evaluating evidence, inducing rules, and performing more experiments, is referred to as *scientific discovery learning* (e.g., Klahr & Dunbar, 1988; Reimann, 1991; De Jong & Van Joolingen, 1998). This term arises out of the idea that a computer simulated discovery learning environment imposes special conditions on the learning behavior of students. Students are encouraged to imagine that they have been put in the position of a scientist, who is interested in studying some new phenomenon worthy of experimentation. To discover patterns in their data, they have to design experiments, interpret experimental results, and apply these results to new experiments. De Jong and Njoo (1992) studied the behavior of students working with a computer simulation, using protocol-analysis. Analysis of the scientific discovery learning behavior of these students resulted in two important processes: *transformative processes*, and *regulative processes*. Transformative processes (like analysis of the domain information and hypothesis generation) yield information that learners can use in subsequent steps. Regulative processes do not yield knowledge, instead they are used to keep the discovery task itself under control; these

processes are, for example, planning and verifying. As De Jong and Njoo point out, what sets discovery learning apart from other types of learning are the regulative processes. Thus, this type of learning has a broader range of processes than other types of learning.

In an unguided environment, students are not tied to any criteria: they can design experiments at random, and do not have to work with specific ideas in mind with respect to the outcome of their experiments. When this procedure is used, learning certainly is not an easy task to accomplish, because an unguided discovery learning environment does neither give learners feedback concerning the logical soundness of their experiments, nor of the validity of their conclusions, nor of the quality of the conclusions they draw. In discovery learning, students have to decide for themselves whether they already have sufficient information to draw valid conclusions about the model underlying the simulation, or whether more information is needed. As such, learning without any external aid can be compared to juggling. The 'balls' (performed experiments, already found facts) have to be kept up in the 'air' (the learner's mind), which is a complex task because it carries a high cognitive load (Sweller, 1994). One redeeming factor may be though, that a clear-cut goal is absent in free discovery learning. It has been found, that the use of such goal-free problems can enhance learning (Sweller & Chandler, 1991; Sweller, 1994). The reason for this finding is explained by cognitive load theory, as forcing students not to use a means-ends analysis for problem-solving (a task that carries a high cognitive load). However, there is always the risk that discovery learning may lead to the construction of 'wrong' conceptions. These resemble the kind of intuitive conceptions students have about elementary physical phenomena, which are based upon children's experiences with the real world and their representation of these (Driver & Easley, 1978; Gilbert & Watts, 1983). In any case, without guidance discovery learning is not without risks. Because of this, usually only guided forms of discovery learning are used in school settings (Novak, 1980).[10]

[10] It has been argued that the teacher in a guided discovery situation restricts and frames the learning process, and that this means that learner's individual intuitions and ideas, where they don't fit in the lesson plan, are discouraged or neglected by the teacher. This goes against the principles of learning by experience, since learners' interpretations are governed by the teacher's interpretations in terms of learning goals (Edwards & Mercer, 1987).

1.4 General research question

As may have become clear from the discussion in the previous sections in this chapter, a vast amount of research has been carried out on the effectiveness of computer simulations and the learning processes that their use fosters. Before we come to the research questions that are the focus of this thesis, first a global distinction between two experimental methods will be discussed. The distinction this is made here is between studies that focus on the effectiveness of computer simulations, and studies that focus on the learning process itself.

Effectiveness studies focus on the product of discovery learning. The question these studies ask is: what is the educational gain of using a computer simulation for learning? In the beginning of this chapter, we have already discussed this point in some detail. One issue that deserves attention is the *measurement problem*. The discovery learning process is a special type of learning. Knowledge results from the interplay between experimenting, hypothesizing, and making inferences. Swaak and De Jong (1996, p. 346) suggest, that the knowledge that is constructed by learners during work with a discovery environment has special qualities. Swaak and De Jong developed a test, called the 'what-if' test, that attempts to provide a more valid measure of 'discovery knowledge' than standard tests. In this thesis, the what-if test is used in various incarnations. A more detailed description of this test is given in Chapter 4.

Process studies focus on what happens during discovery learning. In the previous section, the study by De Jong and Njoo (1992) used protocol-analysis to analyze scientific discovery learning processes. This led to the formulation of two types of processes: transformative and regulative. Other process studies compared the strategies successful and unsuccessful learners used while working with a computer simulation. For example, Shute and Glaser (1990) used the Smithtown learning environment (which deals with the subject of economics) to distinguish between good and poor discovery learning behavior. Shute and Glaser found that better learners generalize more elaborately, perform more complex and systematic experiments, generate better hypotheses, plan their experiments better, formulate more predictions, and take more focused notes. A study by Lavoie and Good (1988), showed that successful learners work more systematically and state better predictions of outcomes, reported similar results. Finally, Glaser, Schauble, Raghavan, and Zeits (1992) showed that discovery learning skills are not necessarily re-

stricted to one domain, but are part of general evidence-generating and interpreting skills.

The studies that were carried out as part of this thesis are all process studies. The main purpose was to study if and how discovery learning processes are regulated by the presence or absence of knowledge. In Chapter 2, a distinction is made between *generic* knowledge and *domain-specific prior knowledge*. Basically, domain-specific knowledge involves knowledge about variables and the relations between them. Generic knowledge involves knowledge about the type of mathematical relations that can exist between two or more variables. In all experiments, both types of knowledge are measured working with the computer simulation commences.

Our main research question was: how does prior knowledge influence learning in a scientific discovery learning domain. The main subject is learning with computer simulations; more specifically, the role of prior knowledge on learning processes that are elicited in the context of these simulations. All the empirical studies that we have carried out were attempts to shed some light on this relation.

1.5 Conclusion and overview of this dissertation

In this chapter, we have introduced various issues that pertain to the topic of this dissertation. We started with a discussion on the influence of technology on learning in the classroom. Running parallel with the wide-scale introduction of computers in classroom settings is the increased interest in theories (collected under the generic header of *constructivism*) that view the learner as an active participant in the learning process. The combination of theoretical insight in possibilities for active learning with technological advances that change the learning experience, have prompted the need for appropriate software tools to instruct students in a creative and innovative way. This thesis focuses on the use of computer simulations that foster scientific discovery learning. In Figure 1-1, a graphical overview of the structure is given. In the figure, a division into two core parts, A and B, is made. These parts together form the theoretical and empirical discussion of this dissertation. In part A (chapters 2, 3, and 4), the theoretical background and the description of the method for the empirical studies, is described. In part B (chapters 5, 6, and 7), the experimental studies are described.

Chapter 2 covers the theoretical background to the studies in this thesis. A theoretical framework, the Scientific Discovery as Dual-space Search theory

(Klahr & Dunbar, 1988) is described in detail. SDDS can be viewed as an appropriate theory to explain empirical data.

Chapter 3 gives a detailed description of the Optics computer simulation. This simulation was used in all the empirical studies that were carried out in this research project. The description focuses on the design choices that were made to create a suitable environment to study scientific discovery learning processes.

Chapter 4 is dedicated to describing the measurements that were used in the empirical studies. The chapter focuses on two types of measurement: firstly, the type of tests that were administered to measure prior knowledge, and secondly, the methods that were used to measure (and subsequently analyze) behavior in the Optics computer simulation.

Chapter 5 reports on the Optics 97 study. In this study, the Optics computer simulation was first put to use. Prior domain-specific knowledge about the subject of the simulation (geometrical optics), and general mathematical knowledge were measured prior to working with the computer simulation. Next to the Optics simulation, the Optics 97 study made use of another computer simulation, called Bubbles. The Bubbles simulation dealt with a fictitious subject about which no prior knowledge could exist among subjects. This enabled us to study discovery learning in complete absence of prior domain knowledge. Students at the level of pre-scientific education participated in the study.

Chapter 6 reports on the Optics 99 study. This study was a follow-up study to the Optics 97 study, and provided a more controlled environment. Students worked for four short, controlled, episodes in the Optics computer simulation. For this study, the knowledge measures were improved, and an additional prior knowledge test was introduced: the Peter task for discovery knowledge. This time, both students at the level of pre-scientific education and vocational education participated in the study.

Chapter 7 reports on the Optics 2000 study, which carried on from the results of the Optics 97 and Optics 99 studies. In this study an attempt was made at actively influencing the knowledge of students about geometrical optics. The study focused on the question whether students would be able to employ the knowledge they were given into their learning process. The participants in this study were students at the level of vocational education.

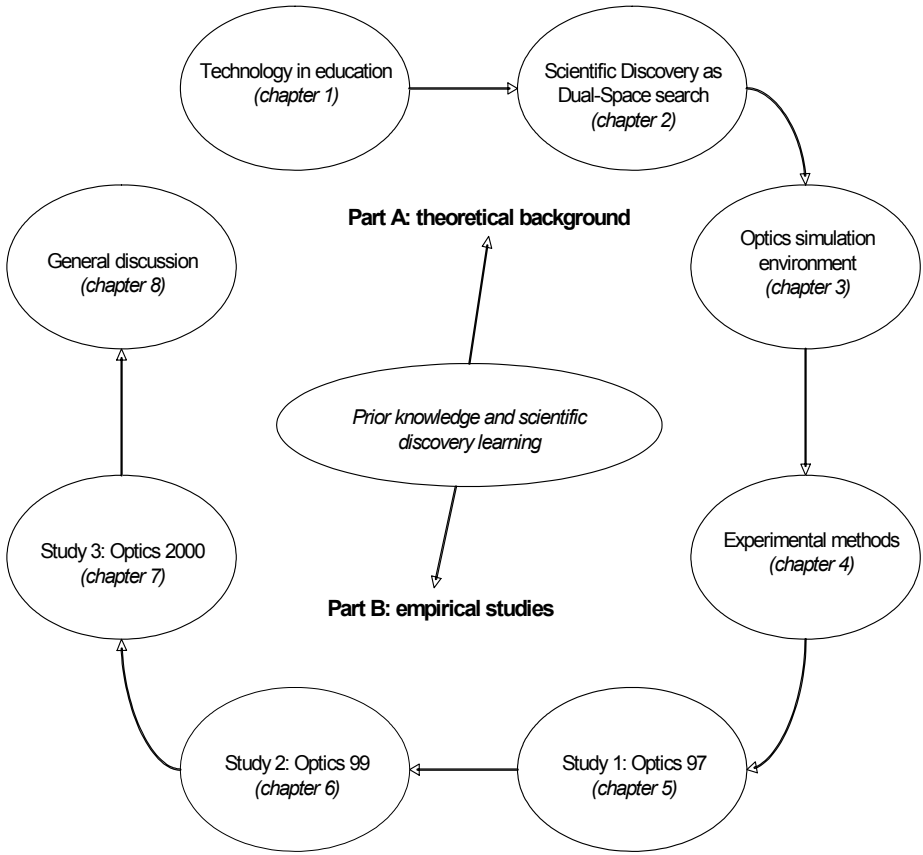


Figure 1-1. *Overview of the contents of this dissertation*

Chapter 8, finally, combines and discusses the results of the empirical studies. The ramifications are discussed: what are the consequences for educational practice and research? An attempt will be made to fit the results into a theoretical framework. The framework that will be used for this is the topic of the next chapter.

2

A theoretical background to scientific discovery learning

> Once I went professionally to an archeological expedition—and I learnt something there. In the course of an excavation, when something comes up out of the ground, everything is cleared away very carefully all around it. You take away the loose earth, and you scrape here and there with a knife until finally your object is there, all alone, ready to be drawn and photographed with no extraneous matter confusing it. That is what I have been seeking to do—clear away the extraneous matter so that we can see the truth—the naked shining truth.
>
> *Agatha Christie, Death on the Nile (1937)*

2.1 Introduction

We are told by none other than Sherlock Holmes himself, that it is a capital mistake to theorize before one has data.[1] Indeed, in many cases the origins of a theory can be traced back to an empirical observation, one that may have caused feelings of surprise. However, though Sherlock Holmes was a professional detective, he was only an amateur scientist. We will follow his advice only partly. In the current chapter, a theoretical background to the studies that form the empirical body of this dissertation is presented, and described in detail. In short, the theory of discovery learning that is presented here, can be traced historically to the pioneering work of Newell and Simon (1972) on human problem solving, and on Simon and Lea's (1974) description of the Generalized Rule Inducer. Problem solving is a basic cognitive activity (Anderson, 1990), and it has been shown that a general theory of problem solving can be powerful enough to unify seemingly unrelated cognitive phenomena (Simon, 1996). One of the most significant insights about discovery that cognitive science has brought to light, is the notion that the act of discovery may not be a 'magical' phenomenon (as has been suggested in some anthologies of scientific progress), but a phenomenon that, because of its relation to problem-solving processes, can be explained in a similar

[1] Sir Arthur Conan Doyle, *A scandal in Bohemia* (1892)

way as other cognitive phenomena. However, when it is stated that discovery is partly related to problem-solving behavior, this implies that another part of discovery is not related to problem-solving. It may be this part of discovery (the relative size of which is a matter of debate) which makes it a more or less complex phenomenon to explain. On the one hand, from an analysis of the behavior of the AM and BACON computer programs (both computer models that aimed at simulating discovery processes), Simon (1996) optimistically concluded that, "Discovery processes do not introduce new kinds of complexity into human cognition (p. 108)". On the other hand, a critical analysis of inductive processes by Holland, Holyoak, Nisbett, and Thagard (1986) led them to the conclusion that ''Such performances [such as BACON's], while very impressive, capture only part of the problem-solving activity crucial to scientific discovery. (p. 324)". In this chapter, the question: what exactly *is* scientific discovery? is addressed, in order to better construct a platform for a theory of scientific discovery learning.

As was stated in the introduction of Chapter 1, the processes that accompany (scientific) discovery have characteristics that are of potential use to education. The qualitative knowledge change that is the result of successful discovery, and the processes that lead up to discoveries, can be used as a means to an end in education. The type of discovery that can be used in an educational setting is usually referred to as scientific discovery learning. Learning situations that make use of scientific discovery learning can be modified to suit many different domains and learner personalities. As was stated in the previous chapter, the type of discovery learning that is central to this thesis, is learning in the context of a computer simulation. As was explained in the previous chapter, computer simulations are suitable for incorporating many principles that arise in the context of a constructivistic approach to learning. In general, what makes the use of computer simulations in the context of discovery learning particularly interesting, is the combination of a 'practical' side (the planning and the execution of experiments) and a 'theoretical' side (discovering rules from gathered data). The interactive experience that this type of learning affords has been found to result in a 'special' type of knowledge, which deviates from 'normal' conceptual knowledge in a number of ways (Swaak & De Jong, 1996; Swaak, 1998). The idea that knowledge that results from discovery learning has special qualities, has far-stretching consequences for the methods by which it should be measured. The knowledge test that is presented in Chapter 4, and versions of which were used in subsequent experiments, was intended to measure this type of knowledge.

Structure of this chapter. The current chapter sets out to map out the theoretical aspects that are behind scientific discovery and discovery learning. First, principles of discovery learning are discussed. Central to the idea that generic principles exist is the idea that discovery is, in a sense, a logical process. Simon's (1973) assertion that discovery has a logic can be seen as the inception of a cognitive theory of the complete discovery cycle. In Section 2.2.3, a theoretical framework that describes discovery learning as a special type of problem solving is presented. Central to the framework is the idea that discovery learning involves a search through two distinct spaces. The notion of these search spaces is explored in detail. In the final section of the present chapter, the relation of knowledge to discovery learning is examined. The role of prior knowledge on discovery learning processes as it has been found in other studies is discussed. In addition, the benefits of discovery learning compared to other types of learning are compared at the knowledge level. The chapter ends with a summary of the theoretical framework that is described.

2.2 A framework for scientific discovery learning

In this section, a theoretical framework for discovery learning processes is discussed. This framework describes the different steps that are involved in (successful) discovery. The remainder of this chapter is divided into two parts. The first part describes the logical steps that make up the process of discovery. An important part of this description is devoted to a discussion of the search spaces that are involved in the discovery process. The second part describes the role of prior domain-specific knowledge and generic knowledge in the discovery process. The purpose of the empirical studies that are described in this thesis is to test expectations with respect to the influence of prior domain-specific and generic knowledge on discovery learning processes. This section ends with a discussion on the fruits of discovery learning: what are the benefits of discovery learning, when compared to other types of learning? The answer to this question will influence the type of measurement that is needed to get a valid indication of students' knowledge level. The operationalization of domain-specific and generic knowledge is discussed in Chapter 4.

2.2.1 Logical steps in discovery

Is there a logic behind scientific discovery? In the aptly-titled book 'The logic of scientific discovery' (1958), Karl Popper answered this question with an emphatic "no". According to Popper, scientific procedures can only be studied in the *context of justification*; that is, there is a logic behind the way in which new ideas are (or should be) evaluated. What these new ideas themselves originate from, that is, studying the *context of discovery*, is not open to logical analysis. According to Popper, every discovery contains an irrational element, largely based on intuition. For this reason, the creation and processing of scientific data has been largely ignored in the past. Gigerenzer (1991) argues, that there is not so much a logic of scientific discovery, as a set of heuristics that guide both the discovery and justification stages. Gigerenzer also argues that Popper's gloomy view has led to a certain one-sidedness in the discussion about discovery:

> "Inductivist accounts of discovery, from Bacon to Reichenbach and the Vienna School, focus on the role of data but do not consider *how* the data are generated or processed. Nor do the numerous anecdotes about discoveries – Newton watching an apple fall in his mother's orchard while pondering the mystery of gravitation; Galton taking shelter from a rainstorm during a country outing when discovering correlation and regression toward mediocrity; and the stories about Fechner, Kekulé, Poincaré, and others, which link discovery to beds, bicycles, and bathrooms. What unites these anecdotes is the focus on the vivid but prosaic circumstances; they report the setting in which a discovery occurs, rather than analyzing the process of discovery. (p. 254-255)".

Popper's view on discovery was contradicted by Simon (1973), in an analysis of the problem of induction. According to Simon, the creation of a normative theory of scientific discovery is a possible venture, if it is stated as a set of criteria for evaluating law-discovery processes. Law-discovery processes are defined by Simon as those processes that recode sets of empirical data. Simon concludes, "The greater efficacy of one process compared with another in discovering laws need not be attributed to chance, irrationality, or creative intuition. Rather, it is a matter of which process is the more capable of detecting the pattern information contained in the data, and using this information to recode the data in more parsimonious form. (1973, p. 479)". By redefining the creation of hypotheses as the recoding of data, Simon managed to remove at least part of the suggestion of 'magic' from the discovery process. However, Simon (also in later studies, see for example Qin & Simon, 1990) chooses to neglect the question of where relevant data stems

from: its presence is taken for granted[2]. The important thing to note is that throughout recent history, discovery has been recognized as a complex process. Only recently has there been progress in dispelling some of the myths surrounding the creation of new ideas. This is due in part to theoretical insights, which have culminated in a broad framework for discovery. Also, attempts to create computational models of scientific discovery processes have needed a rigorous and complete description of (optimal) discovery processes, which has boosted research on these processes (Shrager & Langley, 1990; Langley, 2000).

The discussion about the logic of discovery has been permeated by the separation between discovery and justification. For a theory of discovery to be acceptable, it will need to be able to explain both sides of the coin. As will be shown, the scientific discovery as dual-space search framework of Klahr and Dunbar (1988) still carries the weight of the past in its distinction between experiment space and hypothesis space. However, because the dual-space search framework considers both the discovery and justification parts of scientific discovery, it succeeds in sketching a more or less complete picture of scientific discovery. In the following section, principles of discovery learning are discussed. This formal look at the activities that make up discovery can be used as an introduction to the dual-space framework that is described in detail next.

2.2.2 Principles of discovery learning

The process of discovery learning can be described as involving a number of different activities, the most important of which are on the one hand data collection (by an active search for information, using experiments), and on the other hand an inference of hypotheses from this data. Creating hypotheses can be seen as an 'internal trial and error' activity in which a learner searches for, and selects one or more relevant hypotheses (Gagné, 1966). This means that both experimenting and hypothesizing can be considered to be active search processes. As has been argued before in this chapter, these search processes (one of doing experiments, another of searching for the

[2] However, in other studies (e.g., Simon & Kulkarni, 1988), Simon and colleagues go beyond this data-centered view of discovery, and examine the type of heuristics used by scientists to guide experimental research.

right hypothesis) can be related to the theory of problem solving that was first advanced by Newell and Simon (1972). This theory describes problem solving as moving from one problem state to another through a problem space. In the theory, the direction in which movement through the problem states takes place is the result of an active search process. A distinction is made between three types of problem state: an initial state, one or more intermediate states, and a goal state. The theory of Newell and Simon deals with problems for which a specific goal can be stated, and which can be solved by employing a certain procedure (e.g., by application of an algorithm, or by using a certain heuristic). For more complex and more loosely structured problems (that is, problems that do not have a clearly specified goal), this view on problem solving quickly becomes too limited to warrant a description in terms of a search through a problem space. Therefore, in their research on discovery learning, Simon and his colleagues have tried to extend the theory, by describing the way people move from one problem state to another (e.g., Langley, Simon, Bradshaw, & Zytkow, 1987; Qin & Simon, 1990). They consider the first phase of making sense of a problem to be a process of induction (also see Holland, Holyoak, Nisbett, & Thagard, 1986). Research has centered on the heuristics people use when they induce rules from the information at their disposal and when they perform experiments. The usefulness of this kind of research lies in the domain-independent character these heuristics have (Kulkarni & Simon, 1988). In the previous section, we have criticized this approach to discovery, because it ventures only on one end of the spectrum: that of rule induction. Discovery learning encompasses both discovery and justification, and accounts such as De Groot's (1969) empirical cycle grant other discovery learning processes (e.g., rule evaluation) a lot of importance[3].

The following (formal) description of the complete process of discovery learning can be given. A student who is engaged in discovery learning, is basically exploring a domain. The boundaries of the domain are determined by the discovery learning environment, in the sense that the learning environment usually contains a model of the domain. The model describes relations between different variables. Although the discovery learning environment discloses a few or even all variables to the student, the relations are kept hidden. It is the learner's task to discover them, and he can do so by examining individual combinations of variables and 'guessing' their relation

[3] Other characteristics that pertain to discovery processes are, for example, a) defining a scientific problem, and b) making predictions on the basis of results (Friedler, Nachmias, & Linn, 1990).

from the outcomes that are given. Further experiments can be carried out to confirm or disconfirm a guess (or: hypothesis). In the former case, the now-proven hypothesis becomes a 'discovered' rule. In the latter case, the hypothesis turned out wrong, and hence should have been discarded[4]. In an ideal case, two sets of hypotheses are gradually formed: a set of hypotheses known to be true (that is, hypotheses for which evidence has been found that proves them likely to be true, and that have not yet been discarded), and a set of hypotheses known to be false. A third set of hypotheses can be distinguished: the set of hypotheses not (yet) known to be true or false. The notion of different 'sets' of hypotheses is captured in the 'discovery learning as dual-space search'-framework that has been put forward by Klahr and Dunbar (1988). An explanation of this framework, and a discussion on the role of different types of prior knowledge in the context of the dual-space search theory will form the main body of the next sections.

2.2.3 Scientific discovery learning as a dual-space search process

> Traditional scientific method has always been at the very *best*, 20-20 hindsight. It's good for seeing where you've been. It's good for testing the truth of what you think you know, but it can't tell you where you *ought* to go, unless where you ought to go is a continuation of where you where going in the past.
>
> *Robert Pirsig, Zen and the Art of Motorcycle Maintenance (1974)*

During computer discovery learning, students carry out experiments in a virtual environment, in order to get new information, and to modify their existing beliefs (Kulkarni & Simon, 1990). When both the process of experimentation and the process of deriving a hypothesis are described as search processes, it is possible to extend the general theory of problem solving to

---

[4] cf. Chinn & Brewer (1993), who found that learners are unwilling to perform an experiment, when they expect it will lead to a negative result, that is, a refutation of a hypothesis. This phenomenon is also called 'fear of rejection' (Van Joolingen, 1993; also see Brehmer & Dörner, 1993).

incorporate both. Simon and Lea (1974) proposed such an extension, by dividing the (standard) problem space into a 'rule space' and an 'instance space'. Simon and Lea's Generalized Rule Inducer (GRI) provided a mechanism, in which experimentation and discovery played a role, although only in relatively simple tasks. Klahr and Dunbar (1988) have provided an extension to the GRI (which was meant to serve as a description of a concept learning task) to accommodate more complex situations that resemble the kind of situation scientists find themselves in. In the framework by Klahr and Dunbar, rule space is replaced by 'hypothesis space' and instance space is replaced by 'experiment space'. Based on earlier research with the programmable BigTrak vehicle (e.g., Shrager & Klahr, 1986), in which free interaction with a complex object was studied, Klahr and Dunbar proceeded to analyze the behavior of learners whose task was to find out one rule. The particular rule to discover in the task had turned out to be a difficult one in the earlier experiments. The model that Klahr and Dunbar have proposed to accommodate their findings, describes the process by which subjects search for (and eventually discover) this rule. They have called this model **SDDS** (Scientific Discovery as Dual-space Search). This model describes discovery learning as a problem solving-like search through the two problem spaces: *hypothesis space* and *experiment space*. Basically, the model states that discovery learning consists of three components: search hypothesis space, search (or: test) experiment space, and evaluate evidence. Each of these processes has its own goal. In hypothesis space search, the goal is to form a fully specified hypothesis (Van Joolingen & De Jong, 1997). This hypothesis can be tested by performing an experiment. The resulting evidence is evaluated, the goal of which is accepting or rejecting a stated hypothesis.

This concise description of the SDDS model provides a brief outline of the way Klahr and Dunbar (1988) describe the process of rule discovery. One of the strengths of this framework is that it contains an elegant and useful description of the discovery learning process as it occurs in a number of domains. Also, this parsimonious framework can serve as a general framework for studies on discovery learning processes. Klahr and Dunbar have used and tested the model in a number of studies, most of which dealt with the BigTrak[5] domain (Dunbar & Klahr, 1989; Klahr & Dunbar, 1988; Klahr, Dun-

[5] BigTrak is a programmable robot toy. The four-wheeled toy is operated by programming its movements, which are then carried out in sequence. The research studies by Klahr and Dunbar focus on the discovery of the effect of one special programming button: the RPT key. This key has the effect of repeating the last *n* operations, where n is the number that is keyed in after the RPT key.

bar, & Fay, 1990; Klahr, Fay, & Dunbar, 1993). The theory has also been applied to the domain of genetics (Dunbar, 1993). In these studies, developmental differences in the depth and complexity of the search processes have been examined, and also the effect of supplying students with relevant information, prior to discovery learning. It turns out that for the domains that have been used in these studies, the model can be used effectively. Although these domains are considered to be complex by Klahr and Dunbar, the number of variables and relations of which the properties have to be discovered is actually very small. For example, in the Klahr and Dunbar (1988) study that made use of the BigTrak vehicle, there was only one rule to be discovered. Discovery learning environments that have been used in other contexts, such as Shute and Glaser's (1990; also see Shute, Glaser, & Raghavan, 1989) Smithtown, and Schauble, Glaser, Raghavan, and Reiner's (1991) Voltaville employ more intricate domains, in which a large number of variables interact in a variety of ways. The simulation environments that were used in the present studies are of this type, too: a set of underlying formulas describes relations between a number of variables. The formulas themselves are quite complex, involving division and multiplication. It is assumed, that in these intricate domains search processes follow a somewhat more complex path than described by the SDDS theory. For example, hypotheses will be more varied, and also differ in their level of precision (Plötzner & Spada, 1992). In addition, the SDDS model provides only a global account of the way search processes take place: neither the structure of hypothesis space nor the structure of experiment space are described. Finally, there is no thorough description of the way discovery learning processes are influenced by prior knowledge. There is only the understanding that prior knowledge in general influences the prior configuration of hypothesis space. In the next section, the role of prior knowledge on scientific discovery learning is discussed in more detail. First, however, two useful extensions to the standard SDDS framework are discussed. These are Schunn and Klahr's (1995; 1996) 4-space framework of discovery learning, and Van Joolingen and De Jong's (1997) extended SDDS model. A discussion of these extensions to the framework, will prove useful when the role of prior knowledge on discovery learning is discussed in Section 2.3.

2.2.4 A four-space model of discovery learning

> Four legs good, two legs bad.
>
> *George Orwell, Animal Farm (1946)*

In a number of recent publications, Schunn and Klahr (1995, 1996) have discussed what they call the 'magic number' of search spaces that a framework for scientific discovery learning needs to incorporate. They argue, that the 2-space framework can sometimes be too limited in explaining various processes that occur in complex discovery tasks. In the extension to the 2-space framework that they propose, the experiment space and hypothesis space are complemented by two new search spaces: a *data representation space* is added to hypothesis space, and a *paradigm space* is added to experiment space[6]. In the data representation space, experimental data is abstracted and processed through various possible representations. The paradigm space represents (global) classes of experiments. A search through paradigm space involves selecting a class of experiments, thus identifying a set of independent and dependent factors. Schunn and Klahr argue that, as in the case of two search spaces, processing within each space is dependent on the state of the search in the other space(s). The information flow between the search spaces is depicted in Figure 2-1. As can be seen from the figure, the spaces do not all influence each other.

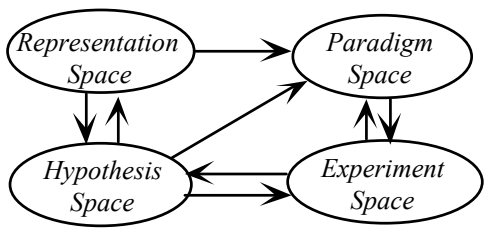


Figure 2-1. *Proposed information flow between search spaces in the four-space search model. From Schunn and Klahr, 1995. Reprinted with permission.*

[6] In an analysis on the discovery of the origin of ulcers, Thagard (1998) proposes three search spaces: experiment space, hypothesis space, and a space of instrumentation.

For example, the search through representation space is influenced by the hypotheses that a learner has arrived at. In turn, a change in the representation of experimental data may influence the hypotheses that are induced, and also the class of experiments that are carried out (by influencing the paradigm space). The model that Schunn and Klahr present, is based on research in the MilkTruck domain (for an extended review of the research with Milktruck, see Klahr, 2000, chapter 7), in which the task was to discover a (complex) rule by experimentation.[7] Evidence for activity in the data representation space was found in a change in the description given by subjects for the experimental outcomes. At first, programmed data was represented to subjects as a collection of single steps. After experimenting with different programs, this representation would change. Instead of single steps, programs were now seen as collection of segments of data. Evidence for the occurrence of activity in the paradigm space was found in a change of the type of experiments subjects planned to perform. The type of experiments that were carried out gradually grew more complex. According to Schunn and Klahr, this indicates that the database of experiment types, that learners made use of in the MilkTruck experiments, changed over time as a result of activity in the paradigm space.

Schunn and Klahr (1996) argue, that certain criteria need to be met to allow a search space to be added to a model of scientific discovery. Of these criteria, two are relevant to the present discussion, namely logical and empirical criteria. Logical criteria imply that search spaces should be *distinct* from each other. This means that search spaces should involve distinct entities and distinct goals. Empirical criteria imply that the use of different search spaces should be *observable*. This means that it should be possible to make a reliable classification of discovery learning behavior into search behavior in two or more spaces. As a sidenote, Schunn and Klahr state that the knowledge that a learner possesses can partly determine whether a particular problem space exists or not in a particular case.

[7] Milktruck can be seen as a more complex variation of the BigTrak domain. It dealt with a milktruck with a programmable route. The subjects' task was to discovery the function of a mystery command that could be given to the truck. This command was a function with three arguments, one continuous, and two dichotomous.

2.2.5 The extended SDDS model

Van Joolingen and De Jong (1997) have tried to extent the dual-space search model by examining the structure of the experiment space and hypothesis space. Here, the detailed look at hypothesis space that they provide is focused on, because the structure of hypothesis space as it exists prior to a (discovery) learning task is equal to (the amount of) domain-specific knowledge.

A domain can be defined as a collected body of knowledge. If this knowledge is described in the form of a collection (or set) of statements, two types of sets can be distinguished: the set of all true statements about the domain and the set of all untrue statements. The number of possible statements is virtually unlimited. Together, the set of true and the set of untrue statements form a general knowledge space that Van Joolingen and De Jong call the *universal hypothesis space*. A domain is always bound by a set of rules and a set of variables. This implies that it is possible to describe a domain with only a limited set of generic statements about relations between variables. Van Joolingen and De Jong (1997) call the (limited) set of true statements about a domain the *target conceptual model*.

Exploring a domain means that a learner tries to find out the boundaries of the target conceptual model. Like the knowledge that is encapsulated in a domain, the knowledge a learner has about the variables and relations in the domain can be considered as a knowledge space. All statements that a learner can assert about relations between two or more variables in a domain (independent of their truth value) are called the *learner hypothesis space*. The *effective learner space* is a subset of the learner hypothesis space. The effective learner space encompasses only those statements in the learner hypothesis space that a learner finds worthwhile to test. Testing for a certain relation between variables will only be considered by a learner when he or she has knowledge of the existence of that type of relation. For this, sufficient generic knowledge of mathematical relations is needed. In the next section, this idea is worked out in detail. A main goal in discovery learning about a domain is to bring the learner hypothesis space and the target conceptual model closer together. An ideal end state would be when the learner hypothesis space and the target conceptual model would overlap. Prior to a learning task, when a domain is relatively new to a student, this will of course not be the case.

Both the learner hypothesis space and the effective learner space may change in the light of new findings about a domain. The difference between these two spaces can be determined by the plausibility of hypotheses (Klahr, Fay, & Dunbar, 1993). It is always the case that some hypotheses are more

plausible to a learner than others. For example, gravity does not sometimes go in another direction, so the plausibility of such a statement is low for a system that simulates falling objects. Effective learner space can also be constrained by the frame in which a learner works. For example, Klahr and Dunbar (1988) found that subjects always worked with the idea in mind that the RPT key worked in a certain way (controller or selector). Changing to a different representation proved to be very difficult. Further research showed that it is easier to switch from an implausible frame to a plausible one than vice versa (Klahr et al., 1993), and that it is a useful heuristic in experiment generation to first consider different frames (this is the 'breadth first' type of search, in contrast to the 'depth first' type).

2.3 Prior knowledge and scientific discovery learning

The previous section can be read as a chronological account of the development of a theory of discovery learning: from the controversy surrounding the distinction between discovery and justification, to the description of the extended search space model. In this section, we will discuss the role played by prior knowledge in the scientific discovery learning process.

A number of factors determine individual behavior in a discovery learning environment and subsequent performance on a knowledge task. Learners do not enter a learning situation with an empty mind; they come to a learning situation with their ideas about a specific subject, and differ in learning ability. Of these determinants, prior knowledge is generally considered one of the most important ones. Schauble (1996) argues that "Most studies from the experimentation strategies approach underemphasize the role of domain-specific knowledge". This statement succinctly sums up a prevailing notion in recent research on experimentation and scientific reasoning: the lack of understanding about the knowledge that learners bring to the learning task. Further on, Schauble adds to this conclusion, "To decide which of several potential causes are plausible, people bring to bear both specific knowledge about the target domain and general knowledge based on experience about the mechanisms that usually link causes with effects…prior knowledge guides observations, as surely as new observations lead to changes in knowledge (p. 103)". In a number of studies, the relationship between prior (domain-specific) knowledge and learning effectivity in computer simula-

tions was studied. All the studies that are cited here looked at learning results, but also at processes that occurred during discovery learning.

Njoo and De Jong (1993) found that subjects who scored high on a domain-specific knowledge test administered prior to working with a simulation learning environment also scored high on a posttest. It was harder to find a relationship between the knowledge learners had prior to working with a learning environment and their subsequent interaction behavior. Njoo and De Jong did not find such a relation, but Lavoie and Good (1988) reported that students with high prior knowledge showed better discovery behavior. Glaser, Schauble, Raghavan, and Zeits (1992) found that higher prior knowledge led learners to consider more alternative hypotheses. Because the type of hypotheses that learners had in mind varied (e.g., both correct and incorrect hypotheses were used by subjects), Glaser et al. argue that prior knowledge can be of help to learners as well as mislead them. Not only the amount of prior knowledge, but also its specific quality that has to be taken into account. It is hypothesized by Chinn and Brewer (1993) that the deeper rooted prior knowledge is, the easier students will refuse to change their conceptions in the light of anomalous data. Instead, ad-hoc arguments ('I probably did something wrong') will help them do away with the anomalous data. Based on results from the BigTrak study, Klahr and Dunbar (1988) argue that prior knowledge can make the discovery of a correct hypothesis more difficult, because subjects' prior knowledge can suggest 'misleading and conflicting analogies' (p. 11). Students may get confused by a mismatch between their knowledge and the data they find (Chinn & Brewer, 1993). Klahr, Fay, and Dunbar (1993) showed that prior domain-specific knowledge rendered some hypotheses more plausible than others, which sometimes hampered successful discovery. Background knowledge can lead students to find some hypotheses more plausible than others, which influences the type of experiments they carry out to gather data.

On the whole, the studies that are mentioned here have shown that in certain cases prior knowledge can hinder student's performance in a learning environment. However, the presence of prior domain-specific knowledge can also have a positive influence on the learning strategy that students use while interacting with a learning environment. Schauble, Glaser, Raghavan, and Reiner (1991) made a division of prior knowledge into four categories, from a low, superficial understanding of a domain (electrical circuits) to deeper and more integrated knowledge. Schauble et al. found that a higher level of prior knowledge coincided with higher learning gains after working with a learning environment. Prior knowledge also affected learning behavior, with higher level knowledge subjects stating more predictions and gen-

erating more explanations, searching more broadly through all possible experiments, and working more systematically through all assignments.

What the research we have mentioned in this section indicates is that prior knowledge plays an important, but to a large extent unspecified role in discovery learning. Two reasons can be given to explain why it is not clear exactly how prior knowledge influences interaction behavior. First, in most studies prior knowledge has been used as a term to denote something like domain-related knowledge. Exactly what type of prior knowledge was measured, and how this knowledge stood in relation to the learning task was not specified. Second, it has not been made clear in previous studies what theoretical relation was expected between the prior knowledge that was measured and interaction behavior.

The extended version of the SDDS model can be used to look in more detail at the influence of prior knowledge on discovery learning processes. In the framework, hypothesis space is divided in four sections: the universal hypothesis space, the target conceptual model, the learner hypothesis space, and the effective learner space. The contents of the universal hypothesis space and the target conceptual model are dependent on the domain, and have a static content. The configuration of the learner hypothesis space and the effective learner space is influenced by prior domain-specific and generic knowledge. These two spaces are dynamic, that is, their contents can change over time. As was discussed in the previous section, a 'goal' of discovery learning is to change the leaner hypothesis space so that it comes to resemble the target conceptual model. This means that a successful search through experiment space and hypothesis space should lead to the conversion of the learner hypothesis space to a space of true and false statements. Figure 2-2 illustrates the process that is described. During discovery learning, the learner hypothesis space should become more like the target conceptual model. The initial makeup of the learner hypothesis space is determined by the level of domain-specific knowledge and generic knowledge (which is discussed next). Acquiring the target conceptual model implies that knowledge of true and false statements about a domain is gained.

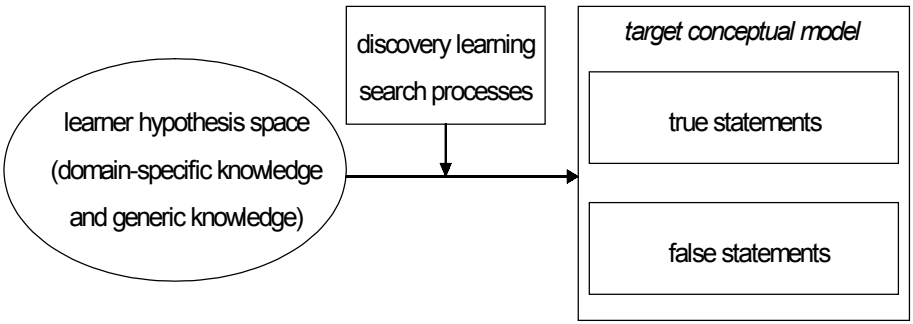


Figure 2-2. *Description of the search process in which the learner hypothesis space comes to resemble the target conceptual model*

Prior to entering the learning task, a learner can already have some knowledge about what is true and what is not true within a domain. This means that the learner hypothesis space will already be partly filled with statements about variables and relations. Thus, it can be said that the prior knowledge, or more specifically the prior *domain-specific* knowledge that a student has available will determine what is already available in the student's learner hypothesis space. Prior domain-specific knowledge provides learners with information about the relevant variables and relations in a domain, some of which might otherwise pass unnoticed. In addition, prior knowledge can help learners decide upon the plausibility of certain hypotheses, and can inform them about the right framework. Another type of prior knowledge determines learner's knowledge of the type relations that are possible. For example, a learner may not know what an asymptotic relation entails, and may therefore experience trouble in interpreting findings that seem to point at just such a relation. In our research, we have found it useful to make a distinction between *domain-specific knowledge* (knowledge of specific variables and relations in a domain) and *generic knowledge* (general knowledge about the type of mathematical relations that exist).[8] We define domain specific prior knowledge as knowledge about the existence and characteristics of variables in the domain. In the domain of optics this

[8] The term 'generic' knowledge was chosen to distinguish it from (domain-)specific. In Chapter 8 (Discussion) the concept of generic knowledge as a middle road between domain-specific and general knowledge about mathematics is discussed.

would, for example, be knowledge about the existence of focal points, and knowledge about the variables that a focal point relates to. Generic knowledge can be explained as the knowledge that is needed to recognize and work with various types of mathematical relations (that is, relations between two or more variables that can be numerically or graphically depicted). This knowledge is has a generic nature because it can be applied in all domains that consist of relations between a number of variables. Examples of mathematical relations between two variables are asymptotic and periodic relations. The idea of generic knowledge of mathematical relations is related to an argument by Plötzner and Spada (1992): learners have to utilize in a learning environment, amongst others, "mathematical knowledge about functional relationships and various arithmetical procedures (p. 107)". Generic knowledge is distinct from domain-specific knowledge in the sense that the latter type does not refer to the ability to recognize a relation as a certain type of mathematical expression, but to knowledge *of* a specific relation or set of relations. De Jong and Ferguson-Hessler (1996), in an overview of knowledge qualities, also made a distinction between domain-specific and generic knowledge. They showed that this qualitative distinction applies for different types of knowledge (situational, conceptual, procedural, and strategic). The distinction that is made here between generic and domain-specific knowledge is not the same as the distinction that is made by Alexander and Judy (1988) between domain-specific and strategic knowledge. Our concept of domain-specific knowledge contains the notion of knowing how as well as the notion of knowing that. This means that domain-specific knowledge contains *all* knowledge someone has about a domain, including declarative, procedural, and conditional knowledge (Alexander, 1992, refers to this as 'domain knowledge' or 'topic knowledge'). Knowledge about relations can exist at different levels, from a purely qualitative level ('a is related to b') to a quantitative level ('$a = \frac{1}{b}$'). In the extended SDDS framework, Van Joolingen and De Jong (1997) divide hypothesis space into a space of variables and a space of relations. Generic knowledge is needed to understand particular types of relations, which means that the level of generic knowledge will influence the precision of relations that a student can work with.

In Chapter 4, both domain-specific knowledge and generic knowledge are operationalized. Tests are created to measure both types of knowledge before students start working with the Optics discovery learning environment, which is described in Chapter 3.

2.4 Summary

In this chapter, an overview was given of the theoretical considerations that were taken into account in designing the studies that are described in the next chapters. The SDDS model of scientific discovery learning forms a basic framework, which lately has seen a number of extensions in the number of search spaces that are proposed (e.g., Klahr, 2000) and in a detailed description of the configuration of these search spaces (Van Joolingen & De Jong, 1997). An overview was given of ideas about the role that prior knowledge can play in scientific discovery search processes. In the empirical part of this dissertation, these ideas are put to the test. First, a detailed outline of the method that was used in the experimental studies will be given. In Chapter 3, the Optics computer simulation that was used as a discovery learning environment is described. In Chapter 4, the two types of knowledge that were distinguished in this chapter, domain-specific and generic, are operationalized. The resulting knowledge tests were used as measures of prior knowledge in the empirical studies that are described in Chapters 5, 6, and 7.

3

The Optics computer simulation: design and implementation

> It may well be doubted whether human ingenuity can construct an enigma of the kind which human ingenuity may not, by proper application, resolve.
>
> *Edgar Allan Poe, The Gold Bug (1843)*

3.1 Introduction

In the Middle Ages, a torturer would always make a point of showing to the victim the torture instruments first, before starting out on the other tasks. In a similar fashion, in the present chapter and in Chapter 4 the instruments that were used in the studies are presented, before the studies that form the main aspect of this dissertation are described. In the present chapter we describe the computer simulation that was used in the studies in chapters 5, 6, and 7. Specifically, the choice of topic, the type of interface, and the underlying model of the computer simulation are discussed. Chapter 4 describes the experimental setting of which the simulation is a part. It is important to notice that there are some fundamental differences between computer simulations that are used to promote, enhance, and amplify the learning experience (e.g., discovery learning processes), and computer simulations that are used purely for the purposes of scientific interest. In both contexts, the use of a simulation is viewed upon as a beneficiary to the learning experience, because of its unique properties and the learning processes it elicits (as has been explained in Chapter 2). On the one hand, in the context of the research that is carried out in this project, features that support discovery learning, and that would otherwise be added to the instructional context, are left out, because they may interfere with measurements. On the other hand, a number of features that would normally not be present can be added to the simulation to help online measurement of computer operations while students work with the learning environment. In short, the context of the experimental research that is presented in the next chapters deviates from other studies that utilize a computer simulation as a learning environment. The focus lies not on studying the effectiveness of using a computer simulation (versus, for example, classroom lectures). Instead, we focus on the learning process itself, in a 'bare bones' discovery setting. This implies that, in this case, the

learning situation is not designed to enhance learning, but rather to optimize measurement of discovery learning processes. The simulation that was developed for this research project possesses a number of characteristics that make it suitable for studying discovery learning processes by examining the how learners operate the simulation. The method that was used to measure learner-computer interaction is described in Chapter 4 (Section 4.3).

In Chapter 1 (Section 1.3), an overview of simulation based discovery environments was given. It was stated that a computer simulation incorporates a model of some real or fictitious system. The type of system that is modeled by the simulation can be chosen to fit in a certain curriculum, or to fit the demands of a research study. Van Joolingen and De Jong (1991) make a global distinction between three types of systems that can be simulated: physical (systems present in the real, physical world), artificial (systems that present an artificially created artifact or situation), and hypothetical (systems without a counterpart in the real world). The foundation for using a simulation is the idea that a student will try to understand the underlying, hidden rules and thus gain better understanding of the simulated system. As was explained, getting to understand the rules of a simulation involves a process of hypothesizing and experimenting, and is referred to as learning by discovery. The set of underlying rules is called the simulation's *model*. The characteristics of the model of a simulation determine the nature of the variables and relations underlying the simulation. Relations between variables can be expressed quantitatively or qualitatively in the model. Variables can be independent (that is, when they can be manipulated by the learner), dependent (when their value is dependent on the simulated model and/or on the settings of one or more independent variables), or intermediate (when they are used in the underlying model, but not visible to a learner). In addition, based on the type of simulation that is used, variables in the simulation can be presented as containing discrete or continuous values. Van Berkum and De Jong (1991) argue that the model of a simulation can be either operational or conceptual. A conceptual model contains facts about the variables and relations between them that are related to the system that is simulated. An operational model contains a series of (cognitive) operations that have to be applied to make the model work. The knowledge needed to work with an operational model is usually of a conceptual nature.

The model of a simulation is not disclosed to students while they operate the simulation. Instead, students operate on the visible part of the simulation, which is called the simulation's *interface*. The interface provides the means to interaction between students and the computer by, for example, the use of icons to refer to certain operations and tools. Knowledge of a simulation's

interface is a prerequisite for learning about the underlying rules. Therefore, it is important to use an interface that is easy and intuitively to understand. The *fidelity* of a simulation roughly determines how it appears to a learner. Hays and Singer (1989) make a distinction between physical fidelity (how realistic does a simulation look) and artificial fidelity (how realistic are the manipulations that a simulation allows). In designing a computer simulation, choices have to be made about both types of fidelity.

3.2 The Optics simulation

A computer simulation learning environment (which we will henceforth refer to by the name 'Optics') was designed and created with the experiments that are described in this thesis in mind.[1] Different versions of the Optics simulation were also used in other studies that were carried out in the context of the Inductive Learning project (described in Section 1.1.1). In this section, the topic of the simulation, the interface, the underlying model, and the built-in support functions ('learning aids') are described. The purpose was to develop a computer simulation that could be processed by a normal desktop computer, and preferably on computers that are in use in classrooms. Indeed, the finished computer program can be run on a standard desktop computer that uses the Windows operating system.

3.2.1 Topic

The specific 'feel' of a simulation is determined by the model and type of interface that are chosen. When a simulation is being designed, choices have to be made with respect to the fidelity and complexity of the underlying model and the interface. The type of variables and relations, and the complexity of a domain are both influenced by the topic that is chosen for the simulation. For example, a simulation of running a chemical plant is far more complex and intricate than a simulation of the chemical reaction between two fluids. In choosing the topic of a simulation, the first important thing to consider is the type of system to simulate. For our studies, we

[1] The Optics simulation and other computer programs that were used in the experiments were developed by Jan Wielemaker (University of Amsterdam). The simulation has also been used in a study by Veenman, Prins, and Elshout (submitted).

wanted to use a simulation of a real system, preferably a topic that is taught in secondary education. Many simulations revolve around the natural sciences (physics, chemistry), because these lent themselves best to formal description. Also, because our focus lies on the influence of (conceptual) knowledge on the learning process, it was decided to build a simulation around a conceptual model. Choosing a topic that originates from secondary education has the added advantage that it may facilitate students' understanding of the computer simulation itself, because to some extent they are familiar with the things they see on the screen. For many students, learning with a computer simulation is a novel experience, but the use of a real-world topic is expected to facilitate understanding of the computer simulation itself. This is important since our focus lies not on the simulation itself, but on the discovery learning processes that it evokes. A final consideration in the choice of topic for the computer simulation is the requirement that students differ in their knowledge on the subject, prior to working with the program.

As its name already suggests, the Optics computer simulation represents the domain of optics, more specifically, geometrical optics. Optics is a physical simulation that deals with the principles that govern image formation properties of lenses. It simulates an optical workbench with which the behavior of optical systems can be studied. In a number of studies by Reimann (1990, 1991), the subject of refraction (which is an important concept in optics) was also taught to students, using a computer simulation interface called REFRACT. Geometrical optics is appropriate for use in a computer simulation, because the domain can be described by only a small number of rules and concepts. However, a number of studies in the field of science education (e.g., Galili & Hazan, 2000; Langley, Ronen, & Eylon, 1997; Reimann, 1991; Säljo & Bergqvist, 1993) have shown that geometrical optics is a challenging subject for students, and one that is prone to misconceptions. A probable reason that students find it difficult to grasp the different concepts, is the fact that light rays and the different illumination patterns they cause are, in real life, not visible to the eye. As was discussed in Chapter 1, one of the special characteristics of a computer simulation is the possibility to make objects or relations between objects visible that would otherwise remain invisible. The consequence for optics is that making light rays visible in a computer simulation should enhance understanding of the properties of light and light rays.

Geometrical optics deals with phenomena concerning light propagation through an optical system, and the creation of illumination patterns (Langley, Ronen, & Eylon, 1997; Hecht, 1998). In our case, the focus lies on optical systems that demonstrate light propagation through one or more (thin)

lenses. Light that is emitted from a light source (such as a lamp) is refracted by a lens. The angle by which a beam of light is refracted when it passes through a lens, is dependent on the incoming angle of the beam, the distance of the beam to the optical axis of the lens, and on properties of the lens. A parallel bundle of light that propagates parallel to the optical axis of a thin convex (or: positive) lens converges on one point of the principal axis; this point is called the focal point of the lens. Light shining through a concave (or: negative) lens diverges. In this case, the light beams seem to originate from one point, called the virtual focal point. The way and the extent to which this refraction occurs can be described by a number of formulas. These formulas constitute the underlying model, which is described in Section 3.2.3.

An important difference between a focal point and a virtual focal point is that the first can be readily measured using tools that are available in the environment, whereas the existence of the second can only be inferred from the refraction of light through a lens. The Optics computer simulation does contain a number of tools that aid in making visible things that are not visible in real life. In geometrical optics, this is especially the case with the so-called *virtual image*. A virtual image is one that can be looked at (by observing the image through the lens), but cannot be projected on a screen. Real images are formed by converging rays, virtual images by diverging rays (Park, 1997).

When restricted to the case of light through lenses, geometrical optics poses an appropriate real-world system that can be modeled in a computer simulation. It was decided early on in the design process to use abstract representations of light beams and lenses. Such a way of representing the information should be familiar to students, because it is similar to the way the topic is taught in schools. Both light sources and lenses were given recognizable shapes. Examples of the representation of light sources, light beams, and lenses are shown in Figure 3-1 and Figure 3-3.

3.2.2 Interface

It was decided to make use of an attractive graphical user interface (GUI) that could be operated primarily by the computer mouse. This was done to simplify operating the computer simulation, and to promote students' interest in working with it. An important part of the visual appearance of a GUI

is the use of icons to represent possible operations in the learning environment (Macaulay, 1995). Three examples of icons that are used in the Optics simulation are shown in Table 3-1. In Appendix C, a complete overview of the available icons in the Optics simulation (divided into objects and tools) is given. The style of the Optics simulation interface is designed to be intuitive, so that mistakes are kept at a minimum. The icons that are used in the simulation resemble the actual objects as they would appear in the real world. Also, the interface is designed to be as simple and uncluttered as possible, so that students are not overwhelmed by the multitude of options available (Jones & Okey, 1995).

Table 3-1. *Examples of icons and their function in Optics*

Icon	Explanation
	*Lens*. This icon represents an object. When it is clicked, a lens can be placed in the working area of the simulation. The number of lenses that can be used simultaneously is limited.
	*Lamp*. This icon represents an object. When it is clicked, a lamp can be placed in the working area. The number of lamps that can be used simultaneously is limited. There are different types of lamps, with either one (laser)-light beam, three parallel light beams, or three divergent light beams.
	*Measure distance*. This icon represents an operation. When it is clicked, a distance between two objects in the working area can be measured. The distance measure stays in the working area, which means that when one of the objects is moved around, the measured distance changes as well.

The whole learning process takes place within the Optics computer simulation itself, no external aid (paper, pencil) is needed to work with the computer program. The main features of the Optics simulation interface are summed up in Table 3-2, together with a short description of each feature.

Table 3-2. *List of characteristics of the Optics interface*

Optics feature	Description
Colorful	Interface makes use of colors to distinguish between objects
Dynamic	Updates are carried out in real-time, no 'experiment' button
Point-and-click	Operations are carried out by easy selection of objects
Icon use	Select object or operation by clicking on a recognizable symbol
Windows-like	Simulation features resemble the Microsoft Windows™ interface
Attractive	Emphasis on visual appearance
Multifunctional	More than one way to perform some operations
Animations	Some events are embellished by animated features

Figure 3-1 shows an example of the interface of the Optics computer simulation interface.

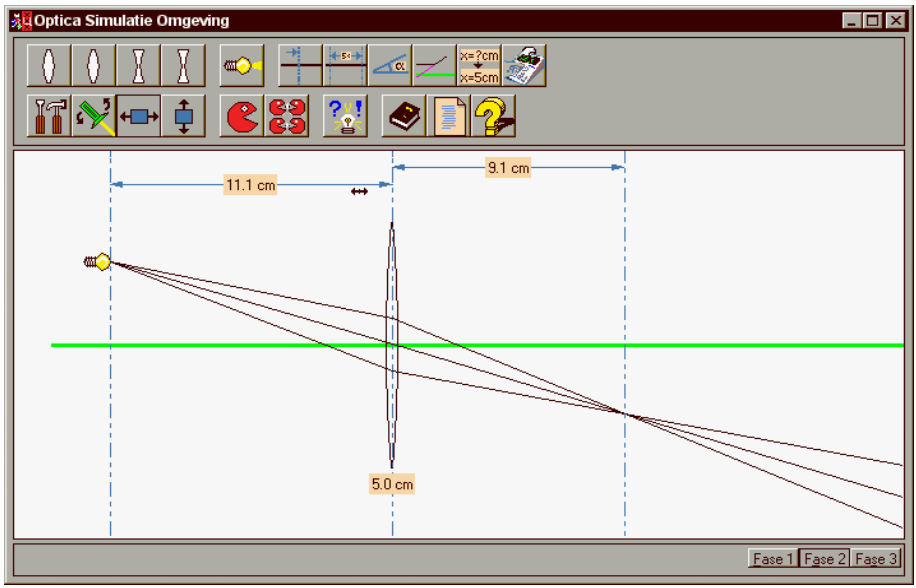


Figure 3-1. *Example interface of the Optics learning environment. Divergent light rays are refracted in different ways by a lens. Distances are shown, as well as the focal distance of the lens (5.0 cm)*

The interface consists of two parts. In the upper part icons that represent objects (that can be added to the simulation) and icons that represent operations (that can be carried out on objects in the simulation) are shown. The lower part of the interface consists of the working area. At first, this area only shows a flat (green) horizontal line on the screen. This is the principal axis (the 'base line') on which objects are placed; properties of these objects can subsequently be altered. All objects can also be removed from the working area (to allow new objects to be placed, for example). The example shown in Figure 3-1 is derived from a situation in the first study (presented in Chapter 5). The figure shows how light beams are refracted by a lens. This particular lamp has three laser-like light beams, each of which strikes the lens at a different angle. The focal point of the lens is shown underneath it (showing the focal distance of a lens is optional in the design of a learning situation). It is possible to manipulate the distance and properties of different objects that are shown on the screen. In addition to manipulating objects themselves, the distances between objects can be measured; while moving an object, the measured distance is dynamically updated in the display. In the figure, the distance of the lamp to the lens and the distance between the lens and the point where the light beams converge are shown (the last point is marked by inserting a special construction line). By moving the lamp or the lens, or by changing the outgoing angle of the light beams, the occurrence of interesting regularities and irregularities in these measures can be explored. Findings can be noted down in an online notebook. This notebook displays the current situation in the working area, along with text that the learner types. The notebook is further discussed in Section 3.2.4. In Table 3-3, an overview is given of the operations that can be performed in the Optics simulation. As can be seen from the table, what can be done in the simulation is limited to nine operations. However, since most of the operations operate on one or more objects, the actual number of possible operations is very large.

Table 3-3. *Overview of operations in the Optics simulation*

Operation	**Objective**
Add	Add an object to the working area
Delete	Delete an object from the working area
Move	Move an object around the working area
Measure	Measure the distance between two objects
Rotate	Rotate a light beam or a set of light beams
Object tool	Attach label to an object, for use in formula
Formula	Compute and display a relation between one or more objects
Notebook	Access the notebook feature
View	View instructions about the assignment or the interface

In Chapter 4, the method that was used to measure learning behavior in the empirical studies that made use of the Optics simulation is described. Two features of the Optics simulation, the first of which is visible in the interface, make it possible to register certain operations. Firstly, when an object is moved around any numerical values that are present in the working area of the Optics simulation are replaced by question marks, and are only made visible again when a 'show values' button is pressed. This makes it possible to register the moments when a student is interested in one or more values. Second, a difficulty with the Optics simulation is that it is possible for a student to perform measurements by moving an object, performing the measurement, and then to move the object to another position, continually keeping the mouse button pressed. To register this type of behavior, instances of holding an object in one position for some time were also registered.

3.2.3 Underlying model

The underlying model follows the actual physical laws of light refraction through one or more lenses. The angle of refraction is computed such that the ideal properties of the lens are met, i.e. that for a positive lens a parallel bundle of light rays converges in a single focal point. By working with the environment it is possible to discover these laws. In physics texts, the characters, *v*, *b*, and *f*, are usually used to denote object distance, projection distance, and focal distance. Figure 3-2 shows a schematic description of the conceptual nature of *v*, *b*, and *f* in geometrical optics. Moving an object around changes the properties of *v* and *b*. By measuring distances between light sources and projections, the relation between these variables can be revealed. Important facts can be discovered about the differences between positive and negative lenses, and about the existence and position of the focal points of a negative lens. There is a quantitative relation between the focal point distance and the refraction of light beams through a lens.[2] In addition, the relation between the order of magnification and object distances can be sought out. This relation is shown by the formula

1) $N = \left| \frac{b}{v} \right|$

[2] For the description of formulas in this section we made use of Hecht (1998).

N expresses the order of magnification. When *N* has a value between 0 and 1, a projection is smaller than the projected object. When *N* is greater than 1, the projection is larger than the projected object. *N* is exactly 1 (no magnification) when *b* and *v* are equivalent. In addition to the magnification formula, it is possible (in principle) to discover the so-called Gaussian lens formula. This formula specifies the relation between the distance of objects from a lens and focal distance of the lens. The Gaussian lens formula is expressed by the relation

2) $\frac{1}{v} + \frac{1}{b} = \frac{1}{f}$

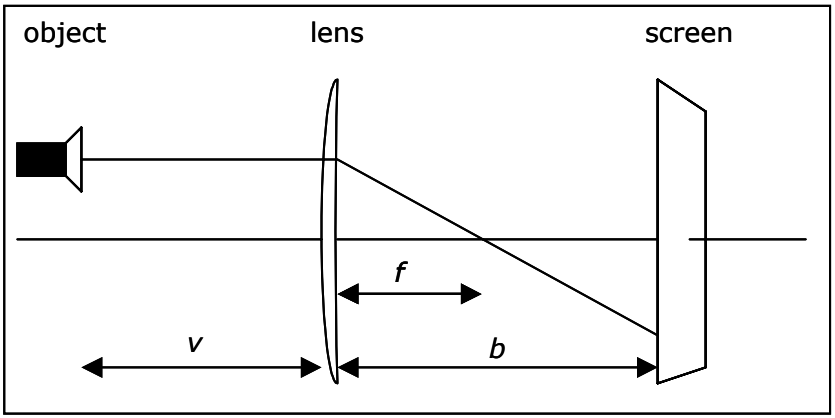


Figure 3-2. *Schematic drawing of the properties of v, b, and f*

When measurement tools are used, the variable that is measured (such as the distance between two objects) is displayed as a quantitative value. Without measurement tools, qualitative relations between variables can be inferred by varying the distance of objects from a lens, and by comparing different lenses. The magnification and general lens formulas represent the complete formal description of the optics domain as it is used in the Optics simulation. It might seem that the underlying model for the Optics simulation is, at best, limited in its scope. However, we feel that this is not the case. Understanding the properties of the lens formula means understanding the properties of light through a lens. Under special circumstances, for example when *v* is smaller than *f* (which means an object is closer to a lens than its focal distance), a real image changes into a virtual one. From the formula, it can be seen that in such a case *b* holds a negative value. To discover these

properties from working with the computer simulation presents quite a challenge. In reality, we do not expect learners to find out all relations between objects in Optics at a quantitative level. As was explained in Chapter 2, learners are expected to get a global, qualitative understanding of the properties of the simulated model. In Chapter 4, a test for domain-specific knowledge about optics is introduced. This test operationalizes the ideas that have been put forward in Chapter 2 by focusing on qualitative aspects of knowledge.

3.2.4 Learning support

Scientific discovery learning with a computer simulation can be supported in various ways. This can be done, for example, through the use of specific assignments, by using some form of model progression, by guiding learners in their decisions, and by monitoring the behavior of learners and providing feedback (Swaak, Van Joolingen, & De Jong, 1998).

In the experimental setting that the Optics simulation offers, learner support is kept at a minimum level. This so-called 'bare bones' approach was chosen for the following reason. As was explained in Section 1.1, the goal of this research project is to study the influence of determinants of scientific discovery learning processes. To get an understanding of the influence of prior knowledge on discovery learning processes themselves, it is considered to be of importance to keep the processes that take place during discovery learning 'clean', that is, free from interfering influences, such as available support. In the first two studies that are described in this thesis, the support for learning in the Optics simulation that is given to learners is (apart from instructions about the Optics interface) limited to the availability of assignments, and the ability to make notes while working with the learning environment. In the study that is described in Chapter 7, the effect of (limited) online support is tested. In that study, a group of subjects can make use of information which they have available while they are working with the computer simulation.

Here, we shortly discuss two types of support that are used in the Optics simulation: model progression and the online notebook.

Model progression. Different situations in Optics consist of different configurations. This means that in each configuration, a different set of tools can be made available to learners. Also, for each configuration a different goal can be set, or an overall goal can be used. When model progression is used, there is an option that forces learners to work through the different situations in sequence. When a 'forced sequence' is used, a new situation in the simulation is only made available to a learner after a certain amount of time has passed. In the experiments that made use of model progression (the Optics 97 and Optics 99 studies), this option was used. The Optics 2000 study made use of only one situation. Possible configurations in Optics range from very simple to complex. At its most simple, the only available objects are a lamp with one light beam, and a lens. This is the type of situation that was depicted in Figure 3-1. A more complex type of situation is shown in Figure 3-3. The special feature of this situation is that a big lamp (at the left) shines light through a plate that has holes in it (the holes in the plate form the letter L). The light through the plate is projected (by means of a lens) on a screen (at the right of the figure). This way, the properties of focused and unfocused images can be examined. By using model progression, the learning situation can gradually be extended in its complexity.

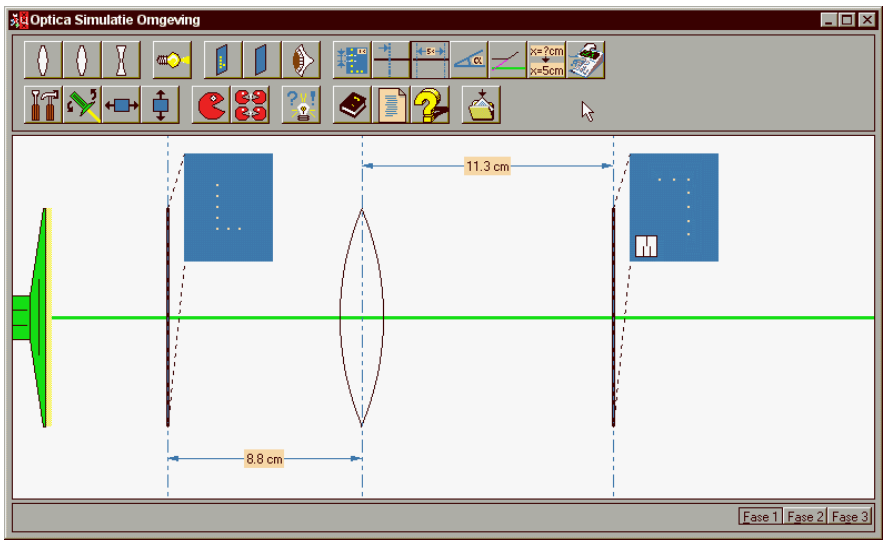


Figure 3-3. *Situation in Optics, where a figure is projected through a lens on a screen*

Online notebook. For the Optics simulation, an online notebook was developed which allows learners to make notes about their findings. The online notebook can be accessed while the simulation is running. In the notebook, the situation that is visible to the student in the working area at the time of making the note is visible. An example note is shown in Figure 3-4. In the figure, the top part shows the current situation in the Optics simulation. Current note number is shown at the bottom left, and text can be typed in the area at the bottom right Allowing learners to make notes while they work with the Optics simulation makes it possible for them to keep track of results that they have found. They can browse the notes they have made, and get a snap-shot of the relevant situation in Optics. In practice, the notebook serves as a memory aid, which could possibly reduce cognitive load (Sweller, 1994).

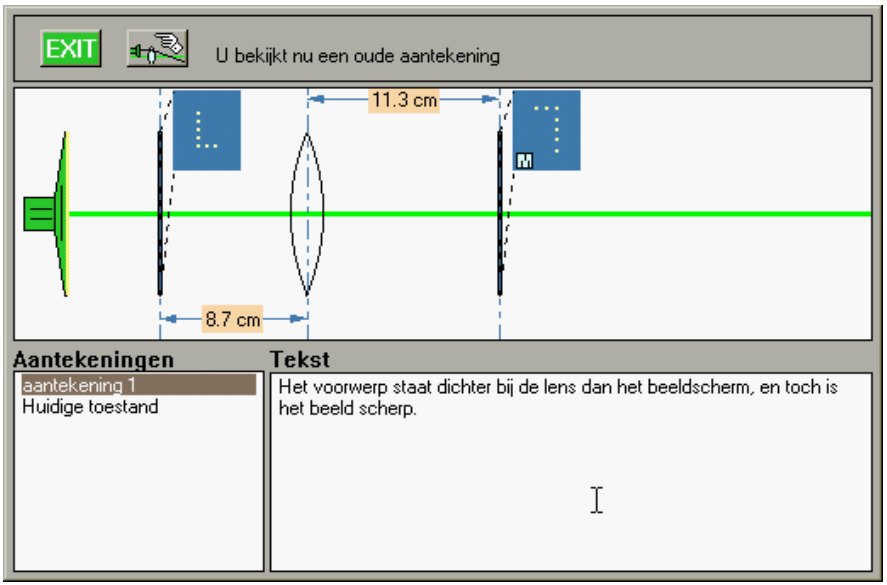


Figure 3-4. *Notebook example in Optics. The note states 'The object is positioned closer to the lens than the screen is, but still the image is sharp'.*

3.3 Summary and conclusions

In this chapter, an overview of the Optics computer simulation was given. The Optics simulation was developed for the studies that are described in the following chapters. Optics deals with the topic of geometrical optics. The underlying model and the interface of the computer simulation and the learner support that was used, were determined with the research goals in mind. This has implications for the instructional efficacy of the simulation. Because the goal of using the simulation was to enable us to study the influence of prior knowledge on discovery learning processes, a suitable topic with a simple underlying model was chosen. The interface was designed to be intuitive and easy to use. In the experiments, learner support was kept at a low level. The Optics simulation allows situations to be designed in which online support through model progression and an online notebook are used. Both of these types of support were used in the first two experiments. However, the configurations of Optics that were used in the experiments were not meant to provide an optimal and effective learning situation. Rather, the design was optimized for experimental study. This does not imply that the Optics simulation is not, in principle, suitable for learning about geometrical optics. When the simulation is suitably configured, it can be included as part of the standard physics curriculum. A curriculum into which the Optics simulation would be embedded, would involve using more guided assignments, more levels for model progression, and feedback about the quality of experiments that are performed.

4

Experimental measures

You know my methods. Apply them.

Sir Arthur Conan Doyle, The Sign of Four (1890)

4.1 Introduction and research questions

In Chapter 2, the hypothesis was stated that the amount and quality of prior knowledge that learners have at their disposal can influence scientific discovery learning processes. This idea is explored in the empirical studies that were carried out in this project. The Optics computer simulation that students worked with in the experiments was described in Chapter 3. In these experiments, domain-specific and generic knowledge were measured before students started to work with the simulation. In some cases, knowledge was manipulated during the discovery learning task. In the present chapter the measurement of both types of knowledge is described. Also, the method that was used to measure learner-computer interaction in the empirical studies is discussed in the present chapter.

Structure of this chapter. First, in this section an overview of the research questions will be given. By using the discussion on prior knowledge and discovery learning processes in Chapter 2, more detailed research questions can be stated than was the case in Chapter 1. In each of the experimental chapters further detail will be given. In Section 4.2, the ideas about prior knowledge and discovery learning that were discussed in Chapter 2 are put into practice, and the operationalization of domain-specific and generic knowledge is discussed. The knowledge tests that resulted from this operationalization were used in the empirical studies. In Section 4.3, the operationalization of learner behavior and the design of the analysis of Optics interaction data are discussed. While students worked with the Optics simulation, their interaction with the computer was measured. Data about the interaction behavior of students was analyzed in order to find differences in discovery learning processes between groups of students. Considerations that led to using the method of online registration of interaction behavior (as opposed to other experimental methods, such as protocol analysis) are also discussed.

Research questions. The arguments that have been put forward in Chapter 2 pose a number of questions. It was argued that from the perspective of the SDDS framework of discovery learning (and also the extended version of the framework), domain-specific knowledge and generic knowledge determine the configuration of the hypothesis search space which is traversed during learning. In our view, the search spaces that make up the hypothesis space are not static instances in a student's mind, but have a dynamic nature in that they can change over time as a result of learning. One question is if and how discovery learning processes are regulated by the presence or absence of prior domain-specific and generic knowledge. To be more precise, the chapters that cover the empirical studies deal with the following research questions:

1) Do the domain-specific and generic knowledge influence discovery learning processes? Are there differences between the influence of domain-specific knowledge on the one hand, and generic knowledge on the other hand?
2) Do domain-specific and generic knowledge influence the effectiveness of discovery learning? Does more prior knowledge lead to more successful discovery learning?
3) Does the discovery learning behavior change over time? How is the strategy that is followed dependent on domain-specific and generic knowledge?

These broad questions have served as a guideline for the design and performance of the empirical studies that are described in the following chapters. To measure students' domain-specific and generic knowledge, tests were developed. In the following section, considerations that went behind the design of the tests are discussed.

4.2 Measuring generic and domain-specific knowledge

The empirical studies that are described in this thesis all made use of a test for generic and domain-specific knowledge. Over experiments different versions of the tests were used. The main thoughts that went behind the original design of the tests are discussed in this section.

4.2.1 Generic knowledge

In Chapter 2, the concept of generic knowledge was explained. Generic knowledge encompasses knowledge that is needed to recognize and work

with numerical and graphically depicted relations between two or more variables. Knowledge of mathematical relations means for instance knowledge about the behavior of relations, like asymptotes and periodic relations, and knowledge that a linear relation is also a monotonic relation. Thus, generic knowledge encompasses all knowledge that is needed to understand qualitative or quantitative relations in a very general sense. Mathematical relations can be represented in various ways, for example, as points in a graph, in a formula, or in generic terms such as 'monotonic decreasing' or 'exponential'. Knowledge about different types of relations can exist at various levels. Van Joolingen and De Jong (1997) state that a main distinction can be made between qualitative and quantitative relations, with knowledge on a qualitative level (e.g., 'A and B are related in some way') being less precise than knowledge on a quantitative level (e.g., A = 2 B). A higher level of generic knowledge implies that more understanding of quantitative relations exists, and thus a better ability to understand these relations. Knowledge about the type of relation that exists between two or more variables is needed to be able to state a hypothesis (either qualitative or quantitative) about it. Also, to understand relations that are implied by experimental findings, knowledge about the type of relation they represent is necessary. When findings from a series of experiments are plotted in a graph, generic knowledge is needed to understand the mathematical relation that is depicted. For example, if experimental findings point to an exponential relation between two variables, generic knowledge is needed to recognize the presence of such a relation to understand the implications. Goldenberg (1988) has noted that students often experience difficulty in interpreting what they see in a graph, and that this can cause them to induce wrong relations.

The considerations that are discussed here led to the creation of a test for generic knowledge. The test items were inspired by two tests for 'scientific processing': TIPS II (Test of Integrated Science Process Skills II; Burns, Okey, & Wise, 1985) and TOGS (Test of Graphing in Science; McKenzie & Padilla, 1986). Although all test items used a multiple-choice format, the items tapped various abilities. In Figure 4-1, three example items from the test for generic knowledge that was used in the Optics 99 experiment are shown.[1] In the figure, item 1 refers to a quantitative relationship between two variables

[1] The items that are shown were translated from Dutch. The complete test for generic knowledge is shown in Appendix A.

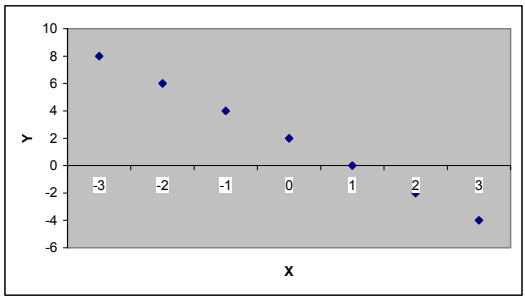
X and *Y*. The relationship is represented in numbers. Item 2 also uses numbers to represent a relationship between two variables *A* and *B*, but the question refers to the qualitative relationship between the variables. Item 3 uses a diagrammatic presentation to depict a relationship between *X* and *Y*, and the question refers to a qualitative relationship between the variables. Each variation of a question (qualitative/quantitative/diagrammatic relationship represented, qualitative/quantitative question asked) was used a number of times in the test.

1. Which function is schematically described by X and Y in the table?

X	Y	*Answer*
1	1	□ $Y = \frac{1}{2}X + 1$
2	1,4142	
3	0,7321…	□ $Y = \frac{1}{X^2}$
4	2	□ $Y = \frac{1}{X}$
		■ $Y = \sqrt{X}$

2. The table shows for some values the relation between variables A and B. Which description on the right best summarizes this relation?

A	B	*Answer*
10	4	□ The more A increases, the slower B increases
20	8	□ When A is zero, B is also zero
30	16	□ The more A decreases, the faster B decreases
60	64	■ The more A increases, the faster B increases
100	1024	



3.
What type of function describes a line that goes through all the points in the graph on the left?
□ A quadratic function
□ A function with an asymptote
■ A monotonic decreasing function
□ A logarithmic function

Figure 4-1. *Example items from the generic knowledge test. Correct answers are made black in the figure.*

4.2.2 Domain-specific knowledge

Domain specific knowledge is defined as knowledge about the existence and characteristics of variables and relations in a domain. For example, in the domain of geometrical optics this would entail knowledge about the existence of focal points, and knowledge about the variables that a focal point relates to (such as focal distance). A student's level of domain-specific knowledge can be tested for in a number of ways. Usually, a distinction is made between open-ended questions and multiple-choice questions. However, in the context of scientific discovery learning with the Optics simulation, the test that was developed to measure domain-specific knowledge differed on a number of characteristics from a standard test. The reason for using a different type of domain-specific knowledge test is based on findings of research on effects on knowledge of discovery learning. In a large number of studies it has been argued that discovery learning has important advantages over the classical expository kind of teaching (Banger-Drowns, Kulik, & Kulik, 1985). The idea is that the knowledge structure acquired in active learning is qualitatively different from the knowledge that is acquired in other kinds of learning contexts. For example, Ausubel (1963) argues the following about meaningful discovery learning:

> 'When the learning task is more difficult and unfamiliar, autonomous discovery probably enhances intuitive meaningfulness by intensifying and personalizing both the concreteness of experience and the actual operations of abstracting and generalizing from empirical data.' (1963, p. 143)

A few decades before Ausubel made this statement, Vygotsky (1987) had already argued that 'Pedagogical experience demonstrates that direct instruction in concepts is impossible. It is pedagogically fruitless'. The special properties of the knowledge acquired in discovery learning situations are referred to in a number of ways in the literature. The knowledge that results from discovery learning is either considered to be intuitive (Fischbein, 1987; Lindström, Marton, Ottosson & Laurillard, 1993; Swaak & De Jong, 1996), implicit (Berry & Broadbent, 1988), contextual (Tennyson, 1990; Elmore & Tennyson, 1996), functional (Leutner, 1993), or tacit (Wagner & Sternberg 1985; 1986). It is generally agreed that the knowledge that results from discovery learning is different from learning in an expository context in the sense that the knowledge is deeper rooted and more flexible (Bruner, 1974; Wittrock, 1966), more readily viable (Bruner, 1961), more qualitative (Peterson, Junck, Sharp, & Finzer, 1987), more contextual (Tennyson & Rasch,

1988), leads to enhanced transfer (Bruner, 1974), and encourages reflective thinking (Norman, 1993).

Thomas and Hooper (1991) have argued that most knowledge tests that are used to measure the effect of simulations ask for the wrong things. Instead of asking for definitions that require recalling information, tests should try to measure 'transfer and application'. Developing new ways of assessing abilities is also a point of argument for Grabinger (1996), who speaks in this respect of the 'more realistic and holistic form' that assessment must take in general (p. 667). To this end, Swaak and De Jong (1996; also see Swaak, 1998) have developed a test, called the 'what-if'-test that aims at measuring intuitive knowledge. The test finds its origin in a dissatisfaction with the link between the context of learning in a discovery situation and the methods that have been used to assess resulting qualities of domain-specific knowledge. Swaak and De Jong argue that the quality of knowledge that results from discovery learning is not the same as knowledge that results from learning in a more expository context, and that a standard test cannot provide a valid measure of domain-specific knowledge.

The main idea behind the 'what-if'-test is the notion that the knowledge that is constructed by learners when they work with a discovery learning environment has a special, intuitive, quality, which can be characterized as comprising 'quick perception of meaningful situations' (Swaak and De Jong, p. 346). The test items that were developed by Swaak and De Jong try to tap this knowledge by offering a meaningful[2] situation and asking for a quick evaluation of a change in this situation. Also, the items use only simple words and short phrases, and technical terms are generally avoided. Before a what-if test is administered, students are asked to balance their speed of answering with the correctness of the (multiple choice) items, that is, they are asked to answer the items as fast as possible, but still correctly. Administering a what-if test yields data on the answers to the test items, and data on the time taken to answer an item.

The test for domain-specific knowledge that was developed for this research project, and that was used in the Optics 97 experiment, closely followed the format that was introduced by Swaak and De Jong.

[2] 'Meaningful' here refers to the context of the computer simulation that students work with.

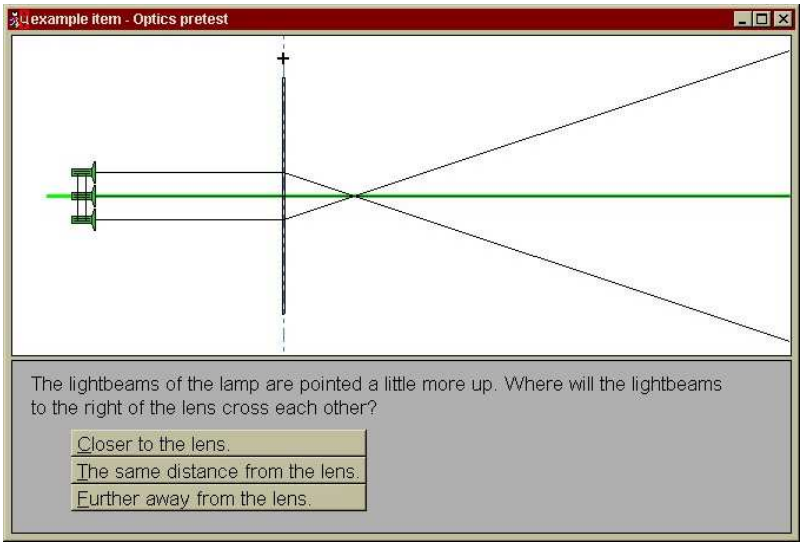


Figure 4-2. *Example what-if test item from the Optics pretest (explanation in text)*

The items used a three-answer multiple choice format, and for all items a similar design was used. Figure 4-2 shows an example of a what-if test item as it was used in the Optics 97 experiment (for the present thesis, the figure was edited a little in order to fit). As can be seen from this example, the items picture a situation that can occur in the actual learning environment. In this case, the question states, 'If we replace the lens [that is used in this situation] by a weaker one, how should we then move the screen [on which the object is projected] to get a sharp picture?'. The answer possibilities are: 1) further from the lens, 2) closer to the lens, and 3) position should remain the same. Other items used the same kind of configurations, and some items used configurations with two lenses. Items were presented in random order: at the moment learners picked the answer of their choice, the item would disappear and the next (random) item appear. It was not possible to go back to a previous question after it had been answered. In the empirical studies, what-if tests were used both as tests for prior domain-specific knowledge and as posttests. In the Optics 99 and Optics 2000 studies, the test was modified in a number of ways to increase variation and reliability. The complete domain-specific knowledge test that was used in the Optics 2000 experiment is shown in Appendix B.

4.3 Measuring learner behavior

> A man, who is a heavy drinker, doesn't feel too well and goes to visit his doctor. The doctor says, "You drink too much. To begin with, you should reduce your daily alcohol intake by ten glasses of beer. In two months we'll see if your condition gets any better." To which the man replies: "Can't I instead drink ten glasses more every day, and see if it gets any worse?"
>
> *Dutch Joke Calendar, June 16th 2000*

The joke that is printed above, though admittedly a bit corny, neatly illustrates an important methodological problem: what is the best way to search for experimental effects? The use of the Optics computer simulation allows us to study discovery learning processes in a variety of ways. It was chosen at an early stage to focus on the operations that learners perform while they work with the learning environment. The studies that are reported in this thesis measure discovery learning behavior by an online registration of operations. This provides an unobtrusive behavioral measure. Although the data that is obtained by using registration of operations is less 'rich' than verbal data, cognitive processes can be inferred by combining different methods. The topic of the current section is firstly, the method by which learner operations were registered in the empirical studies, and secondly, how this data was subsequently analyzed.[3] Our reason for discussing the analysis of learner behavior in such detail, can be succinctly summarized by citing Schoenfeld (1992), who states that

> 'Researchers putting forth new methods, or results based on new methods, have a responsibility to explain the methods—to describe where they come from and how they can be used, and to characterize their strengths and limitations (p. 181).'

The most commonly used measures to study learner-computer interaction are think-aloud protocols and online recorded behavior (Newell & Simon, 1972). The use of thinking-aloud protocols has a long history in research on

---

[3] The discussion that is related in this section can be applied to the Optics simulation, but also to the other computer programs that were used throughout the empirical studies that are described in this dissertation, because they registered learner behavior in a similar way. This implies that parts of the discussion also apply to the Bubbles simulation (described in Chapter 5), and the Peter task (described in Chapter 6).

higher mental processes (e.g., problem solving). Think-aloud data has proven to be extremely useful, since it carries a wealth of information on the way people reason about a problem, and about the specific cognitive steps that are performed during the interpretation and solving of a task. However, in the case of computer simulations, studying learning processes through the method of thinking-aloud measures can be a less appropriate research method (De Mul and Van Oostendorp, 1996). For example, it is not possible to study a classroom learning situation this way. Also, learning processes in the artificial setting of the laboratory may differ from learning in a normal classroom setting.

In the studies that are presented in this thesis, protocol analysis was not used. Instead, all actions students performed while they worked with the Optics computer simulation were registered by the computer. This means that all button presses (that is, clicking on one of the mouse buttons) were recorded, together with the offset time in seconds (a timer started as soon as the student started working with the simulation).

4.3.1 Overview of learner-computer interaction analysis

The analysis of learner-computer interaction poses a number of problems, some of which may until now not have been given fair credit in the literature. One reason for this lack of attention is that the analysis of learner-computer interaction in complex learning environments, through the online recording of actions, is a relatively new phenomenon. The complexity of the task makes interpretation of interaction data a complex affair. We will illustrate this by focusing on two studies, which, though very different in nature, have both drawn conclusions about discovery learning. Firstly, we look at the BigTrak study by Klahr and Dunbar (1988), which was also discussed in Chapter 2, and which is interesting for our purpose because of its task constraints which facilitate the interpretation of student behavior. Secondly, we look at the Smithtown study by Shute and Glaser (1990), in which a complex computer simulation was used to study differences between successful and unsuccessful learners, and which used learner-computer interaction measures to do so.

In the BigTrak study by Klahr and Dunbar (1988), described in Chapter 2, the distinction between two strategies in discovery learning (experiment or hypothesis based) was grounded in the way subjects came up with the right

hypothesis (which involved making a 'frame-shift'): either by use of an experiment that pointed in the right direction, or just by coming up with the idea to make a change of frame. Thus, Klahr and Dunbar based their distinction on the experimental actions that led a learner to a specific conclusion about the underlying rule. In their case, it was possible to do this because the complete experiment space of the BigTrak 'environment' could be described, with each possible experiment falling into a subset of experiment space. The actual frame shift involved either doing an experiment in a new subspace and reaching the right conclusion (the experimenter approach) or reaching the right conclusion without having done any experiments in the 'right' subspace (the theorist approach).

What the BigTrak experiments by Klahr and Dunbar make clear is that in order to be able to make distinctions between different styles of learning, it is useful to have a thorough overview of the experimental domain, that is, the dimensions of experiment space and the actions that are possible within the domain. However, the domains that are normally used to teach a subject are very much unlike the (relatively) simple[4] BigTrak domain. First, there is usually not one rule to discover, but a (complex) set of interrelated rules. Second, the domains themselves offer a lot more possibilities for interaction, and the number of variables that can be manipulated is larger.

Shute and Glaser (1990), in their study with the (also relatively complex) Smithtown domain, were able to derive global learner differences on the basis of learner interaction measures, and without a thorough analysis of the domain itself. The analysis pursued by Shute and Glaser focused on the large-scale extraction of clusters of interaction patterns. Their method was to take operations together (called 'learning indicators') to distinguish between global processes that take place during learning. The learning indicators were derived from both log files and protocols, and they all had to do with the number of times a certain operation was carried out (that is, its raw frequency). They were loosely categorized into three types of 'rational categories': general activity level (e.g., total number of experiments), data management skills (e.g., total number of notebook entries), and thinking and planning behaviors (e.g., number of variables changed per experiment). The categorization of learning indicators is subject to doubt, the main problem being that the categories were too broadly defined. The approach by Shute

[4] 'Simple' here refers to the relatively small experiment space of the BigTrak domain, and the fact that only one underlying rule had to be discoverer. Overall, there appears to be no clear dividing mark between simple and complex simulation environments.

and Glaser can be summarized as involving 1) counting frequencies of actions, 2) categorizing them into 'meaningful' units, and 3) making comparisons across groups (which, in the case of Shute and Glaser's experiments, was done without the use of statistical methods of analysis). The approach by Shute and Glaser can be helpful to compare groups of subjects at a global level. However, it can also be a misleading approach, because differences between learners can even each other out. When good and poor learners are compared, other factors such as domain-specific and generic prior knowledge may have influenced learner interaction. In addition, differences between groups as they occur over time are ignored by this approach.

So far, the discussion of learner-computer interaction has centered on properties of the task domain (BigTrak versus more complex discovery learning environments), and properties of interaction analysis (Shute and Glaser's 'learning indicators'). In the remainder of this section we discuss a more generic approach to the analysis of interaction in discovery learning.

There are roughly two types of analysis of log files. These are *frequency-based* analysis, and *sequence-based* analysis. In a frequency-based analysis, the total number of all operations is counted, that is, the frequency of occurrence of each possible operation in the learning environment is determined. The frequencies of operations can be compared within subjects as well as among subjects. Frequency data provides information on the overall activity of learners, as well as on the global focus of learning. Sequence-based analysis goes deeper than frequency-based analysis into the actual processes that take place during learning. It focuses on finding sequential patterns that occur in log files, and thus focuses on behavior as it unfolds through time (see Bakeman & Gottman, 1997; Bakeman & Quera, 1995). A basic assumption in sequence-based analysis is that certain operations are more likely to follow one another than other operations. This can be checked by determining conditional probabilities for sequences of operations (see Karasavvidis, 1999, for an explanation of this procedure). Sequential data provides information about the learning process at a higher level than that of individual operations. Thus, the unit of analysis in sequence-based analysis lies at a higher level than in frequency-based analysis. However, the distinction between frequency-based and sequence-based analysis does not necessarily have to be as sharply defined as the descriptions given here. One way in which both types of analysis can be combined is when frequencies of sequences of operations are studied. We argue that this is a method that adds up the advan-

tages that each of the two methods has on its own. In addition, an analysis that looks at frequencies of behavioral patterns can provide an account of what happens during learning, which helps in coping with a number of problems that occur in the extraction of relevant data from raw log files. In the analysis of interaction patterns in the Optics computer simulation, most of the time a frequency-based approach was used. However, in a few cases a frequency-based approach was combined with a sequence-based approach. The way this was implemented in the analysis of the empirical studies is discussed in the following section.

4.3.2 Learner-computer interaction analysis in the Optics studies

The operations that were performed in the Optics simulation and the time at which they occurred, were stored by the computer simulation in a log file. The raw (that is, unformatted) log file that resulted from a learning episode in which a learner worked with the Optics simulation can be seen as a continuous stream of individual operations performed in the learning environment. In the Optics simulation, raw log files included information about the values of variables, the general state of the learning environment, and the time (in seconds) when an operation was initiated. An overview of possible operations in the Optics simulation is listed in appendix C and in Chapter 3 (Section 3.2.2). In the present section the analysis of Optics interaction data is illustrated by examining a short excerpt of an actual log file (the excerpt was chosen as a typical example of data from the Optics 97 study). In Figure 4-3 a short excerpt of a log file for one subject is shown. In the excerpt, a number of operations are performed: after everything is deleted from the working area, a lamp is added, its light beam is rotated, a help text is viewed, and finally a lens and a screen are added to the working area at different positions. All these operations together take place in less than 50 seconds (the excerpt begins at 478 seconds, and ends at 524 seconds).

```
@(478.0, action(delete_all)).
@(490.0, mode(one_lightbeam)).
@(492.0,
  add(l2 = lamp3(switch(true),
                 angle(80),
                 divergence(5),
                 pos_x(3),
                 pos_y(0.1),
                 instrument_name(one_lightbeam)))).
@(493.0, mode(rotlamp)).
@(495.0 - 498.0,
  rotate(l2,
         [ @(0.0, drag(-2.00955)),
           @(1.0, drag(-0.498212)),
           @(2.0, drag(0.498212)),
           @(3.0, drag(0))
         ])).
        .
        .
        .
@(521.0, mode(screen)).
@(524.0,
  add(s5 = shield(pos_x(7.65),
                  unit(1),
                  instrument_name(screen)))).
```

Figure 4-3. *Excerpt from an Optics log (a 'raw' log file, translated from Dutch). Explanation given in text*

As can be seen from the figure, a lot of information (time, type of operation, operation parameters) is stored in a condensed format in a log file. The number of possible operations in the simulation was limited, but each operation used its own set of parameters. In Table 4-1, the operations that occur in the log file excerpt are listed together with the type of parameters that go with each operation.

Table 4-1. *Sample of operations and their parameters in the Optics simulation*

Operation	Parameter(s)
Add, Delete, Move	Object type, Object label, end x,y-position
Drag	Object label, distance
Rotate	Object type, Object label, number of degrees
Check values	--
Action	General action type (e.g., delete all objects)
View	Instruction type

As can be seen from the table, the VIEW operation included a parameter for the type of instruction that was (re)viewed by the student. Other operations, notably ADD, DELETE and MOVE, could only be performed on one of the objects in the working area, which means that the specific object on which the operation was performed was a parameter for that operation.

To be able to analyze log files, they first had to be converted to a format that was easier to handle within different computer tools. The important thing in converting the log data to a numeric format was that as few information as possible should be lost in the conversion process. In converted data, some specific information that was not in the original log data was added, such as a number given to students, and the duration of operations. Figure 4-4 shows an example of a converted Optics log file. The data that is shown in the figure is a conversion of the excerpt that is shown in Figure 4-3.

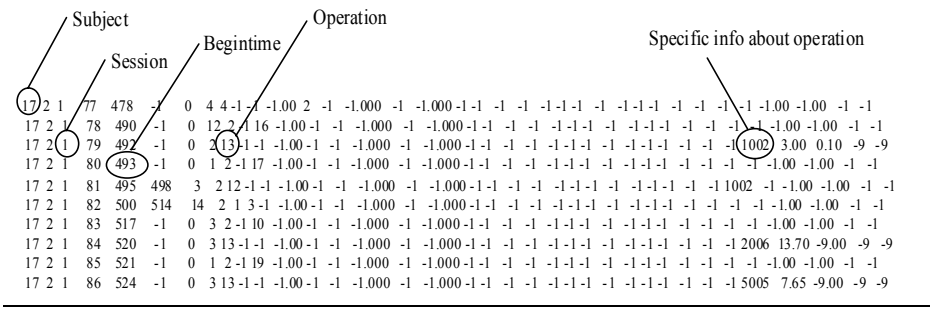


Figure 4-4. *Optics log file, converted to a format ready to be processed*

Log files of the type presented in Figure 4-4 were processed (using a special text processing language) in order to convert them to a more usable format.[5] In the converted version of a log file each operation is treated separately on one line. As can be seen from the figure, extra information that is given with each operation is the subject number, the experimental session number, the starting and ending time of the operation, and details about the operation itself (type of operation, and parameters). Two things were not used in the conversion process: the DRAG operations (that are themselves part of MOVE operations, and that refer to an object being dragged on the screen in different directions without a release of the mouse button), and the content of notebook entries. Notebook entries were extracted from the raw log files. By using the format as specified in Figure 4-4, a frequency-based analysis became a relatively simple procedure. Since each column maps to a specific operation, the main task in this case was to check the frequency of numbers in the columns in the matrix of numbers. The values of different groups of learners could be combined to enable group comparisons of frequencies. In the following chapters, this procedure is used.

Next to frequency-only based analysis, learner-computer interaction was analyzed by checking the frequency of particular *sequences* of operations. This means that a combination of frequency-based and sequence-based analysis was used. The goal was to look for patterns in the learner-computer interaction data that could be linked to the theory of scientific discovery learning. A pure sequence-based analysis (which looks for any pattern that occurs with a higher-than-chance regularity) was not performed. With respect to the behavioral patterns that were devised, we made a distinction between qualitative and quantitative sequences of operations. The experimental behavior displayed by learners in the Optics simulation usually involves adding objects to the working area, moving them around, and deleting them after some time. This behavior can be classified as qualitative, because the data that is obtained by working in this way does not carry numeric information. Qualitative sequences of operations that can be distinguished are:

---

[5] The Perl script language was used to convert the Optics log files. Initial programming of the conversion scripts was carried out by Hedderik van Rijn. Correctness of the mapping between the raw log files and the converted files was checked by a student-assistant for the occurrence of any mistakes: none were found.

1) Move object A, move object B, then move object A again
2) Add object A, and delete object A (immediately)

In this list, objects A and B can be either a certain lamp or a type of lens. Experimental, qualitative behavior, can be opposed to behavior which aims at obtaining quantitative information about the relations between objects in the working area. This can be done by adding distance measures that can be checked. When an object in the working area is moved, displayed values are changed by the simulation into question marks. By performing a CHECK VALUES operation the numbers reappear. This configuration of the learning environment allows for quantitative operations to become visible in the Optics simulation log file. Quantitative sequences of operations that can be distinguished are:

1) Add a distance measure between two objects, then delete it
2) Add a (vertical) construction line, then add a distance measure between it and an object
3) Move an object (values change to question marks), then check values

In summary, analysis of the interaction with the Optics simulation can be performed by looking at the operations that students perform while they work with the simulation. To do this, raw log files can be converted to data that can be further processed. This allows for a frequency-based analysis to be carried out. The interaction analysis will also focus on frequencies of (short) sequences of operations, which can be classified as having a quantitative or qualitative nature. In the empirical studies that are the topic of the following three chapters, the methods that have been described in this chapter will be applied in the context of the Optics computer simulation.

5

Optics 97: The influence of domain-specific and generic knowledge on the learning process in two discovery learning environments

5.1 Introduction

This chapter reports on the first study that was carried out in the context of this research project.[1] In this study, the discovery learning process is studied in the context of two computer simulations. The goals of the present study were stated in Chapter 4. Overall, the goal is to get insight into the role of prior domain-specific and generic knowledge on discovery learning. To this end, students worked for some time with two computer simulation learning environments, one of which was the Optics learning environment (as described in the Chapter 3), and the other a simulation on a fictitious subject.

The previous chapters have respectively dealt with the use of computers in education, the dual-space theory of discovery learning, a description of the Optics computer simulation, and an overview of measurements that are used in the context of this research project. The current study presents an attempt to gather information on the discovery learning process in the context of the Optics computer simulation. In this section the purpose of the study and expectations about the outcomes are discussed. Following that, in the method section the experimental context of this study is described in detail. A complete description of the Bubbles computer simulation that was used in conjunction with the Optics simulation in this study is also given. Student interaction processes are reported on globally as well as in more detail, and discovery learning processes of groups of poor and high perofrmance on the prior knowledge test measures are compared. Finally, the findings of the study are discussed. In Chapter 8, findings from this study are combined with those of the other empirical studies.

[1] The experiment that is described in this chapter was earlier reported in Hulshof, De Jong, and Van Joolingen (1998).

5.1.1 Purpose of the current study

One goal of the discussion in Chapter 2 was to make clear the influence of both domain-specific and generic knowledge on scientific discovery learning processes. As we have shown, it can be fruitful to describe this influence in terms of the SDDS (scientific discovery as dual-space search) framework that was developed by Klahr and Dunbar (1988). The purpose of the study we present here was to relate prior knowledge to interaction behavior in working with a discovery learning environment. For this purpose we used two different learning environments. The largest difference between these learning environments lies in their relation to the real world. As was discussed in Chapter 3, the Optics computer simulation utilizes the real-world topic of geometrical optics. The main consideration for designing the Optics discovery learning environment was that students were expected to vary in how much they would know (or remember) about the subject of geometrical optics. To test this expectation, prior domain-specific knowledge about optics was *measured* in the experiment. The other learning environment that was used, Bubbles, is different in this respect, because its domain is fictitious (that is, the variables and relations between them have no bearing in the real world). In this sense, this learning environment resembles the Detonlab environment that was used in a number of studies by Veenman (1993, also see Veenman, Elshout, & Hoeks, 1993, and Veenman, Elshout, & Meijer, 1997). Because the underlying rules of the simulation did not represent a real-world domain, students could not have prior domain-specific knowledge about this environment. This implied that, by way of instruction, the amount of domain-specific prior knowledge subjects would have when they started working with the learning environment could be controlled to some extent. To manipulate domain-specific prior knowledge, two different instructions sets were created, a normal instruction and an 'enriched' instruction. The latter contained some extra domain-specific information about the underlying rules.

All interaction behavior with the simulation, including the content of notes made during a learning session, was recorded, and written to a disk drive. The complete procedure has been explained in detail in Chapter 4. It was expected that analysis of the the learner-computer interaction behavior in the two learning environments would lead to insight in the influence of prior domain-specific and generic knowledge on discovery learning behavior. In the study, prior knowledge test measures are used to divide students into groups of poor and high knowledge. In addition, comparisons can be made based on performance in the learning environments, as measured by posttests. Similar to Shute and Glaser's (1990) experiment with Smithtown, it was expected that the latter type of analysis would provide information on

the learning processes that lie behind good versus poor performance on these tests.

5.1.2 Research questions and expectations

We can distinguish between two types of expectations that we hold with respect to the present experiment. First, it is important to note that we do not expect students to improve very much in their domain-specific knowledge from pretest to posttest in the Optics environment, because the process of discovery learning in the simulation proceeds mostly unguided. In this study, discovery learning is only supported by the use of (general) assignments, the structuring of the learning environment (as explained in the next section), and the ability to use an online notebook. We do expect any improvement that takes place to be partly dependent on generic knowledge, as well as prior domain-specific knowledge. Furthermore, we expect performance in the Bubbles environment to be correlated with improvement in Optics, independently of prior knowledge about optics, and to performance on the generic knowledge test. We also expect a difference in performance in the Bubbles environment between the normal instruction group and the group that received an enriched instruction, with the latter performing better on a knowledge posttest.

Second, we expect learners to differ from each other in their learning behavior. These differences should be related to the amount of domain-specific and generic knowledge they possess. With respect to the dual-space search theory we expect different experiment-search and hypothesis-search sequences. A distinction can be made between four different groups: 1) high domain-specific knowledge and high generic knowledge, 2) poor domain-specific knowledge and high generic knowledge, 3) high domain-specific knowledge and poor generic knowledge, and 4) poor domain-specific knowledge and poor generic knowledge. We expect subjects in the first group to display hypothesis oriented discovery behavior. They will start with searching Hypothesis space and search Experiment space only to test their hypotheses. Subjects in the second group will start with searching Experiment space, but after this initial period will show Hypothesis space search and systematic behavior. Subjects in the third group will show a constant switching between Hypothesis space and Experiment space, since they have no knowledge which can help them to find the right relation between variables. Subjects in the fourth group necessarily start with searching Ex-

periment space, since they do not know the important variables involved. Since they lack generic knowledge, they will have difficulty switching to Hypothesis space, resulting in unsystematic experimentation behavior.

For the Bubbles environment we expect the group that receives a normal instruction to differ from the group that receives an enriched instruction. The latter will have more domain-specific knowledge, and we expect learners in this group to do better experiments and to perform better on the post-test.

5.2 Method

5.2.1 Experimental setup

In the present experiment, secondary school students worked with two computer discovery learning environments. The learning environments that were used are the Optics simulation, which was explained in detail in the previous chapter, and another learning environment, called 'Bubbles'. The latter environment, which was only used in this study, will be explained in detail in this section.

The types of knowledge that were measured or manipulated in this study are depicted in Table 5-1. This design allowed a direct assessment of the influence of domain-specific and generic knowledge. As can be seen from the table, generic knowledge was only measured, but domain-specific knowledge was measured (for the Optics environment) as well as manipulated (for the Bubbles environment).

Table 5-1. *Setup of the Optics 97 experiment*

	Optics	**Bubbles**
domain-specific knowledge manipulated	-	√
domain-specific knowledge measured	√	-
generic knowledge manipulated	-	-
generic knowledge measured	√	√

Generic knowledge was measured prior to working with the learning environments, by a test that assessed generic knowledge. Domain-specific knowledge about geometrical optics was also measured prior to working with the environments. The generic knowledge test was a paper-and-pencil test that followed a multiple-choice format. The optics domain-specific knowledge test followed a what-if format, which was explained in detail in

Chapter 2. Answers on the test, as well as time to answer (in seconds) were recorded for this test.

Domain-specific prior knowledge on the Bubbles environment was manipulated by providing about half the subjects with an enriched instruction. Performance in both learning environments was measured afterwards with knowledge posttests.

5.2.2 Simulation domains

This section is devoted to a description of the computer simulations that were used in this study. An attempt was made to make the look and feel of both as similar as possible. Of course, because of their very different subject matters, there were differences between the environments in some of the available controls; where possible, however, the same icons were used to indicate similar actions in both simulations (for example, for deleting an object). The Optics computer simulation has been described in detail in Chapter 3. Therefore, in this section only the structure of the simulation as it applies to the current experiment (that is, the options that were available to students) is described. The Bubbles simulation is given a full description in the present section.

5.2.2.1 Optics simulation

Because of the complexity of the underlying formulas, it was decided to structure the learning process by dividing the computer simulation into three parts. It was expected that this would reduce the complexity of the simulation, especially in the beginning. The three 'phases' into which the simulation was divided closely resembled each other; the main differences could be found in the objects that were available to students. For each of the three phases in the environment a general goal was stated. There was a minimum time limit set for each phase, but once a student entered a new phase, it was possible at any time to return one of the previous phases. Of course, the model underlying the environment, that is, the rules that specify the lens refraction rules, stayed the same over phases. Only the assignment and the number of objects that could be used changed in each phase. The design of the three phases was as follows.

first phase. The goal of the first phase was to investigate what happens when the image of a plate with holes in it is projected on a screen. A big lamp was

always present in the environment in this phase, its light turned on. When a plate with holes in it, a lens, and a screen would be placed in the environment from left to right, an image of the plate would be projected on the screen. Initially, this picture would probably look out of focus, but this could be changed by repostioning the different objects. The size and sharpness of the depicted image is dependent on the distance of the plate to the lens, as well as the distance of the screen to the lens. In this phase the properties of three different lenses, two positive (with differing focal distances) and one negative, could be examined. It was not possible to place more than one lens in the working area. To aid the investigation, it was possible to add a lamp with three divergent light beams to the working area.

second phase. The second phase can be seen as an introduction to the third phase. In this phase, the big lamp, plate, and screen were not available. Instead, adding more than one lens was allowed. The same three lenses as in the first phase were featured in this phase, plus an additional negative lens (with a different focal distance). Only the lamp with three divergent light beams was available for use.

third phase. In the third phase, the functionality of the first and second phase were combined. Not only was addition of multiple lenses allowed, the big lamp, plate and screen were also available for use. In this phase it was possible to project the image on the screen through multiple lenses. The lenses in this phase were the same as the ones from the second phase. The lamp with divergent light beams could be utilized to track the light going through more than one lens.

5.2.2.2 Bubbles simulation

The topic of the Bubbles computer simulation is a special fictitious chemical reaction that takes place under certain conditions, when two or more liquid materials are put together and the resulting mix is heated. The names of liquids, as well as the rules underlying the chemical reaction, are artificial. A description of the underlying model of the Bubbles simulation is given in Appendix E of this thesis. To motivate students to work with this artificial environment a science-fiction-like background story was presented as part of the general instruction. The story involved the exploration of a newly-discovered planet. The planet seems to resemble earth in every way, except for the presence of four unknown liquids. The names that were given to these liquids were inspired by the names given by Veenman (1993) to mate-

rials in a similar domain: Magnum, Kryton, Sybar, and Guernic.[2] When certain amounts of them are mixed together and the resulting mix is heated a chemical reaction starts in which 'bubbles' originate. The number of bubbles changes over time. Students are put in the position of the exploring team's scientist who wants to find out the properties of this reaction. This involves doing experiments in a laboratory in which the reaction between different mixtures of the liquids can be observed. It's not only important to study whether or not bubbles appear, the amount of bubbles appearing per second also changes every second, dependent on the amount of liquids added to the mix. The trajectory the reaction follows (that is, the number of bubbles that originates from the reaction each second) is plotted in a graph (as number of bubbles appearing every second; time in seconds is shown on the X-axis, the number of bubbles that appear on the Y-axis). However, students have to specify themselves the points in time where they want to make measurents in the graph. The maximum number of measurement points is four, which means that the amount of bubbles appearing over time is plotted four times in the graph as a maximum. An example of the Bubbles interface is given in Figure 5-1. In the figure, the first two measure points (which have been placed at times 2 and 6) show the number of bubbles that appears each second decreases in this situation. Also, the amount of bubbles that will appear at times 11 and 15 cannot yet be known from the graph.

[2] These names were used in the Detonlab environment. The Bubbles simulation can be seen as an extended version of Detonlab, with the Bubbles simulation being more complex because of a larger number of variables and underlying rules.

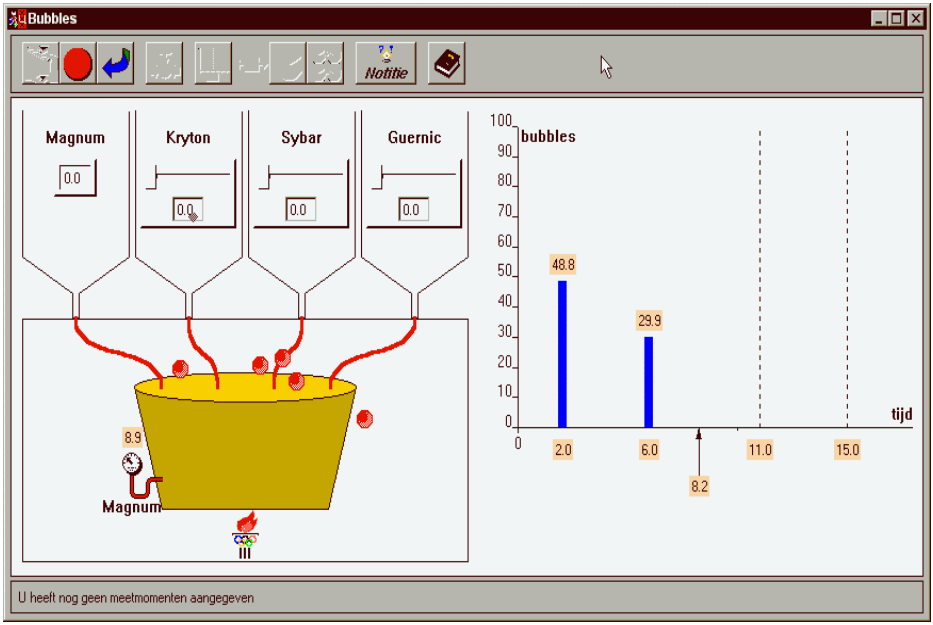


Figure 5-1. *Example interface of the Bubbles learning environment. Explanation given in text*

Because the Bubbles environment explores a fictitious domain, all subjects were given some instructions about the environment prior to working with it. The instructions could also be accessed at any time during working with the environment. Two different instruction sets were developed to manipulate the amount of domain-specific information subjects received: a normal (minimal) instruction and an enriched instruction. The normal instruction consisted of the following parts:

1) Background story
2) Manual (overview of the simulation)
3) Interface instructions (specific operation of buttons)
4) Assignment (statement of overall goal)

The background story was used to help students make sense of the artificial nature of the environment, and of the operations and objects that played a role in it. The manual gave an overview of the learning environment itself, and explained the objects in the simulation. Interface instructions gave an overview of all the possible actions in the environment; a complete list of all possible operations is shown in appendix D. Finally, subjects were given an

assignment. The assignment stated that the goal was to find out how the amount of all four materials determined the amount of bubbles appearing over time.

The enriched instruction set consisted of the same interface instructions and assignment. The background story and manual were different from the normal instruction version in a number of aspects. The background story contained extra information about the necessity of the presence of Magnum and how it is used in the environment (its amount is held constant and cannot be changed). The manual contained extra information about the (fictitious) units used, the minimum and maximum amounts of material that could be used, the way an experiment worked, and a variable 'effect' was introduced to denote the number of bubbles appearing every second. In addition to these changes to part of the instruction, the enriched instruction included an extra section in which a number of hints were given, and students were encouraged to study it carefully. It contained more detailed information about how the materials interacted together to form a reaction, without giving away precise relations. For example, a hint was given about the fact that when the amount of one material was set at a higher level than another, the reaction could altogether stop. Furthermore, it was explained that a material could quicken or delay the onset of the actual reaction. Also included was information on what sort of chemical reaction trajectories could possibly occur.

All in all, precise reading of the extra information in the enriched instruction was expected to give students a head start in the amount of domain-specific knowledge about the learning environment.

5.2.3 Subjects

Subjects in the present study were 51 students who were in secondary school (mean age about 17 years), at the level of pre-scientific education. They came from two school classes of 26 and 25 persons respectively. All subjects were taking physics lessons at school. In the year before the experiment, they had voluntarily chosen physics to become part of their compulsory curriculum. The subject of geometrical optics had been dealt with in these lessons a few years ago, so it could be expected that there would be differences in the amount of domain-specific prior knowledge. All subjects were reasonably familiar with the type of computer interface that was used (e.g., using the

mouse to move the pointer around, select icons, and perform operations in the environment). No special instruction on working with the computer was necessary.

5.2.4 Tests

Subjects were pretested before starting to work with the learning environments. Tests for generic knowledge and domain-specific knowledge were administered. The tests each took about 20 minutes to complete. Both tests were administered in the first experimental session, which means that they were administered about a week and a half in advance of working with the first learning environment.

5.2.4.1 Pretests

The generic knowledge test, which tested for knowledge about mathematical relations, was a paper-and-pencil test. In Chapter 4 (Section 4.2.1), the theoretical background to the design of the test items was discussed. The test that was designed for the present experiment consisted of sixteen multiple-choice items (in the test, some items had three answer possibilities, and some had four). Answers were scored dichotomously. Items for the test were developed by the researchers.

The domain-specific knowledge test consisted of eighteen items. The test was a multiple-choice test that was adminstered by the computer. To test for intuitive knowledge, the test followed a 'what-if' format, which has been discussed in Chapter 4 (Section 4.2.2).

5.2.4.2 Domain-specific posttests

The Bubbles posttest was administered immediately at the end of the second experimental session. The test consisted of twenty items in a what-if format. The items consisted of questions on various levels (from questions about simple effects to questions about interactions). Like the domain-specific knowledge test for the Optics simulation, test items were administered in a fixed order. The Bubbles pretest did not use a domain-specific knowledge pretest, because the underlying rules of the Bubbles environment could not be known to subjects; testing subjects about their knowledge of a set of artificial relations would also have lowered their motivation to work with the learning environment, and would have interfered with the domain-specific knowledge that part of the group received.

The Optics domain-specific posttest was administered immediately after subjects had completed their second session with the Optics simulation. No domain-specific knowledge test was administered after the first session with the Optics simulation. The test that was used was the same test as the Optics pretest, but this fact was not apparent to subjects because both tests were administered in random order.

5.2.5 Registration of behavior

As was discussed in Chapter 4, all operations that were performed by subjects while they worked with the two computer simulations were registered electronically. This means that all button presses were recorded together with the offset time in seconds (a timer started as soon as the learning environment was entered). When an object was moved around, the Optics simulation was updated in real-time. In the Bubbles simulation, the environment was updated as soon as a subject pressed the 'start experiment' button. The difference between the Optics and Bubbles simulation in the way the environment was updated is further discussed in the final section of this chapter.

5.2.6 Procedure

Table 5-2 outlines the sequence of events that was followed throughout the experiment. The subjects who participated in the experiment came from two different school classes.

Table 5-2. *Experimental procedure*

Session	Content	Time taken
1	Generic knowledge test, Optics pretest	50 minutes
2	Bubbles environment and Bubbles post-test	90 minutes
3	Optics environment part 1	50 minutes
4	Optics environment part 2 and Optics post-test	90 minutes

For each class the experiment was spread over four sessions. In Table 5-2, the time schedule and sequence of these sessions is shown for both school classes. The classes performed the experiment in the same sequence, but not at the same time (the reason for this is that the number of available computers was restricted). The first experimental session was devoted to administering the generic knowledge test and the domain-specific knowledge pre-

test, the second session to the Bubbles environment and administering the Bubbles posttest, and the third and fourth sessions were spent on working with the Optics simulation. Also, after the fourth session an Optics domain-specific knowledge posttest was administered. Because the Optics simulation had a richer content than the Bubbles environment (e.g., the three phases used different configurations of the same simulation), it was decided to allow subjects to work for a longer time with this environment, and to arrange the sequence of working with the learning environments so that the Bubbles environment would be worked with first, and the Optics environment second. Because the experiment took place during normal school hours (as a replacement for physics lessons), session times were restricted to the standard length of these lessons, about 50 minutes for one school hour and approximately 90 minutes for two consecutive lessons.

As was mentioned before, both school classes performed the experimental sessions in the same order. Subjects first completed the generic knowledge test and the Optics pretest (the first one on paper, the second one on the computer). Next (a few days later) a session with the Bubbles environment followed, after which a knowledge test was administered. Some time after this (approximately one week), a short session was spent on getting used to the Optics environment. The last session involved working with the Optics environment; at the end of which a post test was administered. There was no difference between the Optics environment of the third and the fourth experimental session; at the beginning of the fourth session the state in which the program had been left at the end of the third session was resumed. This also meant that each subject's notes were saved over sessions.

Before they began working with the computer simulations, subjects were provided with a short explanation of the interface, and were instructed on some of the operations that were possible in the simulations. During the experiment, a fact sheet with a short explanation about the different icons was available to subjects. Also, when needed the experimenter would give further instruction about the operations; however, no guidelines were given about 'good' discover learning behavior or about the underlying rules of the simulations.

5.3 Results

In this section, results of the current experiment are described. Two types of results are given. First, results of the different knowledge tests that were administered are described. Next, these test results are compared to each other. Second, we describe an analysis of students' interaction behavior at

different levels. Test data and data on the interaction with the Bubbles and Optics simulations are then combined, to gain insight in the similarities and differences of the behavior, as based on differences in test results.

In total, 33 male and 18 female subjects participated in the experiment. No gender differences in proficiency at operating the mouse and keyboard of the computer were observed.

5.3.1 Test results

Three different tests were used: a generic knowledge test, a domain-specific knowledge test about geometrical optics (both a pre- and a posttest), and a domain-specific knowledge test about Bubbles (only a posttest). The generic knowledge test used a standard multiple-choice format, the other tests were intuitive knowledge ('what-if') tests that used the format described in Chapter 4. At present, there is no formal method of analysis of intuitive test results, that allows for the combination of test scores and answer times into one generalized score. Therefore, for these tests we analyzed scores and times separately, and computed correlations between them.

5.3.1.1 Generic knowledge

Because of ambiguity, one item was discarded from the test before analysis. The overall average score on the generic knowledge test was 11.8 out of 15 multiple choice items, with a standard deviation of 1.7. The results show a ceiling effect, with only a few of the subjects scoring less than 10 out of 15 questions. Analysis shows that in order to study where this ceiling effect derives from, it is helpful to divide the items used in the test in two sets: one set of 10 items was answered correctly by more than 75 percent of the subjects. The mean number of correctly answered items for this set is 9.35 (SD is 1.92). These items refer to a specific numerical relation (the question being: what number should logically follow these?), and most of the subjects did not experience problems in answering these questions correctly. The five items in the other set are answered correctly by less than 75 percent of the subjects. The mean number of correctly answered items for this set of five is 2.27 (SD is 1.03). In these items a more complex relation is used (the Fibonacci sequence), graphs were used, and the items related to experimental behavior on the basis of results that were shown. What this implies is that

subjects had almost no trouble at all in dealing with simple numerical problems (that could be solved by using superficial strategies) and only experienced difficulty in understanding more complex relations. Further analysis will take the results from the complete test (15 items) into account. Because of the variation in contents of the test, internal consistency is quite low. Removal of one item results in a Cronbach's alpha of 0.40.

5.3.1.2 Domain-specific knowledge

Not all subjects took part in both the pretest and the posttest, because of some students dropping out of the experiment after one of the experimental sessions. All subjects completed the domain-specific knowledge pretest. The results of two subjects were discarded because their answering times were too short to be reliable knowledge measures. This resulted in 49 subjects completing the domain-specific knowledge pretest (26 in class 1, 23 in class 2) and 44 completing the posttest (24 in class 1, 20 in class 2). Both pre- and posttest used eighteen multiple-choice questions (following the what-if format).

Cronbach's alpha for the pretest turned out to be very low (removal of three items of the pretest gives an alpha of 0.29). Cronbach's alpha for the posttest is, like the pretest, low. Removal of 4 items gives an alpha of 0.40. In the discussion section we will go further into the issue of the relevance of alpha values in the case of an intuitive knowledge test. Overall test scores and answer times were recorded, and averages are shown in Table 5-3.

Table 5-3. *Overall mean results of domain-specific knowledge tests (18 items)*

	Average	Standard Deviation	Range
Pretest score	6.82	2.1	1-11
Posttest score	7.84	2.2	4-13
Pretest time (seconds)	672.49	175.2	47-1018
Posttest time (seconds)	427.61	175.1	105-920

As can be seen from the table, mean pretest scores did not differ from chance level, which indicates that prior knowledge was very poor (even the highest score on the pretest was still low). The average number of correctly answered items increased significantly from pretest to posttest ($p<.05$), although by a small margin. In addition, average total time to answer all questions decreased significantly ($p<.05$). There was a large variation in the answer times, with some subjects answering extremely quickly on the whole. We cannot exclude these subjects from our analysis, because it may be that these subjects 'saw' the answer almost immediately as soon as the situation

was displayed on the screen, as a result of intuitive knowledge. Mean answer time per question was 37 seconds (SD=5.4) on the pretest, and 24 seconds (SD=6.1) on the posttest.

Table 5-4. *Comparison of mean optics results of the two participating school classes*

	Class 1	**Class 2**	***p*-value**
Pretest score	7.03 (1.75)	6.58 (2.41)	0.45
Posttest score	8.17 (2.33)	7.43 (2.16)	0.28
Pretest time (seconds)	672.12 (108.02)	646.83 (260.20)	0.65
Posttest time (seconds)	459.96 (193.27)	388.80 (145.83)	0.18

Table 5-4 shows a comparison between the two different school classes that participated in the experiment. As can be seen from the table, there was neither a significant difference on the testscores, nor on the answer times between the two school classes, both on the pre- and the posttests.

There are two ways of computing the correlation between the number of correct items on the pretest and posttest. When computed within-subject and across items no significant correlation between the scores on the pretest and the posttest was found. Regression analysis showed an R square of 0.06, which indicates that almost no variance between scores on the posttest is accounted for by the pretest. This means that performance on the posttest can not be explained by performance on the pretest. We can therefore ascribe changes in performance from pre- to posttest to working with the learning environment. When computed across subjects and within items this correlation was 0.75 (p<0.01). This high correlation means that subjects tended to answer the same questions correctly in the pretest and the posttest, and confirms that overall progress from pre- to posttest was not very high. There also was a positive correlation of 0.32 between the answer times on the pretest and those on the posttest (p<.05) when computed within subjects, which indicates that there is some consistency in an individual's speed of answering.

To elaborate on this analysis, we computed the number of items that were answered correctly in the pretest but incorrectly in the posttest, and the number of items that were answered incorrectly in the pretest but correctly in the posttest. The first measure is, in a sense, an indication of detoration from pre- to posttest, the second a measure of improvement. The mean on the first measure was 2.7 (SD=1.38), and on the second it was 4.1 (SD=2.11).

The results with respect to the decrease in answer times can be criticized on the grounds that we considered the answer times regardless of whether an item had been correctly answered or not. The idea is that knowledge can only be used when it is present.

When we assume that a wrong answer on a question indicates that the knowledge needed to answer correctly was not available, it means that the characteristics of intuitive knowledge are not relevant in that case. This means that, under this assumption, only the answer times for correct answers are relevant.

Table 5-5. *Domain-specific knowledge tests mean answer times*

	Average	**Standard Deviation**	**p-value**
Pretest time, correct answers	258.44	118.0	
Posttest time, correct answers	184.71	114.4	$p<0.01$
Pretest time, wrong answers	419.10	136.1	
Posttest time, wrong answers	241.57	121.6	$p<0.01$

In contrast, it can also be claimed that a wrong answer is not an indication of the absence of knowledge, but of the presence of wrong knowledge. In that case the answer pattern for wrong answers may well be the same as for correct answers. To study this, we divided the test results into correctly and wrongly answered items: the mean answer times are displayed in Table 5-5.

After correction for correct and incorrect answers, there is still a significant overall decline in answer time from pretest to posttest, for both correct and incorrect answers. In addition, there are significant differences between the correct answer times and the incorrect answer times, for the pretest as well as the posttest ($p<.01$). In Table 5-6 the trade-off between the scores on pre- and posttest and the answer times for both incorrect and correct answers is shown. Values that are significant below the 0.05 level are shown in bold type.

Table 5-6. *Within-subject correlations between subject test scores and answer times on the domain-specific knowledge test*

	Posttest score	**Pretest time correct items**	**Pretest time incorrect items**	**Posttest time correct items**	**Posttest time incorrect items**
Pretest score	0.25	**.30**	-.05	**.48**	-.09
Posttest score		.04	-.20	.14	-.04
Pretest time correct items			-.02	**.39**	**.29**
Pretest time incorrect items				.14	.04
Posttest time correct items					.09

Table 5-6 shows that there are a number of significant correlations between the correct answer times and test scores, but no correlations for the incorrect answer times. As can be seen there are moderate correlations between the pretest scores and the answer times for correct items, both on the pretest and the posttest. The posttest scores do not correlate with the answer times. A positive correlation means that the longer a subject takes to answer a question, the higher the chance is that the given answer is correct. Because the intuitive knowledge test taps 'fast' knowledge, no positive correlation is expected. The result that this relation is found for the pretest but not for the posttest gives some indication that the knowledge that was used to answer the pretest items did not have an intuitive quality, and that for the post test it had.

5.3.1.3 Bubbles knowledge

The Bubbles test was administered to 49 subjects. The format of this test was the same as that for the Optics tests, with the exception that the Bubbles test had 20 items instead of 18. The overall average score on the Bubbles test was 11.97 with a standard deviation of 3.14. To test whether subjects learned something from working with the Bubbles environment, we compared their scores on the knowledge test with a set of random data (that simulated the responses of persons who had chosen a random answer for every question). This analysis showed a significant effect of working with the learning environment (p<.01). The mean overall answer time for the test was 402.64 seconds with a standard deviation of 138.11. Cronbach's alpha for the Bubbles test was 0.62, which is moderate.

We compared the scores of the enriched instruction group with those of the group that did not receive extra instructions. The results are shown in Table 5-7.

Table 5-7. *Bubbles test scores and answer times (20 items)*

	Normal instruction (n=21)	**Enriched instruction (n=28)**	**p-value**
Test score	12.4 (3.3)	11.9 (2.8)	0.61
Answer time	385.6 (142.8)	417.4 (149.8)	0.47

As can be seen from the right column, no significant difference between the two groups was found on the knowledge posttest, neither in the scores nor in the answer times. It appears that both groups learned equally well from working with the learning environment.

Like the Optics pre- and posttest, we made a division for the Bubbles test in items that had been correctly answered, and items that had been incorrectly answered. The results show a contrast with the results for the Optics tests. For the correctly answered items the mean answer time was 232 seconds (SD=114.7) and for the incorrectly answered items 157 seconds (SD=88); this difference is statistically significant (p<.01). What this means is that subjects in general take longer to answer questions correctly than incorrectly in the Bubbles test. An explanation for this finding might be that in the Bubbles test it is more difficult than in the Optics tests for subjects to see which answer to a question is correct.

5.3.2 Prior knowledge test effects

We expected improvement from pretest to posttest in the Optics environment to be related to generic and domain-specific prior knowledge. In the previous section it has been shown that there is no significant correlation between scores on the Optics pretest and scores on the posttest. Also, no significant correlation was found between the generic knowledge scores and improvement from pretest to posttest in the Optics environment. An explanation might be that the Optics environment encourages subjects to experiment in a qualitative way: no stress is laid on applying general mathematical skills. In the section on interaction behavior we will discuss this fact further.

We expected performance on the Bubbles test to be correlated with improvement from pretest to posttest in the Optics environment, independently of domain-specific knowledge about optics. A correlation analysis, with prior optics knowledge partialled out, shows no significant correlation however (r=-.15; p=.34). The normal instruction condition and enriched instruction condition do not differ in this respect.

We also expected performance on the Bubbles test to be related to performance on the generic knowledge test. Indeed, in the Bubbles normal instruction condition, there is a significant correlation of 0.39 between the generic knowledge test and the Bubbles test (p<.05). The correlation is not significant for the enriched instruction condition. We can interpret this result, in combination with the equivalent results for both conditions on the Bubbles test, as follows. The scores on the Bubbles test for the normal instruction condition can be partially explained by general mathematics skills: the more skilled a subject is, the better the performance in the learning environment. In the en-

riched instruction condition subjects have some of the more difficult work in coming up with good ideas about the environment already done for them. Because of this it might be the case that these subjects do not rely as much on their generic knowledge about relations as subjects in the normal instruction. A regression analysis gives limited support to this claim. Removal of two outliers (persons who scored extremely low on the generic knowledge test compared to the other subjects) resulted in a correlation between the mathematics test and the Bubbles test of 0.63 for the normal instruction condition. Regression analysis showed that R square was 0.40, which is statistically significant (p<.05).

In summary, we only found limited support for our expectations with respect to the influence of domain-specific and generic knowledge on learning performance in the Bubbles and Optics learning environments. By working with the Optics and the Bubbles computer simulations, students picked up at least some of the properties of these learning environments. This is shown both in the scores on a posttest and answer times. However, in the Optics environment, we found that improvement in performance as shown on the posttest appears to be unrelated to pretest performance. The results from the Bubbles environment suggest that generic knowledge can influence performance in a learning environment that calls for these skills.

5.3.3 Interaction behavior results

In this and the following section, we analyze the characteristics of the interactions subjects performed in both learning environments. We look at both general measures of interaction behavior (e.g., frequencies of operations) and more specific measures (e.g., changes over time in the number of performed operations). An attempt is made at shedding light to some of the questions that were raised earlier. We do this by comparing different groups of learners, for example learners with poor domain-specific prior knowledge versus learners with high domain-specific prior knowledge. For both learning environments, we follow a descriptive approach. The analysis is performed at different levels of description. At a global level, we look at the overall characteristics by which we may define learning behavior as it occurred in the learning environment. For example, how many actions were performed, and what actions were performed more than others? At a more detailed level, we look at the specific path that subjects followed through the

learning environment. For example, what combinations of actions can we distinguish, and did the focus of attention change over time? In between these two levels, we look at the way variables in the environment are manipulated, and the systematicity of working. Also, we look at the kind of situations subjects created in the environment: was every possibility explored, and how much sense do situations make (or: how much information can be obtained by a subject from a situation)? In the Optics environment we focus on the role that different objects such as lenses and light sources play, and the operations that are performed on them. In the Bubbles environment we focus on the type of experiments that are performed.

First, we describe the contrasting groups that we created on the basis of the various test results. Group categorization is based on the results on the different knowledge tests.

5.3.3.1 Classification of subjects

Because of various technical reasons, we could not obtain interaction data for all subjects. For Optics we obtained 48 logfiles and for Bubbles 49. Only these subjects are considered in the analysis.

Subjects were categorized on three different measures: score on the optics pretest, score on the generic knowledge test, and score on the Bubbles test. Because test score differences were quite small, we divided subjects only over two extreme categories. In the case of the interaction with the Bubbles environment we also made use of the manipulation in domain-specific prior knowledge.

To create two groups on the basis of the Optics pretest, we computed a score one standard deviation below, and a score one standard deviation above the mean. The resulting numbers, 4.776 and 8.9, were rounded to the nearest integer, resulting in a minimum score of 5 and a maximum of 9. Selecting only from the subjects for whom we had obtained Optics logfiles, we formed a low scoring group (n=14, score 5 points or less on the pretest), and a high scoring group (n=10, score 9 points or higher). Because of the low variation in test scores, poor and high scorers do not lie far apart. For the generic knowledge test, because the results were subject to a ceiling effect, we decided to distinguish between groups on the basis of the median (the median is 12 out of 15 items correct). We selected subjects from the group for which Optics logfiles were available, and left out subjects who scored exactly 12 (n=13). By doing this we created a 'low-scoring' group (n=18) and a 'high-scoring' group (n=17). Like for the Optics pretest, we created groups one standard deviation below and above the mean on the Bubbles test. The selec-

tion was made from those subjects for whom a Bubbles logfile was available. Doing this resulted in a low-performing group (n=7) and a high-performing group (n=13). An interesting finding when we compare test patterns for the poor and high scoring group is the correlation between the overall test scores and the answer times. The overall correlation is 0.22, which is not significant. However, when we separate the low scoring group and the high scoring group results in correlations of respectively 0.43 (p<0.05) and –0.13 (ns.; p=.55). What this shows is that the longer low scoring subjects take to answer a question, the higher the chance that their answer is correct. As has been argued in the case of the Optics pretest, this may imply that these subjects did not use intuitive knowledge to answer items, instead relying on the explicit concepts they had mastered.

Because the environments differ in the type of interactions that can be performed, we will analyze them separately. We will first discuss behavior in the Optics environment, then behavior in the Bubbles environment.

5.3.3.2 Interaction with the Optics simulation

A large number of individual operations are possible in the Optics environment (all the actions and objects that were available for use in the environment are explained in detail in Appendix C). Most of these operations are performed on one or more objects available. Objects include concrete objects like lamps and lenses, and tools used for measurement like construction lines and distance lines. Our analysis focuses on the way objects are used in the learning environment. We distinguish in our analysis between global measurements and more specific measurements of the use of objects. Global measurements concern differences in frequencies with which operations have been carried out. Specific measurements concern the actual use to which objects are put: in what way can operations be meaningfully carried out. A general problem in the analysis of behavior in the Optics environment is that the environment has a continuous nature, in which experiments are not separated from each other in a discrete way. Rather, if we want to say something about the type of experiments performed by subjects we have to define categories in which to put operations or series of operations that can be called experimental.

From the perspective of object use, we can divide the operations that can be carried out in general in the Optics environment into three types: 1) core ac-

tions (type I), 2) peripheral actions (type II), and 3) irrelevant actions (type III). Core actions are those actions that are necessary and sufficient for working with the learning environment. They are necessary in the sense that without them, the other operations make no sense, they are sufficient in the sense that in principle, no more operations are needed to work meaningfully with the learning environment. These actions are made up of adding and deleting objects, moving them around, and rotating a lightbeam. Moving objects around and rotating a lightbeam are here considered separately, because the kind of information they yield to the learner is different, and because light beam rotation can only be carried out on one object: a lamp. Type II actions are actions that make the learning environment yield more information than can be gathered at first sight and that enable learners to keep track of results. These actions are therefore helpful but not necessary to work with the learning environment. They include making notes, adding measures, viewing instructions, and naming objects. Type III actions have no special meaning, for instance just pressing a button[3], or opening and closing an experimental session. Type III actions are omitted from further analyses, the other two types are discussed separately.

In total, subjects carried out a mean number of 436 (SD=150) type I and type II actions. Type I actions can be carried out on nine different objects, most common among which are lamps and lenses. In Table 5-8, the occurrence of type I actions is compared for threeferent groups we created. In the table mean frequencies of occurrence per subject are shown, with standard deviations between parentheses. The right column shows the sum over operations of the means. Type I actions constitute a very large part of all possible actions in general, about 85%. Typical learner behavior consists of a long series of type I actions, interrupted at intervals by type II actions.

For each of the four categories we separately performed a Manova to check for overall differences between the groups. There is an overall statistically significant difference between the poor generic and high generic prior knowledge group ($p=.014$). A t-test shows that this difference can be attributed to the number of times objects are moved ($p<.05$). Subjects with high generic prior knowledge moved objects less around than subjects in the other group.

[3] This means, selecting an operation (by clicking its corresponding selection button), but instead of actually performing it subsequently selecting a different operation. This 'random clicking behavior' has been left out of the analysis.

Table 5-8. *Mean frequency of type I actions, for different groups in the Optics environment (standard deviations between parentheses)*

	Action				
Group	**Add object**	**Delete object**	**Move object**	**Rotate light-beam**	**Mean total**
Poor optics (*n*=14)	158 (66)	48 (25)	137 (83)	31 (19)	119 (59)
High optics (*n*=10)	193 (146)	55 (31)	121 (93)	23 (13)	98 (75)
Poor generic (*n*=18)	139 (65)	55 (23)	162 (103)	29 (20)	96 (64)
High generic (*n*=17)	201 (125)	41 (34)	96 (73)	24 (10)	91 (80)
Poor Bubbles (*n*=7)	136 (73)	46 (30)	89 (47)	40 (20)	78 (45)
High Bubbles (*n*=12)	161 (124)	60 (33)	147 (73)	21 (11)	97 (68)

In addition, we can state tentatively that the high generic prior knowledge group tended to add more objects (p=.07) than the poor generic prior knowledge group. These findings indicate that learners with high generic prior knowledge were in general more active in doing different experiments, adding an object, performing a few operations on it, and then deleting it again. There is also a statistically significant difference between the groups based on the Bubbles scores (p=.02). A t-test shows the difference can be attributed to the rotation of light beams (p<.05). Subjects who score low on Bubbles perform this operation less than learners who score high on Bubbles. In addition, subjects who score high on Bubbles tend to move objects more than subjects who score low on Bubbles (p=.08), which confirms the idea that moving objects and rotating a lightbeam are qualitatively different operations.

In Table 5-9, the occurrence of type II actions is compared for the three groups, in the same way as in Table 5-8. Again, mean frequencies of occurrence over subjects are shown, with standard deviations between parentheses. Comparing Table 5-9 with Table 5-8 shows not only that type 2 operations are performed a lot less than type 1 operations, but also that there is a large variation in the frequency with which type 2 operations are performed. Because of this large variation none of the groups differ significantly, although it appears that subjects in the high groups perform more type 2 operations than subjects in the low groups. 'Make measurement' in Table 5-9 is the sum of two different operations: both the use of the 'make measurement' button in the learning environment, and the addition of a distance measure. The first operation is needed as soon as an object is moved around (numbers present in the working area are replaced by question

present in the working area are replaced by question marks); the second operation is carried out more in general. In the section on the specific use of objects we will show that subjects did work in a quantitative way, but found a way to circumvent using measurement operations.

Table 5-9. *Mean frequency of type II actions for different groups in the Optics environment (standard deviations between parentheses)*

	Action				
Group	**Notebook entry**	**View instructions**	**Make measurement**	**Use naming tool**	**Total**
Poor optics (*n*=14)	10 (9)	11 (4)	14 (9)	40 (31)	21 (19)
High optics (*n*=10)	14 (15)	11 (2)	29 (18)	31 (30)	21 (10)
Poor generic (*n*=18)	13 (12)	11 (3)	18 (12)	30 (30)	18 (9)
High generic (*n*=17)	11 (9)	12 (5)	18 (12)	29 (28)	18 (8)
Poor Bubbles (*n*=7)	9 (6)	13 (3)	17 (11)	24 (16)	16 (6)
High Bubbles (*n*=12)	14 (14)	11 (4)	28 (17)	34 (31)	22 (11)

Movement of objects

To further look into the activities performed by learners during working with the learning environment, we computed the mean number of movements, that is, the number of movements made with objects during their existence in the working area. Moving objects around is informative because the working area is updated in real-time. Therefore, the effect of moving an object is immediate. If the number of movements was zero (that is, the object was introduced and deleted without any operation having been performed on it) it was excluded from this computation. Because often objects are only added and not moved around, we also counted the frequency with which objects were added to the environment but not moved at all (we call these 'zeromoves'). Both measures are displayed in Table 5-10 for all the different groups, with standard deviations between parentheses.

Table 5-10. *Mean number of movements per object in the Optics environment (standard deviations between parentheses)*

	Group					
	Poor optics	**High optics**	**Poor generic**	**High generic**	**Poor Bubbles**	**High Bubbles**
movements	4.4 (1.40)	4.3 (1.7)	4.9 (1.7)	4.1 (1.9)	3.7 (1.2)	6.1 (4.0)
'zeromoves'	22.4 (40.6)	10.7 (28.6)	23.7 (39.7)	26.7 (49.1)	9.8 (28.0)	13.6 (33.1)

When we compare the higher scoring groups with the lower scoring groups (using t-tests), we find that the Bubbles groups differ significantly in the number of movements ($p<.05$), where higher performance is associated with more movements per object. In addition, it appears that both groups are more extreme than the other groups, that is, the poor Bubbles group has a lower mean number of moves (3.73) per object than the other groups, and the high Bubbles group a higher mean number of moves (6.14). This shows that performance in the Bubbles environment is indeed related to the number of movements performed (as we concluded tentatively from Table 5-7). There are no significant differences in the number of 'zero-moves', because of large individual differences.

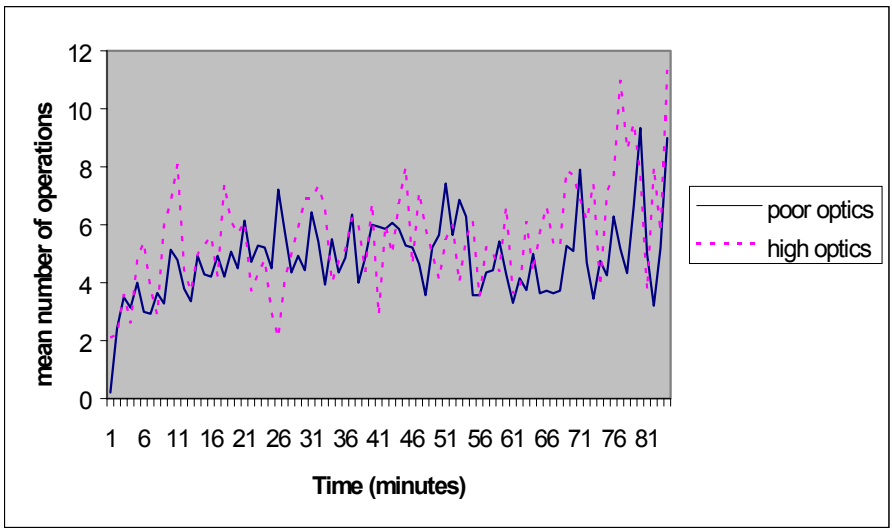


Figure 5-2. *Type I operations over time, for the prior Optics knowledge groups*

Effects over time

In addition to the frequencies of different operations (as given in the previous section), we analyzed changes over time for type I and type II operations. Figure 5-2 and Figure 5-3 depict changes over time in type I operations for the Optics prior knowledge groups and the generic knowledge groups, respectively.

The computations for these figures are done as follows. For each subject operations were categorized as being of type I or not, and the time (in seconds) was registered as well. The results for all subjects were added at each time point. To make the results more concise, frequency of actions was cumulated in 60 second steps.

Because subjects differed somewhat in the amount of time they had worked with the learning environment (a few subjects dropped out after the first session with Optics), we used a dynamic algorithm to compute the mean number of operations. At the moment subjects stop, they are left out of the computation of the mean. When only one subject is left, we stop the analysis. Therefore the graphs differ in total time in minutes, because they take different groups of subjects into account.

As can be seen from the figures, there do not appear to be group differences in the number of type I operations performed over time. In both figures, after an initial rise, the number of type I operations performed roughly stays the same to the end.

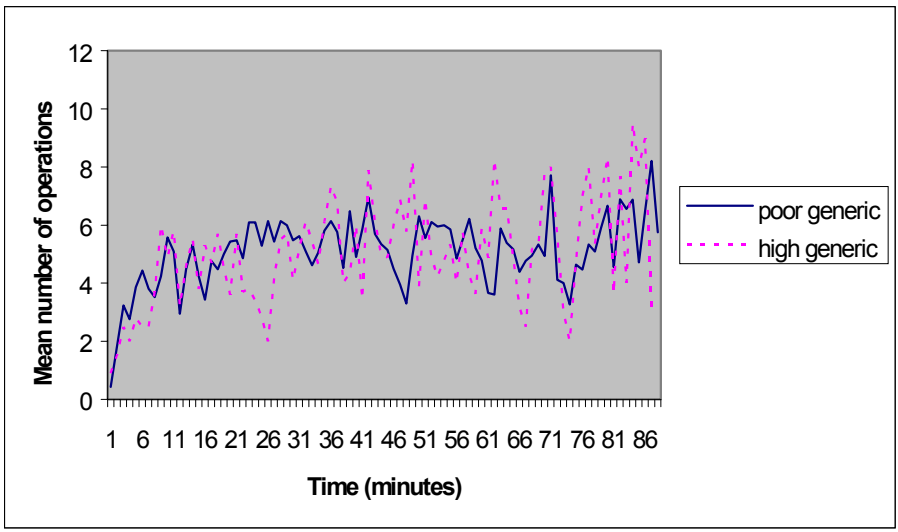


Figure 5-3. *Type I operations over time, for the prior generic knowledge groups*

Figure 5-2 and Figure 5-3 show that the overall lack of difference between the groups with respect to type I operations is reflected in the time course of these events. However, we did find a difference for the generic knowledge groups in the number of movements performed on objects. Because both additions and movements count as type I operations, these two operations

may have cancelled each other out, explaining for the lack of any finding. To check this idea, we computed the mean number of operations over time for movements only. The resulting picture is displayed in Figure 5-4.

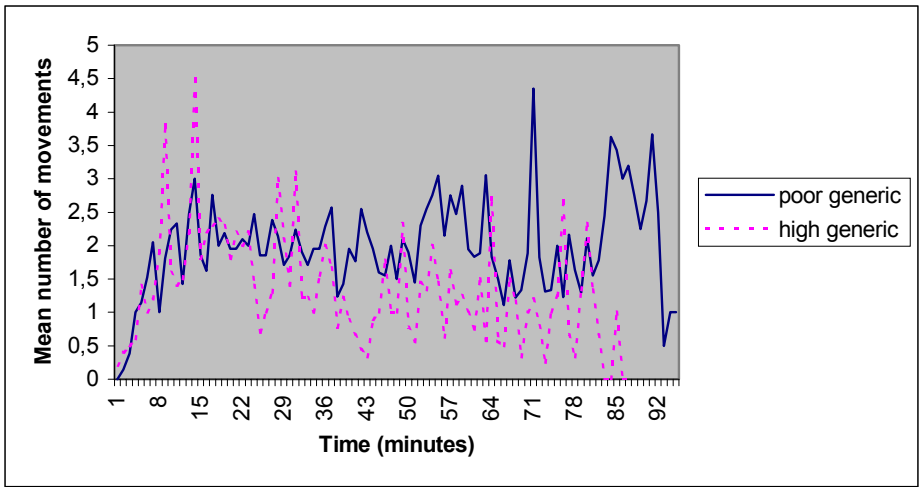


Figure 5-4. *Movements over time, for the generic knowledge groups*

From Figure 5-4 it can be seen that initially both groups act similarly, but at a later stage the groups start to diverge. This might imply that after an initial period, in which both groups perform similarly, the group with high generic knowledge starts working in a different way, in which less movements are needed.

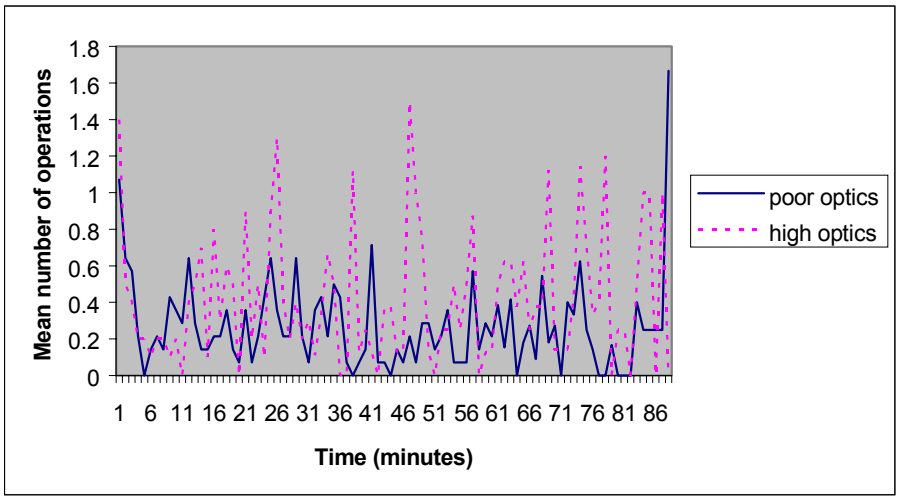


Figure 5-5. *Type II operations over time, for the prior Optics knowledge groups*

Figure 5-5 and Figure 5-6 display changes over time in the number of type II operations. To compute these graphs, the same procedure was followed as in the case of Figure 1 and 2.

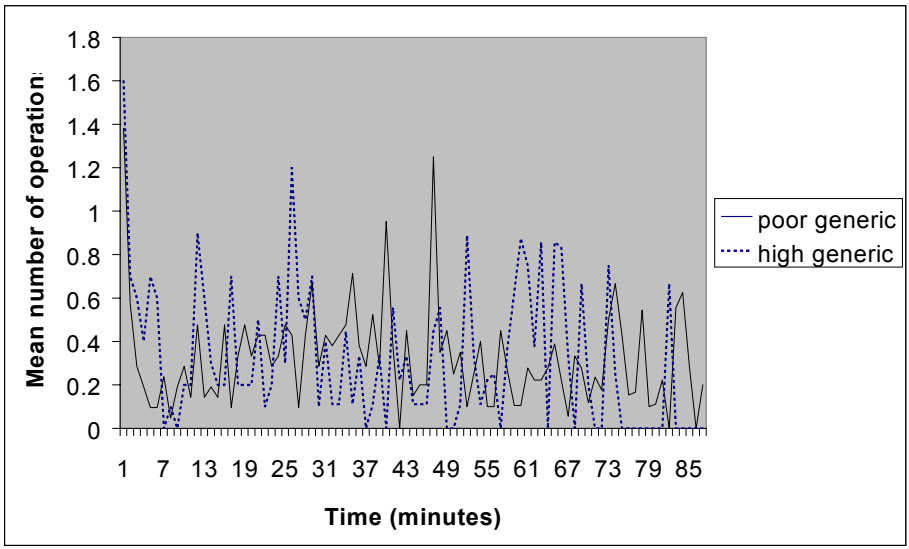


Figure 5-6. *Type II operations over time, for the prior generic knowledge groups*

Two things are apparent from these figures. The number of type II operations is at any time much lower than the number of type I operations, and there is a lot of variation over time, especially in the high optics group. Also, as can be seen from both Figure 5-5 and Figure 5-6, the number of type II actions does not, like the type I actions, settle after some time at a more or less steady point.

Object introduction times

In addition to checking different operation types over time, it is interesting to analyze at what moment, at an earlier or a later stage in working with the simulation, objects are introduced to the working area. Table 5-11 shows for each object after how many time steps it was introduced (time steps were computed in the same way as for the movement analysis). Variation in introduction times is considerable, therefore none of the group differences reach statistical significance. A problem with the this analysis is that we take into account all introductions, mixing real introductions with early, random ones. However, this effect is probably the same for all objects.

Table 5-11. *Mean number of seconds before introducing objects in the Optics environment*

	Object								
Group	**lens**	**lamp**	**screens**	**eye**	**construction line**	**angle**	**distance**	**helpline**	**formula**
Poor optics	80.8	138.1	251.6	256.8	714.7	815.3	900.7	986.8	1832.3
High optics	79.7	122.8	114.9	411.1	769.0	605.7	1389.0	1562.6	1550.7
Poor generic	101.1	195.1	245.7	486.2	924.5	580.3	879.9	1311.3	2385.0
High generic	94.9	154.4	120.8	357.4	1200.0	1035.1	1216.1	1482.8	1415.2
Poor Bubbles	91.3	110.1	163.3	206.3	441.9	428.0	776.6	907.6	1778.0
High Bubbles	64.3	287.8	127.8	310.8	818.1	950.4	1012.3	1648.4	2172.5

As can be seen from the table, the objects that are operated on in type I operations (lenses, lamps, and screens), are more rapidly recognized as being basic to experimenting by subjects than other objects, and are introduced earlier on. Also, these objects are the only objects that are used by *all* subjects. Construction lines, helplines, distances, and formulas are only used by a limited number of subjects; some learners spend the whole learning process only moving objects around. The eye object is introduced relatively early,

which indicates that many subjects start going beyond their basic range of experimentation with this object.

Specific object use

So far, we have looked at the general use of objects. In this section we look at some more specific aspects of interaction behavior. Two different methods to do this are: 1) check meaningful behavior patterns, and 2) check individual interaction behavior.

In addition to viewing interaction behavior as a sequence of individual operations, we can study it as a series of meaningful actions that are made up of more than one operation. These actions can be considered to be informative, which means that they contribute to understanding the workings of the learning environment. We distinguish between two levels of information-yielding actions: qualitative and quantitative. Quantitative actions are those actions that yield quantitative (numerical) information from the learning environment, qualitative actions do not yield quantitative information but are used to experiment in the environment, for example by moving different objects one after another. The list is shown in Table 5-12. As can be seen, all the meaningful actions we mention in Table 5-12 are composed of one or more operations. We analyzed both qualitative and quantitative actions by computing the mean number of occurrences of these actions over time for different groups of subjects.

Table 5-12. *Types of qualitative and quantitative actions*

Action type	Actions
Qualitative	• move object A, then move object B, then move object A again • add object A, delete object A
Quantitative	• add a distance between two objects and remove it immediately • add construction line and add distance measure • move object and check values • rotate light beam and check values

For technical reasons, the time measure we used here is a little different than that used for the type I and type II operations. This means that the unit of time here is not counted in seconds, but rather in 'time steps', where each operation is counted as occurring during one time step. For reasons of clarity, we took the mean number of occurrences of relevant actions over 25 time steps. Figure 5-7 shows qualitative actions for the poor and high domain-

specific knowledge groups. Occurrences of the two qualitative actions were summed, and a mean computed.

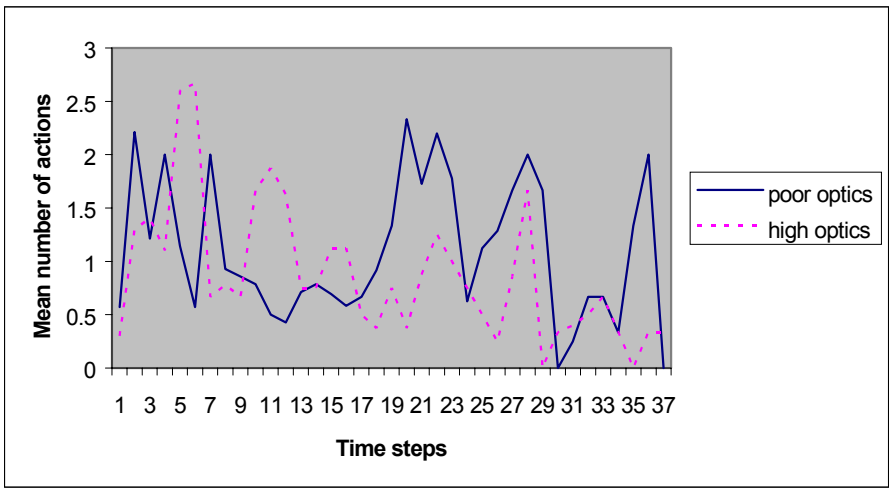


Figure 5-7. *Qualitative actions over time, for poor and high domain-specific prior knowledge groups*

As can be seen from the figure, there is a large variation over time in the number of qualitative actions that are performed. As can also be seen, the two groups start out in a similar way, but after about half the time has passed (at Time step 19) the paths diverge.

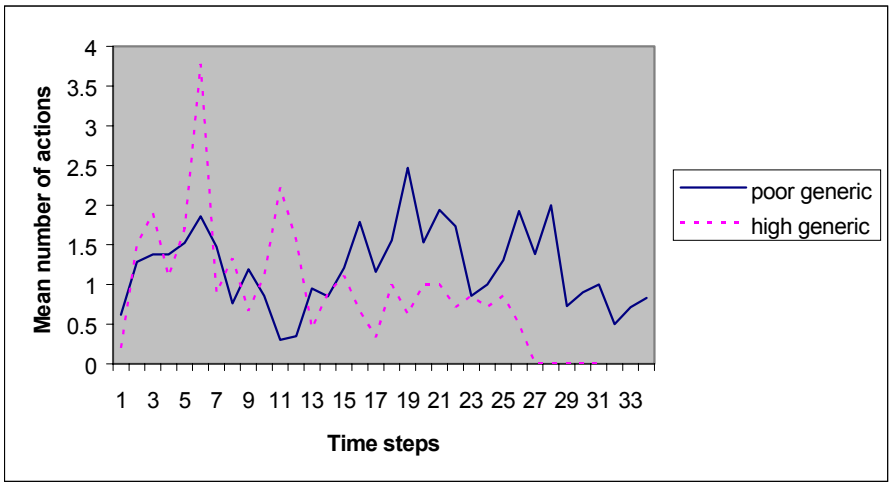


Figure 5-8. *Qualitative actions over time, for poor and high generic knowledge groups*

In Figure 5-7, after about half the time of experimenting has passed, the group with poor domain-specific prior knowledge group performs more qualitative actions than the group with high domain-specific prior knowledge.

Figure 5-8 shows a comparison of qualitative actions over time for the generic knowledge groups. This figure follows a similar pattern as Figure 5-7. After an initial period in which both groups perform roughly the same number of actions, the paths diverge again, with the poor generic knowledge group performing more qualitative actions than the high generic knowledge group.

In Figure 5-9, a comparison of the Optics pretest groups for quantitative actions is shown. Occurrences of all four quantitative actions were summed, and a mean computed.

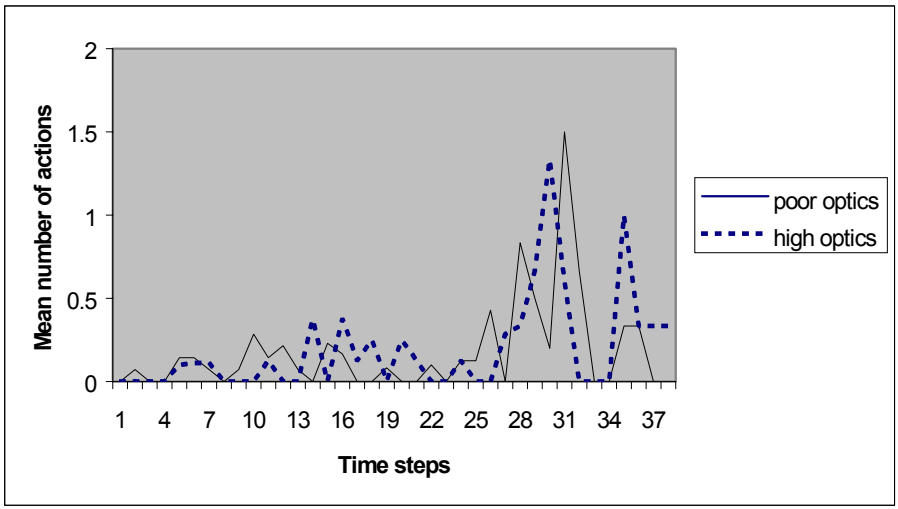


Figure 5-9. *Quantitative actions over time, for Optics pretest groups*

As can be seen, Figure 5-9 is very different from Figure 5-8. Quantitative actions only rarely occur after a long period of experimenting has taken place. This occurs both for the poor and high optics prior knowledge groups. These findings correspond with the results shown in Table 5-11 (introduction times of objects). Objects that are used to perform quantitative actions are introduced at a relatively late stage.

A comparison of quantitative actions for the generic knowledge groups is shown in Figure 5-10. Although the time scale for the high generic knowledge group is shorter than for the poor generic knowledge group, it can be seen that the poor generic knowledge group performs more quantitative actions than the high generic knowledge group. The number of quantitative actions stays at a very low level over the whole time trajectory for the poor generic knowledge group. As can be seen from the figure, the poor generic knowledge group performs more quantitative actions than both optics knowledge groups (as shown in Figure 5-9).

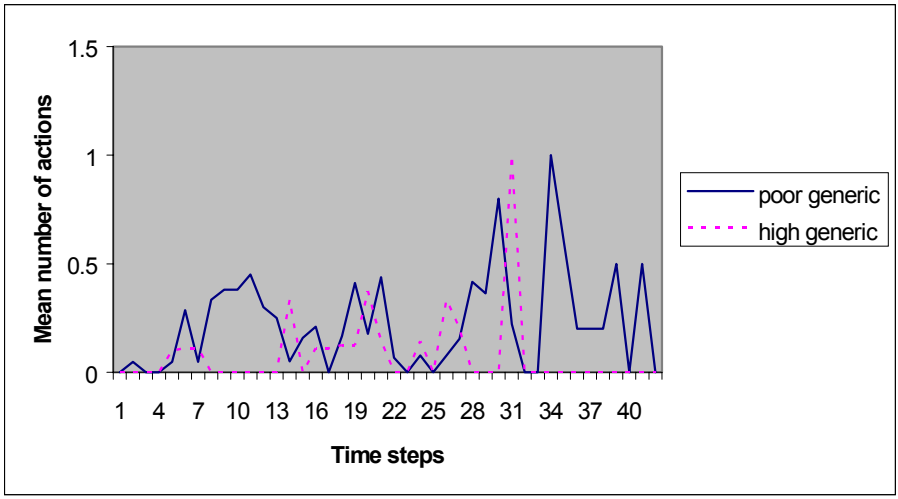


Figure 5-10. *Quantitative actions over time, for generic knowledge knowledge groups*

5.3.3.3 Interaction with the Bubbles simulation

We could not obtain interaction data for all 51 subjects. In total, 49 logfiles were obtained for the Bubbles environment.

The Bubbles environment follows a more straightforward experimental procedure than the Optics environment. The procedure can be described as: set one or more measure bars, fill one or more kettles, start the experiment and view the results. Thus, the environment is not continuous like Optics. By default, the Bubbles environment also stresses quantitative aspects more than the Optics environment. The amount of bubbles appearing is shown in the graph as a bar with a corresponding number. In both the normal and the enriched instruction condition, learners are urged not to focus on the numbers too much, but on overall relation between the amount of liquids and the resulting reaction.

Operations on objects

We compared frequencies of the different actions in the Bubbles environment for a number of groups. Table 5-13 shows the different comparisons for relevant groups. Although on the whole subjects who score high on the Bubbles test do not perform significantly more experiments (87.7 for the low scorers vs. 92.8 for the high scorers), there is a significant overall correlation ($r = .32$, $p<0.05$) between the number of experiments performed and per-

formance on the Bubbles test. This correlation stays significant when instruction group and generic knowledge scores are partialled out. It can be noted with respect to prior generic knowledge that subjects with high generic knowledge perform more experimental actions than subjects with poor generic knowledge, but make less notes. It may be that the subjects with high generic prior knowledge were more focused on finding overall relations compared to subjects with poor generic prior knowledge, who focused on the numerical outcomes of individual experiments.

Table 5-13. *Mean frequency of actions for different groups in the Bubbles environment*

	Action					
Group	**Fill Kettle**	**Move line**	**Start experiment**	**Pause experiment**	**View instructions**	**Notebook action**
Normal instruction (*n*=21)	67.6 (30.9)	21.8 (20.3)	25.6 (21.6)	3.9 (5.4)	6.5 (3.4)	67.2 (40.3)
Enriched instruction (*n*=28)	91.0 (64.2)	12.0 (10.7)	24.5 (15.0)	5.4 (8.6)	6.9 (3.1)	61.5 (39.1)
Poor generic (*n*=18)	68.5 (37.0)	16.8 (17.1)	26.2 (14.6)	2.9 (4.1)	6.2 (2.3)	76.1 (52.3)
High generic (*n*=17)	76.7 (38.9)	20.1 (17.6)	27.4 (24.1)	3.6 (5.7)	7.0 (4.2)	58.9 (27.7)
Poor Bubbles score (*n*=7)	87.7 (48.3)	17.4 (12.8)	31.6 (20.7)	9.9 (12.9)	5.7 (2.3)	78.9 (65.0)
High Bubbles score (*n*=12)	92.8 (63.8)	9.7 (9.1)	26.3 (19.0)	4.9 (7.1)	7.5 (4.0)	68.3 (29.3)

As can be seen from Table 5-13, like in the Optics environment there is a large variation in the number of different operations. Only about one in 5 experiments is paused while it is running, which shows that most experiments are carried out to the end of the trajectory. An interesting finding is the frequency of use of the notebook. Comparison with the results for the Optics environment shows that the notebook was used about six times as much in the Bubbles environment. This includes both making notes and looking back at old ones. An explanation may be that subjects, although urged not to do so, did lay emphasis on the numerical outcomes of different experiments. By systematically noting the numerical outcomes of experi-

ments in the notebook, they could keep track of these outcomes in relation to the different values used. Making notes helped exploring more possibilities. In fact, some learners even went as far as to make tables with outcomes in the notebook module.

Effects over time

In Bubbles, the interesting units of analysis are the experiments that are carried out. Experiments are carried out one at a time, and before an experiment is started the amount of liquid in zero to three kettles can be changed. To determine the effect of one liquid a good strategy is to change only one value at a time. To check how many values are changed per experiment, we computed the mean number of varied values for different groups of subjects. We expect subjects with high prior generic knowledge to show more systematic experimental behavior than the poor generic knowledge group, resulting in a mean number of changed values per experiment close to 1. Also, subjects who perform more systematic experiments should be able to extract better knowledge about relations in the simulation, and therefore perform better on the posttest; this means that we expect subjects in the high Bubbles score group to have smaller mean number of changed values per experiment than subjects in the poor Bubbles score group. The results, including p-values of the differences between groups (computed using t-tests) are shown in Table 5-14.

Table 5-14. *Mean number of value changes over experiments for different groups (standard deviations between parentheses)*

	Mean		p-value
Normal instruction (*n*=21)	1.31	(0.58)	
Enriched instruction (*n*=28)	1.33	(0.32)	0.92
Poor generic knowledge (*n*=18)	1.30	(0.50)	
High generic knowledge (*n*=17)	1.40	(0.45)	0.59
Poor Bubbles score (*n*=7)	1.58	(0.61)	
High Bubbles score (*n*=12)	1.32	(0.47)	0.36

As can be seen from Table 5-14, the expectations about differences between the different groups are not met by the results. There are no significant differences between groups, and the mean number of changed values per experiment lies between one and two for all groups. The values in Table 5-14 are informative about the overall systematicity of experimentation, but they say nothing about how the number of values that are varied changes over time. We expected that poor domain-specific knowledge would lead learners

to experiment erratically at first (to explore the domain and get to know the boundaries), and more systematic later. For generic prior knowledge, we expected that learners with high generic knowledge would show more systematic experimenting behavior than learners with poor generic knowledge. Figure 5-11 and Figure 5-12 show a comparison in the mean number of varied values for the different prior knowledge groups. On the X-axis individual experiments are shown, on the Y-axis the mean number of values that are varied before the experiment is started. The computed means are corrected for the number of experiments a learner has done. In Figure 5-11, the normal instruction group and the enriched instruction group are compared.

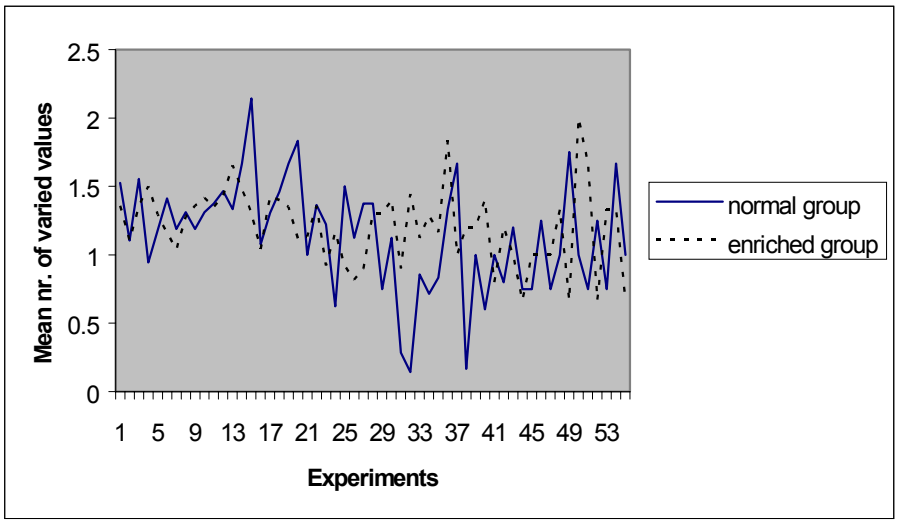


Figure 5-11. *Mean number of varied values in the Bubbles environment, for normal and enriched instruction groups*

It appears from the graph that there is more variation in the number of values that is altered before an experiment in the normal instruction group than in the enriched instruction group, especially from (approximately) experiment 10 to 40. In Figure 5-12, the generic knowledge groups are compared. After the 10th experiment the groups diverge, with the high generic knowledge group showing more variation than the poor generic knowledge group.

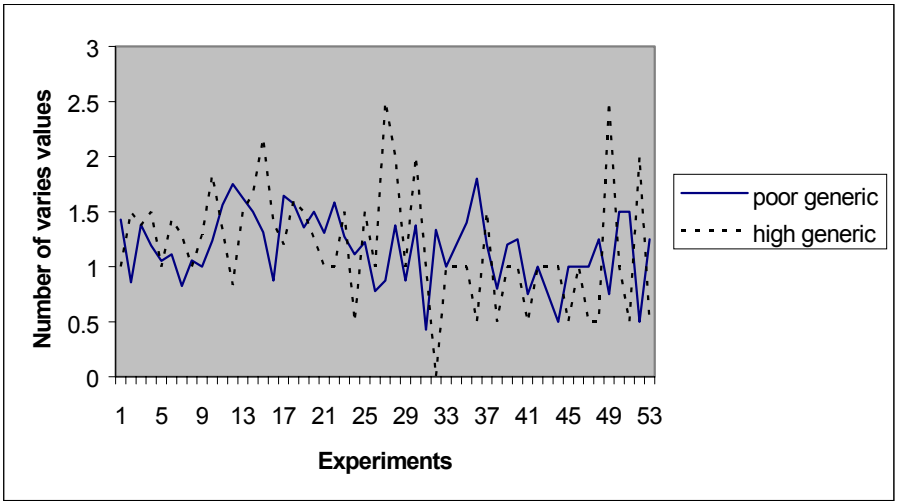


Figure 5-12. *Mean number of varied values in the Bubbles environment, for generic knowledge groups*

The result for the generic knowledge groups is a bit contrary to expectations. The poor generic knowledge groups stays close to changing only one variable at a time, whereas the high generic knowledge group behaves more erratically. However, an explanation can be that the high generic knowledge groups are quicker in recognizing relations between variables, and therefore more quickly attempt to discover more complex interactions.

To complete the analysis of interaction with the Bubbles environment, we computed the percentage of experiments carried out when no value had been changed from the situation before. There are two plausible reasons to perform the same experiment twice or more. Subjects may want to replicate an experiment, to make sure the same findings are obtained. Most of the time however, the measure bars are moved and the same experiment replicated to obtain new values (at different time spots). This can be considered as a strategically useful experimental manipulation. The percentages are given in Table 5-15, with standard deviations between parentheses.

Table 5-15. *Percentages of 'same-type'-experiments in the Bubbles environment*

	Group					
	Normal instruction	**Enriched instruction**	**Poor generic**	**High generic**	**Poor Bubbles**	**High Bubbles**
Percentage	27.1 (18.2)	18.2 (12.8)	25.5 (13.9)	17.2 (23.3)	21.2 (20.4)	24.3 (14.5)
Statistics	p=0.004		p=0.08		p=0.53	

As can be seen from the table, a reasonable part of all experiments, about one fourth, was carried out without changing any value prior to starting it. There is a statistically significant difference between the normal and enriched instruction groups, with the enriched instruction group performing less of the 'same type'-experiments. It appears that learners in the poor generic prior knowledge group also performed more of these experiments, than learners in the high prior generic knowledge group. These results are a bit paradoxical. One the one hand, subjects who received extra instruction may perform less 'same type'-experiments because they have received information on the type of relations in the environment in advance. These subjects did not have to rely as much on their generic knowledge of relations as subjects who received a normal instruction.

Earlier, we saw a significant correlation between prior generic knowledge and performance on the Bubbles test for the normal instruction group, but not for the enriched instruction group. On the other hand, subjects with high generic prior knowledge also tended to perform less 'same type'-experiments than subjects in the poor generic prior knowledge group. This indicates that subjects with high generic prior knowledge were quicker at recognizing a certain type of relation than subjects with poor generic prior knowledge. Because they had a better overview of the relations in the environment, they did not need to make as many notes as the poor generic knowledge group (as shown in Table 5-13).

5.4 Discussion

In the present chapter we have given an overview of a first experiment, in which the influence of domain-specific and generic prior knowledge on the process of discovery learning was studied. In this section the results are dis-

cussed. In the final chapter of this dissertation, the results of this study are combined with results of the other studies.

In this study, the scientific discovery learning behavior of students was investigated in two different learning environments, one on the (real-world) subject of geometrical optics ('Optics'), and one on a fictitious subject ('Bubbles'). The simulations allowed various properties of a simulated model to be manipulated by learners, in a quantitative as well as in a qualitative way. Before working with the computer simulations, knowledge about the subject of geometrical optics was *measured* and knowledge about Bubbles was *manipulated*. We separated between high and poor performance groups on the tests used in order to compare groups of different ability. In this section, we will discuss the different outcomes of the experiment with respect to discovery learning behavior and learning outcomes. An attempt is made to put the results in the perspective of the dual-space search model of Klahr and Dunbar (1988), and the extended dual space search model of van Joolingen and De Jong (1997). A number of limitations and shortcomings to the current experiment will be put forward, and possible improvements that are brought into practice in the Optics 99 experiment (described in Chapter 6) are discussed.

5.4.1 Prior knowledge and learning outcomes in Bubbles and Optics

In both the Optics and the Bubbles learning environments, no influence of domain-specific knowledge on learning outcome was found. For the Optics simulation this was shown by the absence of a relation between domain-specific knowledge about optics and learning performance, and for the Bubbles simulation by the absence of a difference in learning performance between the normal and enriched instruction groups. The result for prior generic knowledge turned out to be dependent, to some degree, on the level of domain-specific knowledge. Part of the subject group was minimally informed about the domain of the Bubbles simulation; the other part was given an enriched set of instructions, to increase domain-specific knowledge about the domain. It was found, that in the poor domain-specific knowledge group, prior generic knowledge played a more important role in determining learning outcome than in the enriched domain-specific knowledge group. Because students in this group only had generic knowledge to rely upon, they may have relied more on this type of knowledge than students in the enriched instruction group did. This, in turn, may have caused these students to perform 'better' experiments (as shown, for example, by the number of varied values per experiment, in Figure 5-12). In addition, the extra domain-specific knowledge of students in the enriched instruction

group, may have been compensated by students in the normal instruction group who relied on their generic knowledge. This may have been a cause of the equal learning outcomes that were found for the Bubbles simulation. However, the lack of difference in learning outcome between the normal and enriched instruction groups may also have been caused by the type of enriched instruction that was used in the experiment. The extra knowledge that was provided was, to a certain extent, both domain-specific and general in nature (for example, information about the type of relations to be expected in the simulation is not primarily domain-specific information). Learners may have focused on the 'general' aspects of the extra instruction, which caused them to shift their attention during learning from the generic knowledge they already had to the information provided in the instruction. Although the enriched instruction was effective in providing learners with knowledge about the background behind the Bubbles simulation and the type of relations to expect, when they made use of it learners may have neglected their available knowledge on how to interpret the results of experiments. Unfortunately, we cannot be certain of this since domain-specific knowledge of Bubbles was not tested after the instruction, prior to working with the computer simulation.

A number of explanations are possible to account for the lack of positive relations between prior knowledge measures and learning outcomes that were found for the Optics computer simulation. Because students worked only for a relatively short time with the simulation, and discovery learning was not guided by support measures (e.g., feedback) improvement in domain-specific knowledge was expected to be low. Also, the distinction between high and poor domain-specific prior optics knowledge suffered from the fact that domain-specific knowledge, as measured on the pretest, was around chance level. This lowered the reliability of the test considerably, and it indicates that before the experiment students showed at best a very poor understanding of the rules of geometrical optics. This is surprising, because students had received formal education about the topic of geometrical optics two years before. The instruction had included a presentation of slides that depicted similar situations as were presented in the domain-specific knowledge test. Yet, the subjects turned out to have forgotten most of the relations underlying the topic. Interestingly, in a study by Johnson and Lawson (1998), a comparable result was obtained. In the study, Johnson and Lawson compared two instruction methods for a biology course, inquiry teaching and expository teaching. They found scientific reasoning to be a much better

predictor of learning performance than domain-specific knowledge. This did not meet their expectation that scientific reasoning ability would be the more important factor in inquiry teaching, and domain-specific knowledge more important in expository teaching. Domain-specific prior knowledge was a poor predictor of learning performance in both conditions. The knowledge test used by Johnson and Lawson suffered from a low reliability, which was caused, as in this experiment, in part by poor performance on the domain-specific knowledge test.

We did not find a relation between generic prior knowledge and learning outcome of the Optics learning environment. A reason may be that the Optics domain-specific posttest did not ask for specific knowledge about quantitative relations. Available quantitative knowledge may therefore have remained unused in the knowledge tests. As stated before, the Bubbles environment is more quantitative in nature than the Optics environment. Lavoie and Good (1988) hypothesized that 'Computer simulations that quickly link the quantitative effects of changes in an independent variable to a dependent variable may be particularly helpful to persons who have difficulty in making accurate cause-effect linkages. (p. 355)'. When this idea is compared to the results that were found in the present study, and it is assumed that the ability to make an accurate link between cause and effect is related to generic prior knowledge, this may be an explanation why a relation between learning outcome and generic knowledge was found for the Bubbles simulation, but not for the Optics simulation.

The differentiation between the results that were found for prior generic knowledge and domain-specific knowledge provides evidence that studies that only take domain-specific knowledge into account to study discovery learning processes look at only one side of the coin: different types of knowledge may have contrasting effects on scientific discovery learning behavior.

5.4.2 Prior knowledge and scientific discovery learning behavior

Both domain-specific prior knowledge and generic knowledge turn out to be related to scientific discovery learning behavior. The influence of both types of knowledge was revealed by both the frequency with which operations in the computer simulations were performed, and analysis of the time course of different types of actions.

In the *Optics learning environment* we found evidence that after a certain amount of time time, students with poor domain-specific prior knowledge started performing more qualitative actions, whereas students with more

prior domain-specific knowledge performed less qualitative actions. Both the poor and high domain-specific knowledge groups also showed an increase over time in the number of quantitative actions that were performed. Combining the results from both qualitative and quantitative actions, we see that having domain-specific knowledge about geometrical optics meant that less qualitative operations had to be performed to start working on a quantitative level, or, that less experimental actions were needed to confirm hypotheses. This is shown in the decline over time of qualitative actions for the high knowledge groups. We can conclude from this, that learner interaction did not proceed in the same way over the whole learning trajectory. Rather, learner-computer interaction changed over time, in a qualitative as well as a quantitative way; in what way this change occurred was influenced by both domain-specific and generic prior knowledge. In terms of the dual space model of discovery learning, we can explain these results in terms of, for domain-specific knowledge, the initial configuration of hypothesis space, and, for generic knowledge, characteristics of the search process itself. The results can also be related to Klahr and Dunbar's distinction between an *experimenter* and a *theorist* 'strategy' of discovery learning. In their experiments with BigTrak, the rule to discover usually could only be derived after a *frame-shift* had taken place (from the role of the Repeat button as a 'counter' to a 'selector'). Subjects who induced the right frame as a result of an experimental outcome (thus, by searching experiment space), were classified as experimenters. The remaining subjects searched hypothesis space to make a frame-shift. The main difference that Klahr and Dunbar found between the experimenter and theorist strategy is that theorists performed fewer experiments than experimenters did before reaching a conclusion. Klahr and Dunbar did not relate both strategies to prior knowledge (this was not measured in their experiments). Our findings suggest that domain-specific knowledge influenced the configuration of the *learner hypothesis space* (as was hypothesized in Chapter 2). Having knowledge about the domain leads learners to change their experimentation style after a while; this change is reflected in a change from a qualitative to a more quantitative experimentation style. Although we found learners with poor domain-specific knowledge also to progress to a more quantitative level of experimenting, these subjects kept working in a qualitative way, more than the high knowledge groups.

In the *Bubbles learning environment*, we did not find clear results with respect to the influence of domain-specific and generic prior knowledge. There was an overall relation between the number of experiments subjects carried out,

and performance on the posttest: the more experimental results that were collected, the higher performance on the posttest.

A comparison of the normal instruction group with the enriched instruction group revealed no differences in the number of different operations that were carried out. Also, the groups did not differ in the mean number of values that were varied before an experiment was carried out. We did find, however, that subjects in the normal instruction group (poor domain-specific knowledge) showed a higher variation in the number of variables that were changed from experiment to experiment, than subjects in the enriched instruction group (high domain-specific knowledge). This may indicate, that subjects who did not receive extra information in advance behaved more erratically in this learning environment, because they knew less about the type of relations to expect in it. With respect to generic knowledge, it was expected that subjects with poor generic knowledge would need more data in order to see a relation between two variables than subjects with high generic knowledge. It was found that subjects with high generic knowledge showed a larger variation over time than subjects with poor generic knowledge in the number of variables that are changed. The finding that generic knowledge influenced learning behavior in the same way as domain-specific prior knowledge, that is, high generic knowledge caused learners to behave in a more varied way, can be explained by the idea that subjects with poor generic knowledge know less about mathematical relations in general than subjects with high generic knowledge. As explained before, learners with high generic knowledge are quicker at recognizing relations, which means that they sooner change their behavior into analyzing more complex relations, that involve more variables.

Discussion of the results has up till this point focused on two things: the relation between prior knowledge and learning performance, and between prior knowledge and learning interaction.

The overall difficulty we face is relating the negative results with respect to the relation between prior knowledge and learning performance with the (mildly positive) relations we found with learning behavior. The poor test results, and the difficulty in analyzing the interaction behavior (caused, in part, by differences between the learning environments), are two issues that are directly related to these results. Therefore, both these issues will be discussed in the remainder of this chapter. First, the knowledge tests that were used are discussed, after which a number of differences between the two computer simulations are considered. The conclusions have implications on the the studies that are described in Chapter 6 and 7.

5.4.3 Comments on the domain-specific and generic knowledge tests

In the experiment two types of knowledge tests were used. To measure generic knowledge we used a normal multiple-choice test that featured three- and four-answer items, and which was a paper-and-pencil test. To measure optics prior and post knowledge and Bubbles knowledge, we used intuitive knowledge tests that were administered on the computer. On the generic knowledge test, subjects experienced only on a limited number of items difficulty in answering. This reduced the variation in the test scores. A number of items were easily answered by most subjects. These items asked for the next number in a row of given numbers (for example, the sequence 3…5…10…12…24…26…??). These items could in principle all be solved by using the same strategy: that of looking at the quantitative difference between successive numbers. If we assume that the other items rightly tapped generic knowledge about relations, then a reason that we find an influence of this knowledge on learning behavior but not on learning result may be that another factor is involved in determining learning outcome. In the introduction we have argued that generic knowledge can be separated into two components, one of which is the knowledge of functional relations that we have tried to measure, and another which we can call 'discovery skill', and which involves the ability to induce relations. In the next section we will go deeper into this issue.

The other tests used the What-if test format that was developed by Swaak and De Jong (1996). As was explained in Chapter 4, in this test format both correctness and answer time are taken into (equal) consideration. More specifically, subjects are asked to trade off correctness and speed of answering, because both are of equal importance.

Both the Optics and the Bubbles domain-specific knowledge posttests showed a small but statistically significant increase in test scores. The average answer times showed that there was a drastic decrease in the total time taken for the test from pretest to posttest. Whether this finding indicates that learners were more familiar with the special format of the test after they had done the optics pretest, or that they had acquired intuitive knowledge, is difficult to say. It may be that subjects responded faster in the posttests because their responses followed more of an 'all-or-none' strategy. An intuitive knowledge test is designed to tap deep-rooted knowledge that is quickly available. The results on the optics pretest indicate that, prior to the experiment, subjects knowledge of geometrical optics was not only low but also

not of an intuitive quality. Students had learned the underlying rules, but had not developed any understanding of the relation between these formulae and actual behavior of light through lenses. After working with the computer simulation this situation was changed. Performance was still at a relatively poor level, but the knowledge that subjects used may have had a more intuitive quality as indicated by the decrease in answer time.

5.4.4 Similarities and differences between the Optics and Bubbles simulations

In summary, the results that were found in the present study can be described as follows: a number of relations between prior knowledge and interaction behavior in both the Optics and the Bubbles learning environments were found. The only effect of prior knowledge on test results that was found, is in the relation between generic knowledge and performance in Bubbles for the normal instruction group. Another interesting finding is that a few results point at a relation between performance in the Bubbles environment, and behavior in the Optics environment. This may be caused by a common factor underlying learning behavior, but it is not clear whether this factor is domain-specific knowledge, generic knowledge, or another. There are a number of similarities and differences between the Optics and the Bubbles learning environments that shed light on the results.

Both computer simulations were designed to use interfaces that looked similar, given the number of different options that were available to learners. Moreover, the learning environments were introduced in a similar way, as being exploratory environments. For both simulations, it was stated in advance that no knowledge about the actual formulas underlying the simulation was expected afterwards. Finally, similar assignments were used (i.e., to find out the rules underlying the simulation). It turns out, however, that there are a number of important differences between the two learning environments, which may have affected learning behavior and learning performance. These differences have to do with the type of experiments that are possible in the learning environments; these differences are summarized in Table 5-16.

Table 5-16. *Differences between the Bubbles and Optics learning environments*

Bubbles		*Optics*
Mostly quantitative	**Interaction type**	Qualitative and quantitative
Present	**Experiment boundary**	Absent
Only few changes per experiment possible	**Variable change**	Many changes possible

Interaction type, experiment boundary, and variable change influence the type of experiments subjects perform, the type of conclusions that are drawn from these experiments, and subjects' motivation to experiment at all. In the Bubbles environment, although qualitative changes were made possible through the gauges, and the number of Bubbles appearing was presented in a graphical and numerical way, learners focused on the quantitative aspects of the environment. This is reflected in the type of notes that are made, with the notes about Bubbles being more quantitative than the notes about Optics. Furthermore, working with the Bubbles environment proceeds in a necessarily more systematic way than working with the Optics environment, for two reasons. One is that there is a fixed boundary to experiments in the Bubbles environment: they start as soon as the 'start experiment' button is pressed. This forced 'experiment boundary' may have had a supporting effect on learning, because it structures the learning task. This is further influenced by the number of variable changes over time. In Bubbles, only a limited number of variables can change from experiment to experiment. In Optics, the number of variables that change as a result of an operation is dependent on the particular setup of the working area, and therefore can range from zero to many.

As a side issue, it is interesting to note that we observed that subjects showed much more interest in working with the Bubbles environment than with the Optics environment. A reason for this may be that the subject of optics is an actual school subject and therefore 'boring' to the students. Although both learning environments were more or less presented as being games, it is not unimaginable that only the Bubbles environment was perceived as having enough game-like characteristics (completely unknown topic, unrelated to school work) to be interesting to play around with. Nor-

man (1993) states that '...games require just the behavior we wish those same children would apply to schoolwork. What is the difference between these informal experiences and the formal, structured behavior of the classroom? There is something captivating about informal learning. Why can't we get the same devotion to school lessons as people naturally apply to the things that interest them?'. It follows from this argument, that motivation and interest in the learning environment may also contribute to learning performance.

5.4.5 Discovery skill: another influencing factor?

In Chapter 2, we have described a model of scientific discovery learning that relates the learning processes involved in discovery learning to a special type of problem-solving: one that includes two problem-spaces, hypothesis space and experiment space. From the analysis of prior knowledge and both interaction behavior and learner performance, it has become apparent that prior knowledge may not be the only influencing factor on discovery learning behavior. This is shown, for example, by the fact that students who scored high on generic knowledge performed less quantitative operations in the Optics simulation, showed less variable manipulation behavior in the Bubbles simulation, but showed similar performance on the posttests. This means that subjects with high generic knowledge, although they performed different types of experiments, may not have derived relations from experiments more than subjects with poor generic knowledge. These results can be explained by introducing a factor that influences discovery learning behavior and learning performance, next to prior knowledge. We put forward the hypothesis that this mediating factor is discovery skill. As stated before, this skill constitutes the ability to induce a relation from experimental data. For both learning environments, prior domain-specific knowledge, both for the poor and the high knowledge groups, was relatively low, which means that the difference between these groups were relatively small. Therefore, differences in knowledge may be obscured by differences in discovery skills. Although we did not use a measure of discovery skills in the present study, we did find some evidence for these claims if we consider performance in the Bubbles environment to be related to the ability to induce relations. As stated before, interaction in the Bubbles simulation was mostly quantitative and proceeded in discrete experiments. In addition, the relevant variables in the environment were more directly available to students than in the Optics simulation. To collect data in the Bubbles simulation may have been easier for students than collecting data in the Optics simulation, and the emphasis in the Bubbles simulation may lie more on the ability to induce relations.

Although we did not find performance in Bubbles to be related to improvement in the Optics simulation, we did find a relationship with a number of global interaction measures in Optics. This result indicates that, in addition to prior domain-specific and generic knowledge, there may be an influence of discovery skills on scientific discovery learning processes. This consideration has led us to carry out a second study for this research project, which is described in Chapter 6.

6

Optics 99: The role of domain-specific knowledge, generic knowledge, and discovery skills on discovery learning processes

6.1 Introduction

In Chapter 2, the process of scientific discovery learning has been described as a special type of problem-solving. The model of problem-solving that was developed by Newell and Simon (1972) views the process of solving problems as a search process. The model was extended by Klahr and Dunbar (1988) so that both the process of experimentation and the process of generating hypotheses would be incorporated. Thus, the process of discovery learning can be seen as an active search through both a space of possible hypotheses and a space of possible experiments. Van Joolingen and De Jong (1997) have extended the dual-space search model of discovery learning by distinguishing between different subspaces of the hypothesis and experiment space. For instance, the *learner hypothesis space* refers to all statements that a learner can assert about relations between two or more variables in a domain, and the *effective learner space* refers to statements in the learner hypothesis space that a learner finds useful to test. In other words, the knowledge that a learner has about a domain, and the knowledge about the type of relations that are worth testing, determine the contents of the learner hypothesis space and the effective learner space. This implies that prior knowledge should have an effect on the discovery learning process, because the level of knowledge influences the distance between the learner hypothesis space and the final (correct) model (called the *target conceptual model*). From previous research it had become clear that in some cases prior knowledge could greatly affect learning outcomes: persons with a higher amount of prior domain-specific knowledge tended to be more effective learners. However, many studies that investigate differential effects in discovery learning use a post-hoc approach, focusing on individual differences in performance on a knowledge posttest. Thus, prior knowledge is only rarely singled out as a possible influencing factor. In the Optics 97 study (described in Chapter 5) the goal was to explore the influence of differences in prior generic and domain-specific knowledge on the process of scientific discovery learning. Scientific discovery learning was operationalized by using a configuration of

the Optics computer simulation (described in Chapter 3) and by using the Bubbles simulation. Domain-specific knowledge was both measured (for Optics) and manipulated (for Bubbles). The results of the Optics 97 experiment showed that prior domain-specific knowledge did not influence learning outcome for the Optics computer simulation. A relation was found between generic knowledge and performance in the Bubbles environment, but the results were not unequivocal. With respect to interaction behavior with the computer simulations, limited influence was found for both domain-specific knowledge and generic knowledge. A division between qualitative and quantitative actions was made, and differences between groups of poor and high ability on the prior knowledge tests were measured. The analysis showed that initially, students of poor and high ability performed similarly, but after some time the patterns diverged. However, again it turned out to be difficult to obtain unequivocal results.

The experiment that is described in this chapter aims at further investigating differences in scientific discovery learning behavior between students with poor and high prior knowledge. An attempt was made at overcoming some of the difficulties of the Optics 97 experiment. This was done by improving on a number of aspects of the first study. These improvements, as well as the specific objectives for the current study, are covered in the following section.

6.1.1 Study design and objective

In designing a follow-up study to the Optics 97 experiment, a number of considerations were taken into account. Most of these considerations were based on improving on the first experiment, so that better (that is, more clear) results would be found. Improvements were made to two aspects of the study: the knowledge tests and the experimental setup. Also, a more varied group of students participated in the experiment.

To begin with, the tests for prior knowledge from the Optics 97 experiment showed a number of limitations. The optics domain-specific knowledge test consisted of items that were relatively complex. This lowered the reliability of the test[1] as well as the variance in performance. The test for generic

[1] It is unknown, however, whether a 'classical' measure for reliability (in these experiments, Cronbach's alpha) is appropriate to use in the case of a what-if test. In the past, attempts have been made at approaching the relation between speed and accuracy from a psychometric point of view (e.g., Furneaux, 1973). Recent research on this topic suggests that creating an item-response model of speed/accuracy trade-offs may be a possibility (Van Breukelen 1989; 1997).

knowledge contained only a limited number of items that showed variation in subjects' answers. On the other items, a ceiling effect was found, which lowered overall reliability and reduced variance. Both the domain-specific and generic knowledge test were completely redesigned for the present study, and the number of items was extended.

From the discussion of the results of the Optics 97 study, it was concluded that students' knowledge about discovery itself (or: their level of 'discovery skills') might be an additional factor in explaining differences in learning behavior. Therefore, a new test was added to measure *discovery skills*. The test was developed for the present experiment, and involved the use of a combined computer and paper task. The knowledge tests, and the experimental context of this study, are described in detail in the method section.

Secondly, students in the previous experiment worked in three distinctive 'phases' in the Optics simulation. There was no tight control over the moment of switching from one phase to the next. Because not all students worked for the same amount of time in each phase of the simulation, a reliable analysis of the data for groups of subjects was difficult. Also, although the phases were increasing in complexity, they were not related to each other, which reduced the coherence of the learning experience. The current study used a different design which made use of 'situations' in the Optics simulation. The situations were similar to the original phases of the first experiment, but an attempt was made to have students work for a fixed amount of time with a particular configuration of the Optics simulation. This was done to make the experimental situation more controlled. It was decided to let all students meet certain situations in the Optics simulation. Therefore, the present study used four different situations in the simulation. The contents of these situations ranged from (relatively) simple to (relatively) complex.

Finally, the subject group that was used in the first experiment consisted of students who were following pre-scientific education. The students tend to score high on intellectual ability, which reduces variance in the type of experimentation skills that are employed. For this study, it was decided to extent the group of students in pre-scientific education with a similar group of students who were in technical vocational education. By using students from two different types of schools it was expected that the heterogeneity of the subject population would be increased.

The changes to the knowledge tests, the subject group, and the experimental setup that have been described here constitute the main parts of the design of the current study. It was carried out as a follow-up study to the experiment that was described in the previous chapter. The results of the first experiment served as the basis for the design of the second experiment. This experiment constitutes, in many respects, an improved version of the first experiment. The specific configuration of the Optics simulation and the experimental procedure that was followed are both explained in the Method section.

In summary, the current study also aimed at studying the relation between discovery learning processes and domain-specific knowledge, generic mathematical knowledge, and discovery skills. The goal was to find differences in the experimentation behavior of learners with varying amounts of prior knowledge and skills. These could then be related to the SDDS framework. The experiment made use of the Optics computer simulation, which is able to provide a varied, complex, and information-rich discovery learning environment. To study discovery learning processes in a variety of situations, in this experiment students worked with four situations of different complexity in the Optics simulation. Learning behavior was compared over the different situations. All student-computer interaction behavior was registered. The method for analysing interaction behavior that was described in Chapter 4 was modified for the present study. The analysis focused on the operations that had shown in the Optics 97 study the largest differences between groups. Also, the contribution of operations in relation to other operations was examined. In-between each of the four situations, students were asked what they had done in the previous situation, and why. This allowed for quantitative data (learner-computer interaction data) to be compared with qualitative data (reasoning about actions). Finally, the experiment made use of two groups of subjects: students who were in technical vocational education, and students who were in pre-scientific education. These groups differed in intellectual background. It was expected that this would result in a more varied subject population.

The redesigned test for generic knowledge was first piloted in January 1999 with a group of students at a school for vocational education in Hengelo. This led to some alterations to the test. Both the domain-specific knowledge test, the (revised) generic knowledge test, and the situations in the Optics simulation were then piloted in March 1999 at a school for vocational education in Rotterdam (ROC Zadkine). Again, alterations to the design were made and it was decided to include the test for discovery skills in the design. A version of the test for discovery skills was piloted with three students at a

school for vocational education in Enschede. The final experiment was performed in two parts: in June (secondary school students) and November of 1999 (vocational education students).

6.1.2 Research questions and expectations

The primary research question of the current experiment is: how do prior domain-specific knowledge, generic mathematical knowledge, and discovery skills influence discovery learning processes? A second research question focuses on what differences in scientific discovery learning behavior can be found when (relatively) simple configurations of the Optics learning environment are compared with complex configurations.

In Chapter 2, theoretical considerations with respect to the influence of prior knowledge on search processes in hypothesis space and experiment space were discussed. For the Optics 97 study, this led to a number of expectations with respect to discovery learning in the Optics and Bubbles computer simulations. For the Optics simulation, it was expected that students with high domain-specific knowledge would show more hypothesis oriented behavior, and students with poor domain-specific knowledge (combined with high generic knowledge) to show a higher level of experimentation, and that this pattern would change over time. For students with both poor generic and poor domain-specific knowledge, unsystematic discovery learning behavior was expected. The analysis focused on differences in activity (over time) between students with high and poor knowledge (domain-specific and generic). For the present study, it was found useful to be more specific in the expectations. A division was made between expectations about the level of discovery learning activity that students show and expectations about the experimentation style that students follow. These are discussed separately here.

6.1.2.1 Expectations about the level of activity

The level of activity is represented by the number of operations that students perform while they are operating the Optics learning environment. Activity levels can be compared across students and across different situations in Optics. The operations that the Optics learning environment allows for can be grouped in various ways. In the Optics 97 study, a distinction was made

type I and type II operations. Type I and type II operations are respectively related to operations that are performed to directly *manipulate* the environment, and operations that *guide* working with the environment. When one or more objects are added, deleted, or moved around, active manipulation of the learning environment leads to the working area of the Optics simulation being visibly altered. Guiding operations are meant to make working with the simulation easier. For example, notes can be made, or the assignment and interface instructions can be reviewed. Another way to guide experimentation is by adding quantitative measures to the working area. Students can, for example, measure distances between various objects or add vertical lines at any point. In Chapter 4, a distinction was made between qualitative and quantitative operations. Manipulation and guidance operations can both be qualitative or quantitative. Global expectations with respect to the influence of domain-specific knowledge, generic knowledge, and discovery skills on students' level of discovery learning activity (both guidance and manipulation operations) are listed in Table 6-1.

Domain-specific knowledge is expected to influence the configuration of the learner hypothesis space and effective learner search space. Students with high domain-specific knowledge have more variables and relations in their hypothesis space, which means that they should find it more easy to generate new hypotheses. Also, students with poor domain-specific knowledge are expected to be able to understand simple situations in the Optics simulation, but not complex situations. Therefore, students with poor domain-specific knowledge are expected to be more active in simple situations than in complex situations. Furthermore, students with poor domain-specific knowledge are expected to need more guidance in their activities, because they cannot relate the data they find to prior knowledge about the relations in the learning environment. With respect to quantitative guiding operations, we expect students high domain-specific knowledge to perform relatively complex (quantitative) measurements more easily than students who have poor domain-specific knowledge. Students with high generic knowledge are also expected to use more complex quantitative measures.

Generic knowledge influences the size of the learner hypothesis space. Students with poor knowledge of mathematical relations should find it difficult to extend their effective learner search space, because of the restrictions imposed by their limited knowledge of mathematical relations. This is expected to result in less consistent measurements, for example students with poor generic knowledge will perform one type of measurement instead of two that can be related to each other.

Table 6-1. *Expected behavior in the Optics simulation environment, depending on different types of prior knowledge*

Prior knowledge type	*Poor knowledge*	*High knowledge*
Domain-specific knowledge	- More active in simple situations	- Less active in complex situations
Generic knowledge	- More active overall	- Less active overall
Discovery skills	- Less active overall	- More active overall

The level of discovery skills are expected to influence search processes in both experiment space and hypothesis space. Students with poor discovery skills are expected to have more difficulty in performing certain types of experiments (for example, experiments that test a specific hypothesis), and in carrying out specific experimentation strategies (for example, search for the effect of a variable in a series of experiments). Students with poor discovery skills are more limited in the types of experiments they can perform. They are expected to be less active in experimenting than students with high discovery skills.

6.1.2.2 Expectations about the experimentation style

In addition to the overall level of activity, we can look at differences in experimentation style. Experimentation style is here represented by three indicators: the use of operations, object use, and measurements. These indicators can be compared both within one situation in the Optics learning environment, and across situations.

The *relative use of different operations* indicates what students focus on while operating Optics. For example, a student may focus on adding a lamp and a lens to the learning environment, and then perform all kinds of operations on these two objects. In contrast, another student may focus on adding different objects to the learning environment and deleting them almost immediately, without performing many manipulations. The *relative use of different objects* indicates which objects students prefer to use in a situation in Optics. In the learning environment, students have different objects of the same kind at their disposal, e.g., different types of lenses. One student may focus on one type of lens only, while another can focus on comparing different lenses. This is reflected in the number of times objects are used. Therefore,

the relative amount of object use can indicate differences in discovery learning strategy. The *type of measurements made* indicates students' ability to induce quantitative relationships between different variables. For example, a student may only focus on measuring the focal point distance of a lens (a quantity that is set to a value that cannot be changed), while another may combine different measures that relate object distance to magnification.

Students with poor domain-specific knowledge will experience trouble in making good use of the available objects. Their use of different objects in the Optics simulation is expected to be more limited than that of students with high domain-specific knowledge. Students with poor generic knowledge are expected to use more simple quantitative measurements than students with high generic knowledge. Students with poor discovery skills will have more difficulty in creating different situations in the Optics simulation than students with high discovery skills.

6.2 Method

As was explained in the introduction, the present experiment is a follow-up study to the experiment that was described in Chapter 5. These differences are found in the subject groups that were used, the structure of the computer simulation, and the knowledge tests. In this section, the structure of the test for discovery skills (the so-called 'Peter task') is also explained in detail.

6.2.1 Optics configuration

In this experiment we used an implementation of the Optics simulation environment. For the present experiment it is important to note that the subject of geometrical optics is taught both in pre-scientific education and in technical vocational education. All subjects worked with four different situations in the Optics environment. The situations differed in the type and number of operations that were available. In general, they ranged from relatively simple (situation 1) to complex (situation 4). Table 6-2 gives a detailed overview of the objects and operations that were available to subjects in each of the four situations.

Table 6-2. *Availability of objects and operations in the four Optics situations used in the experiment*

	1	**2**	**3**	**4**
Objects				
Lamp – one light beam	X	X	X	X
Lamp – parallel light beams		X		X
Lamp – divergent light beams		X	X	X
Lens A and B	X	X	X	X
Lens C and D		X	X	X
Operations				
Move object	X	X	X	X
Delete object	X	X	X	X
Measure angle		X	X	X
Construction line	X	X	X	X
Distance measure	X	X	X	X
Rotate light beam	X	X	X	X
Show measured values	X	X	X	X
Make note	X	X	X	X
Delete all objects			X	X
Big lamp on screen			X	
Add one of the screens			X	
Two lenses simultaneously				X

All subjects worked with the four situations in the same order (that is, they all started with the most simple situation and ended with the most complex one). Available objects and operations are marked with an X. In addition, help on the Optics interface, and the assignment for current situation, were available in all situations.

To summarize Table 6-2, each situation can be characterized by the following point:

1) Situation 1: Only one lamp and two lenses available
2) Situation 2: Extend the operations of situation 1 and add two more lenses
3) Situation 3: Use of big lamp and screens
4) Situation 4: Two lenses at the same time possible.

Figure 6-1 shows for each of the situations an example of the Optics interface. Notes were not made on paper, but inside the learning environment itself; subjects were not supplied with writing paper. Instructions on the meaning of buttons and operations were available on paper as well as in the learning environment itself.

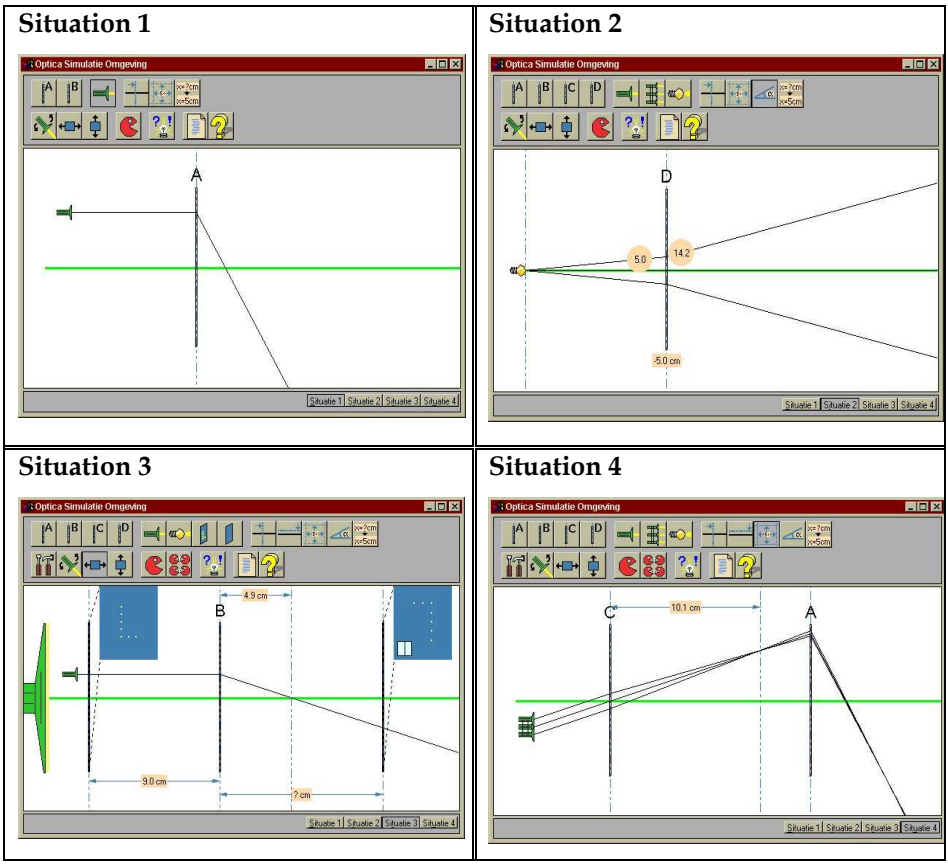


Figure 6-1. *Examples of the Optics interface for the four situations used; options available for each situation as presented in Table 6-2 are reflected in the toolbars for the situations*

6.2.2 Subjects

Two different groups of subjects participated in the experiment, a group of students that was following pre-scientific education (PSE group), and a

group that was following technical vocational education (TVE group). These two groups were different in background education, which implies that they would differ in abilities. In the data analysis, both groups are studied as one combined group, as well as separately. In the PSE group, 29 students, in their fifth year of study (age 16-17), participated in the experiment. In the TVE group, 33 students who were in their second year of study (age 17-18), participated in the experiment. In both groups, most students had previous computer experience: all were proficient at operating the mouse and keyboard. Also, in both groups all subjects were attending physics classes at the time of the experiment.

6.2.3 Prior knowledge tests

Prior to working with the Optics learning environment, three tests were administered: a test for domain-specific knowledge about geometrical optics, a test for generic knowledge about mathematical relations, and a test for experimentation skills in the context of a discovery task. While working with the learning environment, all actions subjects performed were registered. In addition, in between situations subjects answered a question about their behavior in the situation they had just worked with. Finally, after the last situation, a posttest on optics domain-specific knowledge was administered.

6.2.3.1 Generic knowledge test

The test for generic knowledge about mathematical relations was a revised and extended version of the test used in the first experiment. The test was modified to achieve a higher overall reliability, and to allow for a better distinction between poor and high knowledge groups. The test items were evaluated by different teachers. The items covered several topics about mathematical relations, and a number of ways of depicting these. None of the relations that were used in the test were more complex than relations that would explain mathematical relations in the Optics learning environment. The paper-and-pencil test that was administered consisted of 32 items. Of these, 29 were four-answer multiple-choice questions. The other items required a short answer, such as a formula.

6.2.3.2 Domain-specific knowledge tests

Like the test for generic knowledge, the test for domain-specific knowledge about geometrical optics was revised and extended for the current experiment. A number of items from the test that was used in the Optics 97 experiment were discarded because the results showed that they were too difficult for students. The test that was administered consisted of 30 multiple choice questions. The contents of the items closely followed the material that was offered in the four situations in the Optics environment. For each item, a situation was shown as it might occur in the Optics learning environment. About this situation a hypothetical question was asked, and subjects chose one out of three possible answers. The test was administered by computer. In the Optics 97 study, subjects ware asked to trade off speed and accuracy in answering the items. It is hypothesized that students may have had trouble in understanding this demand, which may have negatively influenced their performance on the test. Therefore, in the present experiment the 'time limit' was not used. This means that subjects were not asked to trade off speed and accuracy of answering. In Chapter 8, this issue is further discussed.

6.2.3.3 Discovery skills test

To measure discovery skills, we used a combined computer and paper-and-pencil test, called the 'Peter test'. In this section, the Peter test will be described in detail. The test was not only used in the current experiment, but also in the experiment that is described in Chapter 7. The actual task that was used is based on the 'Peter task', which was developed by Wilhelm et al. (2001). Subjects first worked with the computer task, in which their task was to find out the effect of 5 variables on an outcome variable. Afterwards, a test was administered to measure students' knowledge of these variables.

The task that was used in this experiment provided subjects with a concrete (as opposed to abstract and meaningless) context. They were given a background story about a boy (Peter) who always is too late for school. Peter wants to try to find out how different factors affect the number of minutes he arrives too late (for simplicity, it is assumed that he never arrives too early at school). Five possible causes that could influence the number of minutes that Peter is too late were selected. For each of these variables, two or three different levels could be chosen.

Figure 6-2 shows the computer task interface. Subjects could pick the variables from the small rows on the left-hand side of the screen. They would then appear in the row on the right-hand side.

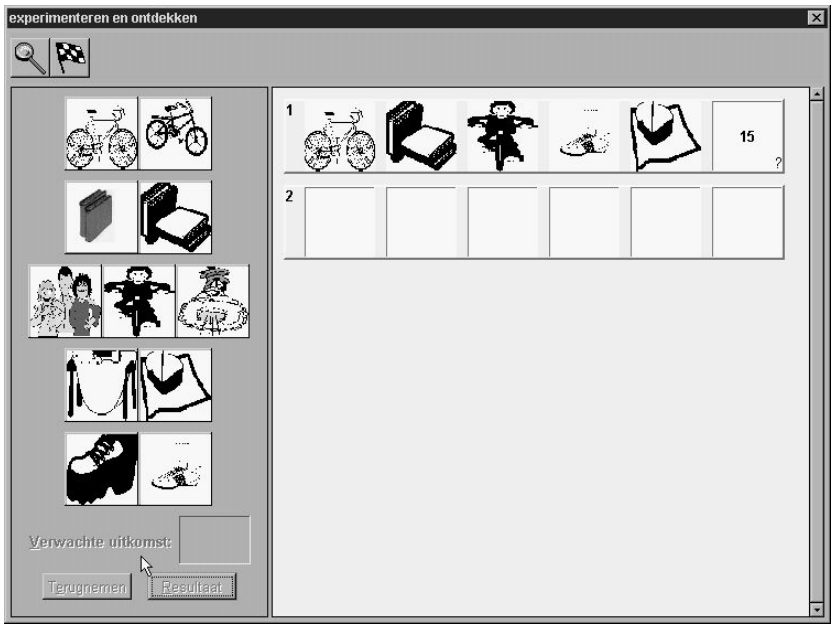


Figure 6-2. *The Peter task computer-interface. On the left, possible options are shown. The row on the right represents an experiment in the task, in which Peter arrives 15 minutes late for school.*

When all 5 variables had been selected, subjects first had to predict the outcome. There were 5 possible outcomes: 0, 5, 10, 15, or 20 (in our specific case this represents the number of minutes that Peter arrives too late at school).[2] In addition, when subjects did not know what to predict, they could fill in a

[2] The fact that the outcome represents the time Peter arrives too late at school, implies that under no condition he can arrive too early. This fact did evoke some comments from subjects, but did not interfere with the task itself.

question mark as prediction. Both the prediction and the real outcome were then shown next to the row with the 5 variables. One row, that is, the 5 chosen variables and the outcome, represented an experiment. By performing different experiments results could be combined and rules inferred. The idea behind the task was that by examining the outcome for different levels for each of the 5 variables in the Peter task, their effect could be discovered. After working with the Peter task, subjects were given a paper-and-pencil test. The paper-and-pencil test was handed out to subjects when they indicated to the experimenter to know all the effects of the five variables in the environment. After 30 minutes, the remaining subjects were given the test. As soon as subjects started with the test, the experimenter made it impossible (by typing a password) to perform any new experiments in the environment. The experiments that were already done remained visible. In addition, subjects were allowed to use any notes they had made on paper. The questionnaire consisted of six open questions. Five questions dealt with the effect on the total score of the five different variables in the task. The sixth question was meant to ask about any further effects subjects had noticed in the Peter task. Three variables in the environment had a main effect on the final score, two had no effect, and two of the main variables interacted with each other. On the basis of their answers, subjects were classified into one of four categories:

1) Subject not able to find one or more main effects; no interaction effect found
2) Main effects found, but no interaction effect found
3) Both main effects and interaction effect found, also additional (nonexistent) effects found
4) Both main effects and interaction effect found, no additional (nonexistent) effects found

Students who were classified in the first group represented the poorest level, and students who were classified in the fourth group represented the highest level of discovery skills.

Because all actions performed by subjects were registered, in addition to data from the written answers, afterwards information was available about the type of experiments that subjects performed in the Peter computer task. These experiments provided information about the number of experiments that were performed, and about the experimentation style that was used. Both results from the student-computer interaction during the task, and from the answers to the paper-and-pencil test, are interesting for this study. Results from the test, as well as some measures from experiments that are

carried out in the Peter task, are informative about the skills subjects have in finding relations of varying complexity (from main effects to a first-order interaction). In addition, the type of experiments that were performed in the Peter task are informative about the discovery process itself. Although the Peter environment appears to be less complex than the Optics task, it offers interesting possibilities to find patterns in scientific discovery behavior because the variables that are changed from one experiment to another can be counted. This provides information about the type of strategy used.

6.2.4 Interaction registration

The operations that subjects performed in the Optics learning environment were registered by the computer the subject was working on. As was explained before, this procedure consisted of storing every mouse-click on disk in a 'log file'. Along with storing the type of operation, time data was saved, so that the exact timing of operations could be analyzed afterwards. From the stored data, it was possible to replay a subject's process of experimenting in the Optics simulation quite accurately. To enable statistical interaction-analysis, the log files were converted to SPSS-readable data. This procedure was explained in detail in Chapter 3.

6.2.5 Procedure

Both the pre-scientific education group and the technical vocational education group followed the same procedure. Table 6-3 shows the sequence of events for both conditions. In the pre-scientific education (PSE) group, subjects were divided over two classrooms. In both rooms the same instructions were given, and the procedures were equivalent. In the technical vocational education (TVE) group, two classes participated at different times. Like the PSE group, the same procedure was pursued for both classes.

Table 6-3. *Sequence of events and measurements in the Optics 99 experiment. Time taken (in minutes) is shown in parentheses.*

First session		**Second session**	
Generic knowledge test	(30)	Four situations in the Optics simulation	(60)
Optics knowledge pretest	(20)	Optics knowledge posttest	(20)
Discovery skills test	(40)		

In the PSE group, the experiment was carried out over the course of one day. The first and second session were divided into a morning and an afternoon session. Between the tests and between the morning and afternoon session there were frequent breaks. Subjects in the PSE group came to the university to participate in the experiment. They were divided over two groups of respectively 11 and 18 persons, and went to two computer rooms (there were not enough computers available in each room to have all the subjects in one room). Each subject was seated behind his or her own computer. In the TVE group, the two experiment sessions were separated by one week. The first session was equivalent to the morning session of the PSE group, and the second session to the afternoon session of the PSE condition. Again, each subject worked individually at the computer.

For all subjects, the Optics situations were presented in a fixed sequence. Subjects started with situation 1, then after 15 minutes a signal was given by the experimenter, and all subjects stopped working with that situation. Next, a paper was handed out to subjects, so that they could write down their behavior in the previous situation, and their reasoning behind their actions (this took approximately 3 minutes). After the paper was handed in, the experiment continued with the next situation in Optics. The procedure of working with a situation, and then answering a question about it, was carried out for all four situations. After the fourth situation, the posttest (for optics domain-specific knowledge) was administered.

6.2.5.1 Assignments

Each situation in the Optics environment was accompanied by an assignment. These assignments stated in very general terms what subjects were supposed to do: find out what happens when a lamp shines through a lens. For situations 1 and 2, equivalent assignments were used, situation 3 asked subjects to find out what happens when light shines through a plate with holes on another plate. Situation 4 asked subjects to find out what happens when light shines through more than one lens. Assignments were given on-screen at the beginning of the situation. While they worked with the Optics simulation, subjects could at any time view the assignment for a situation.

6.3 Results

In this section, results from the experiment are described. The structure of this section resembles that of the results section of the previous chapter. Re-

sults on the different prior knowledge tests are first given. Subsequently, results from the interaction with the Optics learning environment are described. Here, groups with varying amounts of prior generic and domain-specific knowledge and discovery skills are compared on a number of variables. At the end of this chapter, the findings are shortly discussed in relation to the expectations that were presented in Section 6.1.2.

6.3.1 Prior knowledge test results

Three tests were administered, a test for generic knowledge, for domain-specific knowledge, and for discovery skills. The tests for generic knowledge and discovery skills were administered as pretests only.

6.3.1.1 Generic knowledge of mathematical relations

The test for generic knowledge of mathematical relations contained 32 items. Subjects were given 30 minutes to complete the test. A reliability analysis led to the removal of two items from the test, resulting in a Cronbach's alpha of 0.81 for a test of 30 items. Results for this tests are shown in Table 6-4. Both separate results for the groups with different schooling and combined results are shown. Students in pre-scientific education (PSE) showed a higher mean test score than students in technical vocational education (TVE). The difference is statistically significant ($p<.01$). Poor results on this tests were caused both by wrong answers on test items, and by subjects not being able to complete the test in time.

Table 6-4. *Generic knowledge test results*

Generic knowledge test	**Pre-scientific education (n=29)**	**Technical vocational education (n=33)**	**Combined results (n=62)**
Number of items (analyzed)	30	30	30
Mean score	24.21	17.88	21.0
Standard deviation	3.36	3.77	3.57
Range of scores	15 – 28	11 – 25	11 – 28

6.3.1.2 Optics domain-specific knowledge

The optics knowledge pretest and posttest both consisted of 30 items. A reliability analysis revealed that because some items, especially in the pretest, showed an inconsistent answer pattern, they lowered test reliability considerately. For this reason, 6 items were discarded from the pretest and the posttest, resulting in two tests of 24 items. This resulted for the pretest in a Cronbach's alpha of 0.34, which can be considered low, and for the posttest 0.56, which is moderate. Table 6-5 shows results for the pretest and the posttest.

Table 6-5. *Optics domain-specific knowledge test results, pretest and posttest*

Optics pretest	**Pre-scientific education (n=28)**	**Technical vocational education (n=33)**	**Combined results (n=61)**
Number of items (analyzed)	24	24	24
Mean score	13.79	12.09	12.94
Standard deviation	2.99	2.44	2.72
Range of scores	7 – 21	7 – 17	7 – 21
Optics posttest			
Number of items (analyzed)	24	24	24
Mean score	12.18	10.21	11.20
Standard deviation	3.30	2.82	3.06
Range of scores	7 – 19	5 – 17	5 – 19

As can be seen from Table 6-5, there was a drop in performance from the domain-specific pretest to the posttest. Results on the posttest were significantly lower than on the pretest ($p<.01$). Although the range of scores stays the same, an overall decline in scores was observed.

6.3.1.3 Discovery skills

The Peter paper-and-pencil test was administered to students after they had indicated that they knew the effect of the five variables in the Peter task.[3]

[3] In some cases, a subject stated to know all the rules, even after they had performed less than 10 experiments. In cases where this occurred, the subject was asked to per-

The test consisted of six items: five items dealt with the five variables in the Peter task, the sixth item gave subjects a chance to write down any special effects they had noticed in the task. While the test was administered, students were allowed to examine the experiments they had carried out and any notes they had made on paper, but they were not allowed to do any new experiments. The test took approximately 10 minutes to administer. Test items required students to state the effect of each of the 5 variables in the Peter task on the outcome. Based on the answers, students were classified into one of four different levels. Students were classified as level 1 when they did not find correct main effects, and no interaction effect (low performance); as level 2 when they did find correct main effects but no interaction effect; as level 3 when they found correct main effects and the correct interaction effect, but also one or more incorrect interaction effect; finally, as level 4 when they found correct main effects and the correct interaction effect, and no other effects (high performance). The category can be seen as a measure of performance in the Peter task, with level 1 being the lowest, and level 4 the highest level of performance. In Table 6-6, results from the test are shown.

Table 6-6. *Discovery skills test results*

Discovery skills test **('Peter test')**	**Pre-scientific education (n = 29)**	**Technical vocational education (n = 33)**
Mean score	2.83	1.64
Standard deviation	0.97	0.42
Range of classifications	1-4	1-2

As can be seen from the table, PSE students showed a larger variety of scores than TVE students. This is due to the TVE students' inability to discovery the correct interaction effect: about half of the PSE students were able to detect it, whereas *none* of the TVE students reported finding it. The difference in mean test score between the groups was statistically significant ($p<.01$). It is possible that this result was caused by differences in generic knowledge

form some more experiments. However, it did happen that subjects were so sure of themselves, that they refused to perform any more experiments. In that case, the test was administered.

about mathematical relations. In that case, we would expect a high correlation between the Peter test score and the generic knowledge test score. Combining the PSE and TVE groups gives a correlation of 0.51, which can be considered moderate ($p<.01$). To see if the correlation between generic knowledge and discovery skills is really related to a skill measure and not to a general measure of activity, the number of experiments performed in the Peter task was partialled out of the correlation. After partialling the number of experiments done in the Peter task, the correlation drops to 0.46, which is still statistically significant ($p<.01$).

Quality of experiments. In this section, subjects with poor discovery skills (classified into group 1, $n = 11$) are compared with subjects with high discovery skills (classified into group 4, $n = 7$). One way to study the quality of experiments that were carried out by students, is to look at the number of variables that were changed from experiment to experiment. Subjects could either change nothing, one, two, three, four, or all five variables from one experiment to another. For all subjects in both knowledge groups, the occurrence of different variable changes was counted, and this number was compared with the total number by computing the percentage. In Figure 6-3, the resulting percentages are shown. As can be seen from the figure, subjects showed a preference for the 'change one variable' strategy. However, subjects with poor discovery skills showed a lower percentage than subjects with high discovery skills. This difference is partly caused by the use of the (not very useful) 'change nothing', 'change four variables', and 'change five variables' strategy that subjects with poor discovery skills make more use of than subjects with high discovery skills.

Figure 6-3. *Percentage of different variable change strategies in the Peter task, poor and high discovery skills*

There was a small difference in the range of experimental outcomes subjects generated between students with poor and students with high discovery skills. Fifty-five percent of the students with poor discovery skills created all five possible outcomes in the Peter task, whereas this was done by 71 percent of the students with high discovery skills. This result indicates, that students with poor discovery skills explored less of the experiment space in the Peter task than students with high discovery skills.

6.3.2 Interaction results

In this section, the interaction with the Optics computer simulation is described. Results for the influence of generic mathematical knowledge, optics domain-specific knowledge, and discovery skills are described separately. In Section 6.3.2.3, results on interaction with the Peter task is also described. Results for the questions that were answered in between situations are also given.

6.3.2.1 Generic knowledge

Based on the results on the knowledge test for generic knowledge of mathematical relations, two groups were formed: a group with poor generic knowledge (n = 11), consisting of persons who had scored lower than one standard deviation below the mean score, and a group with high generic knowledge (n = 11), consisting of persons who had scored more than one standard deviation above the mean score. These knowledge groups were compared on a number of measures, which are given below.

Overall activity. The main operations that were carried out by subjects in the computer simulation, were Add, Delete, and Move. Other important operations were Read assignment, Make note, and Check values. To get an indication of the activity level of subjects, from situation to situation, a mean activity score was computed from the number of operations performed in each situation. Figure 6-4 shows a graph of the mean number of operations in each situation, for both the poor and high generic knowledge groups.

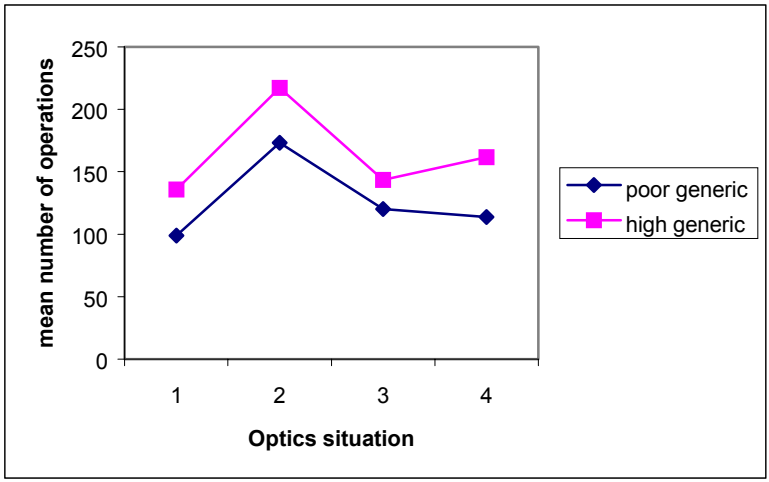


Figure 6-4. *Mean number of operations in Optics for poor and high generic knowledge groups*

The operations that were used to compute the mean values in the figure, were Add, Delete, Move, Read assignment, Make note, and Check values. On the X-axis are the situations that were described in Section 6.2.1. As can be seen from the figure, subjects with poor generic knowledge performed less operations on the average, in all four situations. In all situations, most of the operations that were performed were Add, Delete, or Move operations. In Figure 6-5, the occurrence of each of these three operation is compared. The analysis was performed for both the poor and high generic knowledge groups. As can be seen from the figure, both groups show a similar pattern over the situations. Adding objects to the working area gets a larger share over situations in relation to the other operations. From Figure 6-5, differences between subjects with poor and high generic knowledge are not immediately clear.

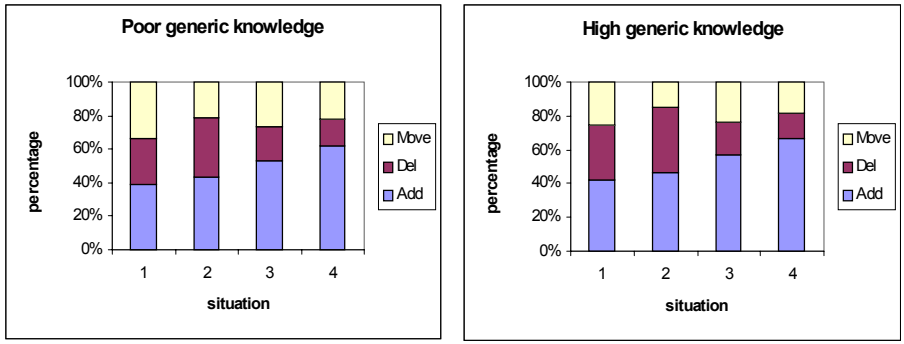


Figure 6-5. *Relative occurrence of Add, Delete, and Move operations for poor and high generic knowledge*

If subjects with high generic knowledge perform more operations, which ones do they focus on? For this, situation 2 can serve as an illustration. Figure 6-6 shows, for situation 2, a direct comparison between the poor and high generic knowledge groups, for the Add, Delete, and Move operations. As can be seen from the figure, subjects with high generic knowledge tended to add and delete more objects than subjects with poor generic knowledge. Both groups perform a similar amount of Move operations. This result is in line with the results of the previous study. Subjects with high generic knowledge create different settings in Optics by adding and deleting objects, instead of experimenting with one particular setting. However, the differences between the groups are not statistically significant, with p=.09 for the Add operation, and p=0.11 for the Delete operation.

In situation 2, the mean number of operations was higher than in the other situations. An explanation for this may be that situation 2 was viewed by subjects as a continuation of situation 1, but with more available operations and objects. It appears that students, when confronted with a new situation (as is the case in situations 1, 3, and 4), take some time to get used to the situation, and, as a consequence, are less active in these situations.

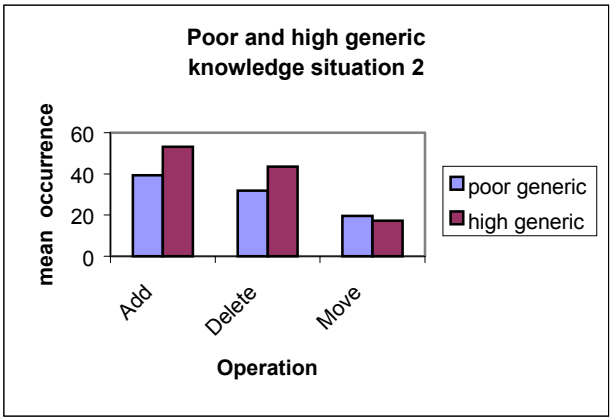


Figure 6-6. *Comparison on Add, Delete, and Move operations in situation 2 for poor and high generic knowledge*

In the Optics 97 study, differences between groups were examined as they occurred over an extended period of time. Here, we have developed a slightly different approach to analyzing differences over time between the knowledge groups. As in the first experiment, for each minute spent in the learning environment the mean number of Add, Delete, and Move operations was computed. However, each situation was then divided into three equal parts, a beginning, middle, and end part (each of these parts representing approximately five minutes of experiment time). Figure 6-7 shows these differences between the high and poor generic knowledge groups, for each period (beginning, middle, and end), and for each situation. Each bar in the graph represents a difference (for a period of time within a situation) between the mean test score of the high generic knowledge group and the mean score of the poor generic knowledge group. Negative values indicate that the poor knowledge group was more active at a certain moment than the high knowledge group.

The results that are shown in Figure 6-7 make clear that differences between the groups with poor and high generic knowledge could be observed both over situations, and within the situations themselves. Overall, subjects with high generic knowledge are more active than subjects with poor generic knowledge. Within the situations, it can be seen that there are only two instances out of twelve in which this pattern was reversed, at the beginning of situation 3 and at the end of situation 4. Subjects with high generic knowledge are a lot more active than subjects with poor generic knowledge in

situation 2, compared with the other situations, but it can be seen that this difference only exists in the middle and the end part of situations.

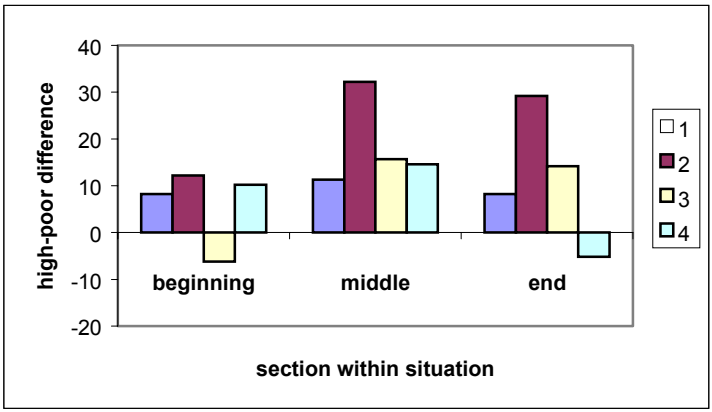


Figure 6-7. *Difference in activity between poor and high generic knowledge groups, in the beginning, middle, and end of situations*

6.3.2.2 Optics knowledge

Based on the results on the optics domain-knowledge pretest, two groups were created: a group with poor domain-specific knowledge (n = 8), which consisted of subjects who scored lower than one standard deviation below the mean score, and a group with high optics knowledge (n = 8), which consisted of subjects who scored more than one standard deviation above the mean. As in the case of generic knowledge, the two groups could be compared on a number of measures, most important of which is the general level of activity in the Optics simulation.

In Figure 6-8, the mean activity level as measured by the number of Add, Delete, and Move operations is shown in all four situations. As can be seen from the figure, only in situation 3 a difference between subjects with high and poor domain-specific knowledge is observed. In situation 3, subjects with high domain-specific knowledge are less active than subjects with poor domain-specific knowledge. The reason for this is not clear. Situation 3 provided a break from the first two situations in that something new (the big lamp and the screens) was introduced in the simulation.

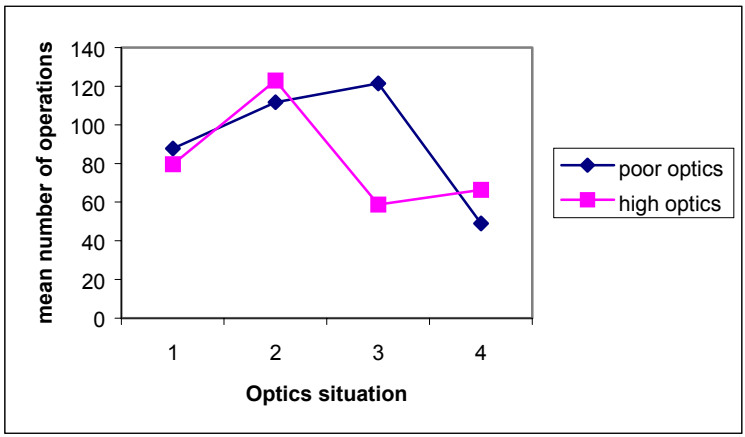


Figure 6-8. *Mean number of operations in Optics for poor and high domain-specific knowledge groups*

In Figure 6-9, the relative occurrence of add, move, and delete operations is shown for each situation. When the knowledge groups are compared, it can be seen that subjects with high domain-specific knowledge perform less move operations in situation 2, and more move operations in situation 3 compared to subjects with poor domain-specific knowledge.

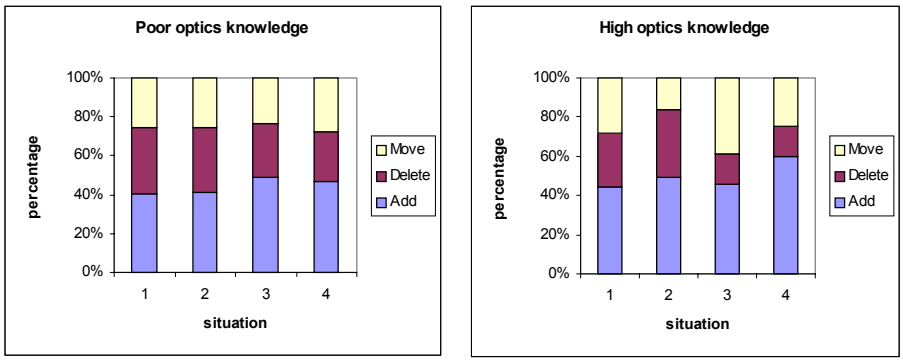


Figure 6-9. *Relative occurrence of Add, Delete, and Move operations, poor and high domain-specific knowledge*

Combined with the finding that subjects with high domain-specific knowledge performed less operations in general in situation 3, it can be concluded that these subjects mostly experimented by moving around objects. From the figure, it can also be concluded that subjects with high domain-specific knowledge showed a larger variation in the basic operations they performed.

6.3.2.3 Discovery skills

Based on the results of the Peter test, subjects were classified into four different groups. In Section 6.2.3.3, the classification procedure was explained. In this section, the interaction behavior of subjects who were classified as group 1 (subjects with poor discovery skills, $n = 11$) and group 4 (subjects with good discovery skills, $n = 7$) are compared with each other. Only results for subjects for which a complete data set was available (test data and interaction data for both experiment sessions), were considered in the analysis. Results with respect to interaction with the Peter computer task are given first. These are followed by the results with respect to interaction with the Optics learning environment.

Peter task interaction

All subjects stated that they were finished experimenting with the Peter task, before the time limit of 30 minutes was reached. A comparison of the number of experiments performed by students from pre-scientific education (PSE) and technical vocational education (TVE), showed that the TVE students carried out less experiments than PSE students. The mean number of experiments performed by PSE students was 35.2 (SD = 15.1), whereas TVE students performed an average of 22.7 (SD = 10.9) experiments. This difference was statistically significant ($p<.01$). Subjects who scored low on the Peter test (group 1) performed a mean number of 18.1 (SD = 8.4) experiments, ranging from a minimum of 4 experiments to 36. Subjects who scored high on the Peter test (group 4) performed a mean number of 34.9 (SD = 16.1) experiments, with a minimum of 13 experiments, and a maximum of 67. The difference is statistically significant ($p<.01$). Removing the subject who performed only four experiments from the analysis slightly changed these val-

ues; the difference in the number of experiments that were performed stayed statistically significant, however.

Figure 6-10. *Relation between Peter test scores and mean number of experiments in Peter task*

The correlation between Peter test score and the number of experiments performed in the Peter task, is 0.40. In Figure 6-10, this relation is depicted. For each score, the mean number of experiments were computed. As can be seen in the figure, there appears to be a general relation between the amount of activity in the Peter task, and performance on the Peter test. The figure also shows that interaction effects can only be found when a large number of experiments is performed.

Optics simulation interaction

As has been explained at the beginning of this section, based on their performance in the Peter task subjects were classified into four groups. Two groups were analyzed: a group with poor discovery skills (n = 11), and a group with high discovery skills (n = 7). In Figure 6-11, the mean activity level as measured by the number of Add, Delete, and Move operations is shown. As can be seen from the figure, subjects with high discovery skills perform more operations in the first two situations than subjects with poor discovery skills. The difference between the groups vanishes in the third and fourth situation. When this finding is compared with the data in Figure 6-10,

it can be concluded that subjects with high discovery skills are not only more active in the Peter task, but also (in the simple situations) in the Optics simulation.

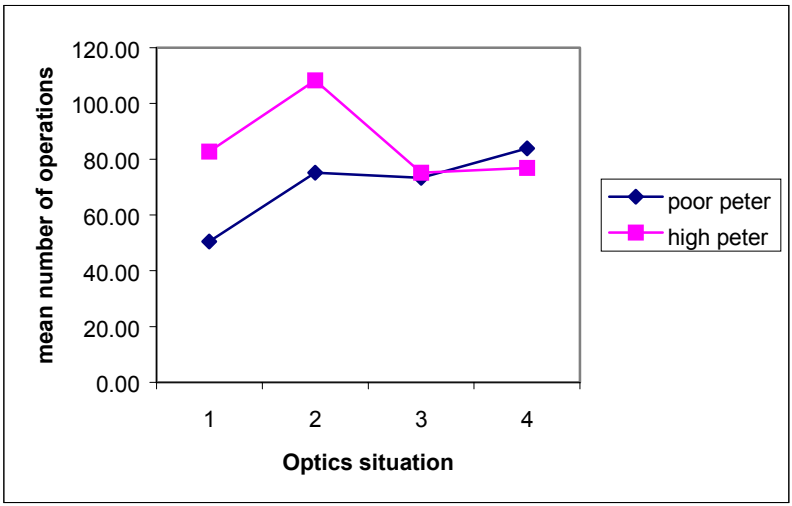


Figure 6-11. *Mean number of operations in Optics for poor and high discovery skills groups*

Figure 6-12 shows for the high and poor discovery skills groups the relative occurence of the basic operations, compared over situations. It can be seen that differences between the groups are small for all situations. In the first two situations, subjects with high discovery skills are more active in adding objects to and deleting objects from the working area.

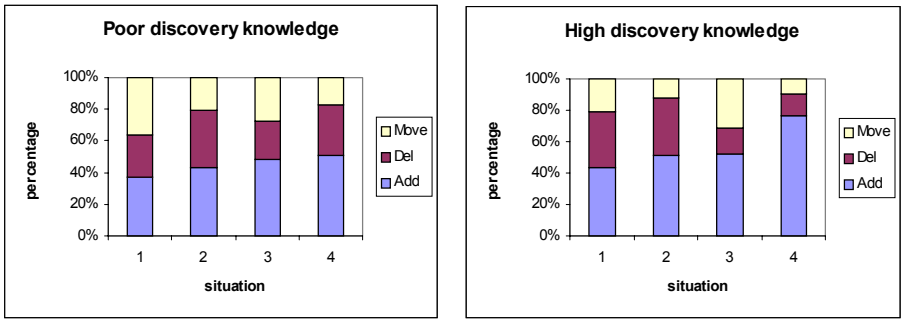


Figure 6-12. *Relative occurrence of Add, Delete, and Move operations, poor and high discovery skills*

6.3.2.4 Question answered after each situation

After working with each situation in the Optics learning environment, subjects answered a question. This question was meant to get insight into subjects' reasons for their behavior in the previous situation. Answers to the question after each situation were checked for a number of characteristics. These characteristics were:

1) Overview of operations
2) Use of geometrical optics concepts
3) Use of qualitative relations
4) Use of quantitative relations

In the answers that were given, all four characteristics were used. However, most subjects limited their report to comments that fall in the first category. Other examples of reasons that were given for behavior in the simulation are: 'I tried to follow the assignment', 'This seemed to me to be the easiest way to work', 'I wanted to get to know the computer program', and 'I acted in this way because I thought it was fun'. It was also found that in situation 3 (the situation that included the projection of an object on a screen), many subjects used an 'engineering' approach. Instead of focusing on the relation between different variables in the simulation, they reported to have only tried to position all objects in such a way that a sharp image would be the result.

6.4 Discussion

In this chapter, we have described the Optics 99 experiment. This experiment was a follow-up study to the Optics 97 experiment that was described in Chapter 5. The goal of the present study was to analyze the relation between discovery learning processes and domain-specific knowledge, generic knowledge, and discovery skills. To study discovery learning processes, a version of the Optics computer simulation (described in Chapter 3) was used. In the experiment two groups of subjects participated: students who were in technical vocational education, and students who were in prescientific education. Two main research questions were addressed in the study: how do prior domain-specific knowledge, generic mathematical knowledge, and discovery skills influence discovery learning processes, and what differences in scientific discovery learning behavior can be found when different configurations of the Optics learning environment are compared with each other. The first question was addressed by testing knowledge and skills before students started working with the Optics learning environment. Generic knowledge and domain-specific knowledge were tested by paper-and-pencil tasks, discovery skills were tested for by using a combination of the Peter task with a paper-and-pencil test. The second question was addressed by making use of different situations in the Optics computer simulation. A total of four situations was used; the situations ranged from relatively simple to complex.

With respect to results on the knowledge tests, an important finding was that students' domain-specific knowledge appeared to deteriorate from pretest to posttest. Because they only worked for a relatively short time in the Optics simulation, it was not expected that the amount of domain-specific knowledge would increase from pretest to posttest. However, the significant decrease in performance was not expected. It may be that the configuration of the Optics simulation in this experiment was confusing to students. Students worked with each situation for only a (relatively) short amount of time, after which they had to adjust to the new situation that was introduced. It may be, that each time students had to overcome a stage of bewilderment, in which they tried to get familiar with the topic of the simulation. Students' initial confusion may be caused by 'cognitive conflict' (Hewson & Hewson, 1984), which means that they went through a phase in which they tried to connect the Optics simulation with their knowledge about geometrical optics. It may be that, because of the short time spent in each situation in

Optics, many students did not pass the stage of cognitive conflict, which may have caused the deterioration in their domain-specific knowledge.

The test for generic knowledge that was used in the experiment focused on knowledge about mathematical relations. The results showed that the redesign of the test that was used in the Optics 97 has been fruitful in the sense that Cronbach's alpha was high. Also, the more varied population that was used in the study increased the range of test scores, which helped in distinguishing between students with poor and high generic knowledge.

The Peter test showed that it is possible to combine a computer task with a paper-and-pencil test in measuring discovery skills. It was shown that students who perform more experiments in the computer task, and who use a different strategy in experimenting (with a higher percentage of experiments in which only one variable is changed) perform better on the paper-and-pencil test.

The analysis of interaction with the Optics simulation showed a number differences in discovery learning behavior between students with high and poor knowledge and skills. The level of activity in four situations in the Optics simulation was compared.

For generic knowledge, it was shown that students with high generic knowledge were on the average more active than students with poor generic knowledge. When we looked at the type of operations that were performed by these students, it could be seen that they performed more Add and Del operations. This result was similar to the result found in the Optics 97 experiment. From this result it appears that students with high generic knowledge tend to focus on creating different situations in the Optics simulation. As was shown, the level of generic knowledge correlated moderately with the results on the Peter task, which indicates that the students with high generic knowledge are more active in experimenting in general. This finding can be related to the extended SDDS framework. Students with high generic knowledge have more knowledge of the type of mathematical relations that exist than students with poor generic knowledge. This means that they have a larger effective learner space, which gives them more freedom in performing experiments in the Optics simulation.

The results for the groups with poor and high domain-specific knowledge did not show many differences. In the third situation in the Optics simulation, students with high domain-specific knowledge performed less operations than students with poor domain-specific knowledge. It may be that the students with high domain-specific knowledge were more familiar with the particular configuration of the Optics simulation that was used in the third

situation (the situation in which the plate with an L-shaped figure served as an object that was projected on a screen). This may have caused these students to perform only a few experiments in this situation.

To examine the relation between discovery skills and discovery learning with the Optics simulation, students who scored in the lowest class (no effects found) on the Peter test were compared with students who scored in the highest class (all effects found). A difference in activity level in the Optics simulation was found for the first two situations. Students who scored high on discovery skills were more active in these situations than students who had poor discovery skills. The first two situations in the Optics simulation were more limited in scope and in the number of available options than the third and fourth situation. It may be, that in the simple situations that were presented first, students' discovery skills were important in guiding discovery learning. The third situation was different from the first two situations and more complex. Here, domain-specific knowledge 'kicked in' and influenced the learning process. No differences were found for the fourth situation. In this situation, the complex situation of using more than two lenses simultaneously was presented. It may be that without the availability of support, this situation was too complex for all subjects, which made differences between groups with a different level of knowledge and skills disappear. In conclusion, it can be stated that generic knowledge asserts an overall influence on scientific discovery learning, discovery skills make a difference in simple situations, and domain-specific knowledge makes a difference in situations that call for specific pieces of (domain-specific) knowledge.

Similar to the Optics 97 experiment only minimal support was provided to the students. Only a general assignment was given. Students in the Vocational Education group frequently reported problems with the interpretation of the assignment, because it did not state an exact procedure that they had to follow. The four situations that were used in the Optics simulation in this experiment, especially the third and fourth situation, each focused on separate topics in geometrical optics. The short time spent in each situation, the complexity of the simulation, and the lack of support that was given to students may in combination have caused the decrease in performance. Differences between students in pre-scientific education (PSE) and students in technical vocational education (TVE) also showed up when the correlation between the domain-specific pretest and posttest was computed. The corre-

lation for the PSE group was 0.45, and for the TVE group -.04. This result indicates that students in the TVE group had more trouble in making use of the knowledge they already had.

It becomes clear from the results of the current study and the Optics 97 study, that the Optics simulation can easily become too difficult for students. Also, the lack of support makes the simulation unsupportive to promote knowledge gain. Most students did not possess sufficient domain-specific prior knowledge to grasp the meaning of all four situations in the simulation. They may have experienced parts of the simulation as confusing, which in turn may have had an effect on their motivation to learn. In the Optics 2000 study, which is described in Chapter 7, the question if discovery learning in the Optics simulation can be supported is explored. It is assumed in this study that one problem that students may have had with the Optics simulation, which caused ineffective learning, was that they had trouble accessing their background knowledge. It is assumed that making knowledge available to students while they work with the learning environment will help them remember background knowledge, and support discovery learning behavior.

7

Optics 2000: Using domain-specific and generic knowledge to support discovery in Optics

7.1 Introduction

In the introduction to Chapter 6, it was argued that, in studies of scientific discovery learning processes, prior knowledge is only rarely singled out as a single influencing factor. The experiment that was described in that chapter used three separate tests to measure prior domain-specific knowledge, generic knowledge, and discovery skills. The experiment involved students with two different knowledge backgrounds, who worked with four situations of varying complexity in the Optics computer simulation. The design of the Optics 99 experiment was in many ways similar to the Optics 97 experiment. However, a number of changes were made to the prior knowledge tests and to the configuration of the Optics simulation. The modified version of the domain-specific prior knowledge test was clearly an improvement over the test for the Optics 97 experiment. This may have been in part the result of removing the demand for trading speed and accuracy from the test. Still, in the experiment no improvement for domain-specific knowledge from pretest to posttest was found. Students' performance even appeared to deteriorate after they had worked with the Optics simulation. It was also observed that some students were confused by the assignment that was given before each situation in Optics. Because only general assignments were used, these students may have been unsure what it was they were supposed to do in the computer simulation, when they were given the freedom to do anything at all in it, and not having to follow some prescribed procedure.

Although these negative effects of using the Optics simulation were found, this has not been a detriment to the primary goal of the research studies. The goal has not been to study effective and optimal types of discovery learning, but rather to specifically study the role of different types of prior knowledge on unsupported discovery learning. Learner support was kept to a minimum to decrease interference with prior domain-specific and generic knowledge. The only type of support that was used in the case of the Optics computer simulation was structuring the learning environment so that learning would progress from a (relatively) simple level to a (relatively)

complex level. It can be said that lack of prior domain-specific knowledge, combined with minimal support, made operating the Optics computer simulation a demanding exercise. In the Optics 97 and Optics 99 studies, prior domain-specific knowledge was measured. In the present study, we will try to manipulate knowledge about optics. In the Optics 97 study, we already attempted to manipulate knowledge about the Bubbles simulation. We did not find an effect of providing students with prior knowledge about the Bubbles domain on the discovery learning process and learning result. It may have been that the prior domain-specific knowledge that was provided about Bubbles was not used by students. It is hypothesized, that when knowledge is available *during* discovery learning (based on a 'just in time'-principle), it can support students in their discovery learning process.

The goal of the current study is to investigate whether the Optics simulation can be made more effective, and performance afterwards improved, by supporting students during discovery learning. This support is implemented by making domain-specific and generic knowledge available to students while they are working with the learning environment.

Although discovery learning can be supported in a variety of ways, support is not always effective. Support can even have the effect of degrading, or hampering, performance (Pieters & Van der Meij, 1994). Problems can occur when either too little or too much support is given. For example, Van Joolingen (1993) tried to support discovery learning by providing students with a structured hypothesis scratchpad. This scratchpad could be used to make hypotheses explicit. It was shown that students who had a structured hypothesis scratchpad at their disposal conducted fewer experiments and stated fewer hypotheses than students who did not. The use of the structured hypothesis scratchpad did not help students in conducting experiments and generating hypotheses. Van Joolingen argues that the structured hypothesis scratchpad provided an extra task to learners, which they may have found difficult to cope with given the short time for the experiment. A similar conclusion was drawn by Shute (1991), who argues that tools that aim at supporting discovery learning can interfere with the learning process itself. This interference disrupts compilation of knowledge and can have an adverse effect on learning. Njoo (1994) states that the cognitive load theory of Sweller (1994) also offers an explanation for the negative effects that an extra task has: the redundant information that the task contains distracts students, and causes wasteful mental processing to occur.

Not only the presence of supporting material such as a hypothesis scratchpad does not always have beneficiary learning effects: the structure of the learning environment itself can pose problems as well. Van der Hulst (1996)

argues that even when students are provided with a logical sequence, there is no guarantee for good performance. According to Van der Hulst, prior knowledge can directly influence the effectiveness of using a particular sequence. Enabling learners to connect new knowledge to previously acquired knowledge is seen by her as an important type of support. Also, it has been shown by Swaak, Van Joolingen, and De Jong (1996) that directive support within one level of exploration can be more effective than providing support through a number of levels (for example, by structuring a learning environment using model progression). A study by Stark, Renkl, Gruber, and Mandl (1998), however, replicated earlier results that showed that intermediate experts (in the domain of economics) did not perform better on a multitude of tests than novices after learning with a computer-based simulation. This result was not caused by a difference in motivation, but more by the inability of the 'experts' to apply their knowledge to new situations. Instead, knowledge from the simulation environment appeared to interfere with already existing knowledge. Stark et al. warn that 'The presence of abstract domain-specific knowledge by no means guarantees successful knowledge application and transfer (p. 402)'.

These results prompt us to be careful in selecting the type of support that students are provided with. The progressive structure that was used in the Optics 99 experiment (and which was not explained to students before the experiment started) may have confused students. In addition, because students had poor domain-specific knowledge, they may have had trouble in understanding the more complex situations in the previous experiment, being unable to connect the situations on the screen with their background knowledge.

7.1.1 Study design and objective

The current study was designed to test the benefits of combining a limited configuration of the Optics simulation with active support for discovery learning. More specifically, the effect of having knowledge available, both domain-specific and generic, on the discovery learning process and performance of students in the Optics learning environment was analyzed (the 'help'-condition). A control group was used, which consisted of students who had no extra information available (the 'no-help'-condition). As in the previous experiment, prior domain-specific knowledge, generic knowledge, and discovery skills were measured. The aim of both measuring and ma-

nipulating knowledge is to compare the relation between prior knowledge and discovery learning behavior for both conditions. In addition, the discovery learning process and progress from pretest to posttest can be compared for the two conditions.

In Section 7.2.1, the configuration of the Optics simulation that was used in the present study is described. In the previous chapter, it was tentatively stated that students may especially have trouble in working with a computer simulation in the first minutes, and may be confused by the simulation's content. Only after overcoming this stage they can start gathering data from the simulation. In the Optics 99 study, students only worked for a short time in each of the four situations that were used. It may be that the time that was used for each situation was too short for students to get adequately used to the simulation. In that case, the use of model progression may have had a detrimental effect on students' experimenting and subsequent learning result. For the current experiment, it was decided to make use of only one situation in the Optics simulation, in which students would work for a longer time. This means that the current experiment did not make use of model progression through a number of situations. The situation that was created in the Optics simulation focused on one aspect of geometrical optics, and was aimed at being neither too simple (for example, a situation with only one lens and one type of lamp), nor too complex (for example, projecting light through two lenses simultaneously). Progression through a number of situations was not used in this experiment. Instead, a number of students received support while they worked with the simulation through the availability of *knowledge tips*. These tips consisted of both domain-specific information (for example, short explanations of concepts), and generic information on how specific instances in the Optics simulation could be used to show certain effects. The content of the tips is explained in more detail in Section 7.2.5.2, which describes the conditions used in this experiment. The tips were not all available at once as soon as work with the Optics simulation started. Instead, tips became available one by one as work with the learning environment progressed in time. Students were not distracted by the availability of a new tip as soon as it became available. They could activate the tip at any moment they preferred, or even not at all. This was done to prevent students being distracted from their work. The tips were not made available at once, so that students who became stuck could use an available tip to get fresh ideas for experimentation.

In summary, this study aimed at improving discovery learning in the Optics computer simulation. Discovery learning behavior and knowledge gain from pretest to posttest was compared for a group of students who had ac-

cess to domain-specific and generic knowledge while they worked with the learning environment, and a group who did not receive these tips. The Optics simulation was configured for this experiment to focus on one topic of geometrical optics. No further progression through a set of situations was used. The study made use of two classes from two schools for vocational education, one situated in Enschede and one in Hengelo. The experiment was performed at the faculty for Educational Science and Technology in April and May 2000.

7.1.2 Research questions and expectations

Since the current experiment sets out to further investigate the influence of prior domain-specific knowledge and generic knowledge on scientific discovery learning in the Optics computer simulation, the research questions are similar to the questions asked about the previous study (see the beginning of Section 6.1.2). However, the setup of this experiment allows us to add another specific research question: what differences can be found between students who are not supported in their learning behavior, and students who have support available in the shape of domain-specific and generic tips during discovery learning? We expect differences between the 'no-help condition' and the 'help condition'. These differences should show up as differences in scientific discovery learning processes and in differences in performance on the domain-specific knowledge test. We expect that students in the help condition will benefit from having knowledge available, which will result in more progress from pretest to posttest on the domain-specific knowledge test. Also, in a study by Hasselerharm and Leemkuil (1990) it was found that students with little prior domain-specific knowledge made less often use of an optional support tool than students with high ability. Therefore, in the current study we expect students in the help condition with poor domain-specific prior knowledge make less use the knowledge tips than students with high domain-specific knowledge.

7.2 Method

This section describes the experimental method that was used in this study. An overview is given of the particular configuration of the Optics simulation, and of the tests that were used. Also, in this section a description of the

type of knowledge support that was provided to part of the subject group is given.

7.2.1 Optics configuration

The configuration of the Optics simulation that was designed for this experiment was more limited in scope than the configuration that was used in the Optics 97 and Optics 99 experiments. The simulation focused on one particular aspect of geometrical optics: the projection of an object on a screen. The underlying formulas included in this configuration are the lens formula and the magnification formula. The simulation also allowed students to make the position and size of a virtual image visible.

In Figure 7-1, an example from the Optics interface as it was used in the current experiment is shown. The figure shows a complicated, but not untypical, situation that might be created by a student. On the left, a large lamp is sending light in all directions. Light passes through a plate with holes (shaped like the letter L). Divergent light rays that originate from the holes in the plate pass through the lens (marked 'C'). Because the plate is close to the lens (within the focal length of the lens), no image is projected on the screen at the right. Instead, a virtual image is formed. This virtual image is made visible by the use of the 'eye'-tool (which is placed at the right). The virtual image is shown with a blue circle around it (to mark its 'virtuality'). It is possible to measure properties of the virtual image, just as if it were a normal projected image. In Figure 7-1, the top row shows on the left three different lenses that were available to experiment with (the lenses had different focal distances). Next to the large lamp (which could not be moved around, but only positioned on the left of the working area), a lamp with one lightbeam, a lamp with three divergent lightbeams, and a lamp with three parallel lightbeams were available. It was possible to measure horizontal and vertical distances, and to measure the angle with which a lightbeam entered or exited the lens. Lightbeams could also be rotated any number of degrees. As an extra aid, lightbeams could be extended in any direction by adding one or more help lines. Subjects were allowed to make notes while they worked with the simulation. For this they were handed writing paper; in this experiment subjects did not have to enter notes directly in the learning environment itself.

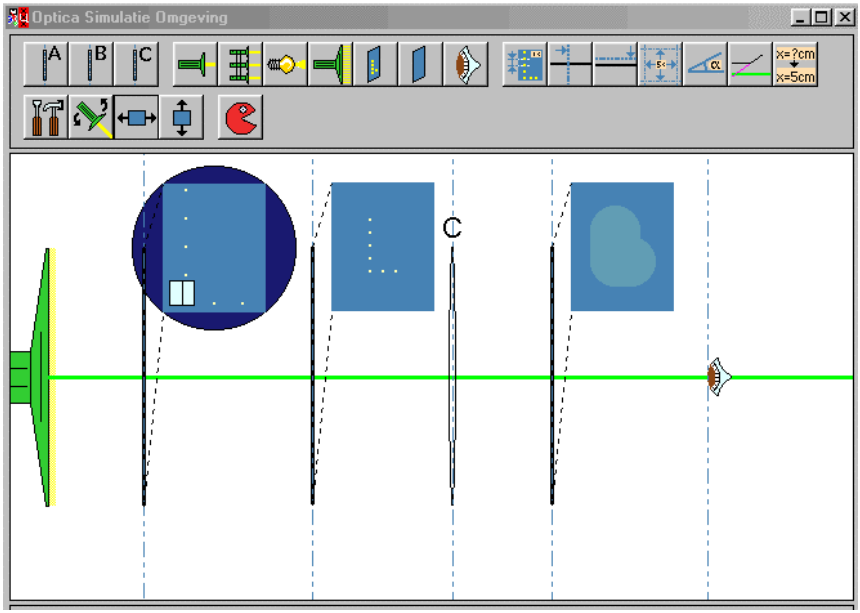


Figure 7-1. *Example interface of the configuration of Optics in the current experiment. Explanation given in text*

7.2.2 Subjects

The subjects who participated in this experiment were 32 students who were following technical vocational education (mean age approximately 19 years). The subject group was divided over two school classes (both in the same year of study). We chose to use students at the level of vocational education, because it was observed in the Optics 99 experiment that students at this level of education formed a more heterogeneous group than students at the level of pre-scientific education. All subjects were taking physics as part of their curriculum. No subjects showed any difficulties in operating the computer, so no special instruction on how to operate the mouse and computer was needed. All subjects received a financial compensation of €23 for their participation in both experimental sessions.

7.2.3 Prior knowledge tests

Prior to working with the Optics learning environment, three tests were administered: a test for domain-specific knowledge about geometrical optics, a test for generic knowledge about mathematical relations, and a test for experimentation skills in the context of a discovery task. While working with the learning environment, all actions subjects performed were registered. Finally, a posttest on optics domain knowledge was administered.

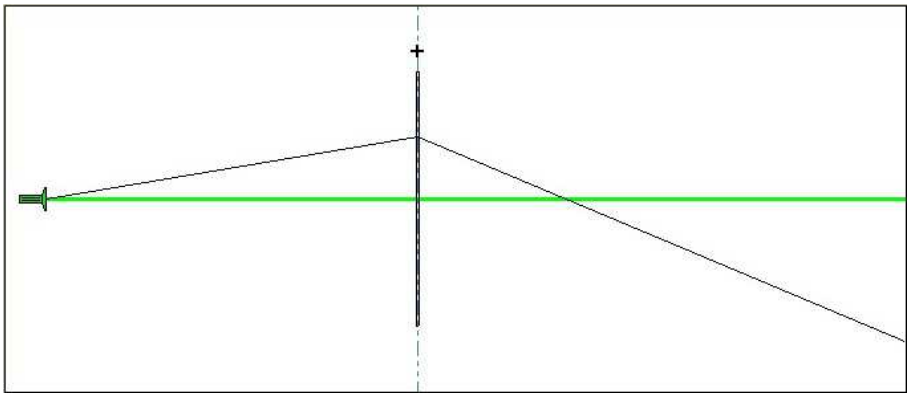
7.2.3.1 Generic knowledge test

The test for generic knowledge about mathematical relations that was used in the current experiment, was, apart from some minor corrections, the same test as the one used in the Optics 99 experiment (described in the previous chapter). The paper-and-pencil test consisted of 32 items. Of these, 29 were four-answer multiple-choice questions. The other items required a short answer, such as a formula. Example items for the generic knowledge test can be found in Chapter 5.

7.2.3.2 Domain-specific knowledge test

The test for domain-specific knowledge about geometrical optics was newly created for the current experiment. The test was based in part on the knowledge test that was used in the Optics 99 study, but consisted mostly of new questions, to fit the configuration of Optics that was used in this experiment. The test consisted of 30 items, and was administered, unlike the tests that were used in the Optics 97 and Optics 99 experiments, as a paper-and-pencil test. A paper-and-pencil test was used in favor of a computer-based test, because this would provide a more familiar type of test to students. The test also made it possible for students to review their answers and to make corrections. In Figure 7-2, two example items from the test for domain-specific knowledge are shown (for the figure, the items were translated from Dutch). The test contained a few items that asked for conceptual knowledge about geometrical optics. For example, two items specifically asked for recognition of the lens formula and the magnification formula. All in all, the test used ten pictures of possible situations in the Optics environment, with an average of three multiple-choice questions for each of the pictures.

figure 1



1. See Figure 1. When the lens is moved a little bit to the left, where will the light beam to the right of the lens strike the base line?

- ❑ Closer to the lens
- ❑ At the same distance to the lens
- ❑ Further away from the lens

2. See Figure 1. When the light beam is aimed a little bit more up, where will the light beam to the right of the lens strike the base line?

- ❑ Closer to the lens
- ❑ At the same distance to the lens
- ❑ Further away from the lens

Figure 7-2. *Example items from the optics domain knowledge test*

7.2.3.3 Discovery skills test

To measure discovery skills, the Peter task was used. The setup of this combined task and test was explained in detail in Chapter 6 (Section 6.3.2.3). For this experiment, no changes were made to the task setup. The maximum time subjects worked with the computer task was 30 minutes. After that time, a paper-and-pencil test was administered, which consisted of 6 short questions. While answering the test, students were allowed to examine the experiments they had carried out and any notes they had made on paper,

but they were not allowed to do any new experiments. The test took approximately 10 minutes to administer.

7.2.4 Interaction registration

Similar to the procedures followed in the previous experiments, all operations that subjects performed in the Optics learning environment were registered by the computer the subject was working on. The operations, along with time data, were stored on disk. The data were afterwards converted to SPSS-readable data for analysis.

7.2.5 Procedure

The experimental procedure that was followed was the same for students in both the no-help and the help conditions. Also, both classes followed exactly the same procedure. The experiment was split over two sessions of about 80 minutes each. In the first session, all the prior knowledge tests were administered. The test sequence was the same as the sequence used in the Optics 99 experiment: first the generic knowledge test, then the domain-specific knowledge test, and finally the discovery skills test were administered. At the start of the second session, subjects were given an instruction sheet which contained an explanation of buttons and tools in the Optics computer simulation. Instructions for subjects in the help condition were the same as instructions for subjects in the no-help condition, but some extra instruction on using the knowledge tips was added. After the experimenter gave a signal, all subjects started working with the Optics simulation. They worked with the simulation for approximately 50 minutes, after which the domain-specific knowledge posttest was handed out and administered in the same way as the pretest. The posttest took approximately 15 minutes to administer.

7.2.5.1 Assignment

Subjects were given one assignment in this experiment. The assignment was written on paper, and was accompanied by a list of available functions in the Optics environment. In the assignment, subjects were asked to perform experiments with the objects they had at their disposal, and to create different situations in the Optics learning environment. Subjects were also encouraged to use all the clues that would be given by the computer program, and

to make notes. The subjects were informed prior to the experiment that a knowledge test would be administered afterwards.

7.2.5.2 Conditions

At the start of the second experimental session, subjects were randomly assigned to one of two conditions. The positioning of the computers in the classroom was arranged in such a way that no persons in the same experimental condition would be seated next to each other. The following two conditions were used:

1) *No-help condition.* No extra information was given by the computer while subjects worked with the Optics simulation.
2) *Help condition*. Extra information was given by the computer, in the form of tips, while subjects worked with the Optics simulation

The no-help condition put subjects in a situation that was similar to the one used in the previous experiments. The only difference lies in the particular configuration that was used for the Optics simulation in this experiment.

Table 7-1. *Example of two knowledge tips used in the help condition*

Tip	Content
1.	a) A thin lens is a weak lens, a thick lens is a strong lens. b) Advice: Compare different lenses with each other and observe differences. c) Appropriate object to use: lamp with one lightbeam. d) Expected outcome: Lightbeams are refracted more through a thick lens than through a thin lens.
6.	a) When an object is standing within the focal distance of a lens, the projected image becomes virtual. This means that the image will lie at the left of the lens b) Advice: The 'eye' was developed for this program to show you the position and shape of the virtual image. Put the eye somewhere to the right of the lens. When the object gets near to the lens, you will see the virtual image appear. c) Appropriate object to use: plate with L-shaped holes, screen, lamp with three divergent lightbeams, distance measures, eye.

In the help condition, tips were used to make domain-specific and generic knowledge available to subjects. The tips became available one at a time, and

the interval between two tips was set at three minutes. In Table 7-1, two examples of tips that were used in the experiment are shown.[1]

In total, nine tips were used, which implies that after 27 minutes of experimenting, all tips were available to subjects The table shows the first tip, and one that was made available later on. As can be seen from the example, the tips consisted of multiple parts. The first part of each tip consisted of domain-specific information on some concept of geometrical optics. The second part of each tip consisted of general advice on how experiments could be used in the Optics simulation to observe the information that was given in the first part. Also, the relevant objects that could be used to perform appropriate experiments were listed. Finally, in some tips the expected outcome of experiments was stated. In the help condition, the bottom row of the Optics simulation contained a set of nine icons, one for each tip. When it was time for a tip to become available, its corresponding icon would change color. The tip could then be accessed by subjects by clicking on the icon. Once a tip became available, it could be accessed as many times as subjects preferred.

7.3 Results

In this section, results from the experiment are described. First, results on the three prior knowledge tests that were used are given. Subsequently, results from the interaction with the Peter discovery task and the Optics learning environment are discussed. Our main focus lies on differences between the group that received no extra help and the group that had extra help available.

7.3.1 Prior knowledge test results

The current experiment made use of three tests: a test for generic knowledge, domain-specific knowledge, and discovery skills. The domain-specific test was administered both as pretest and posttest.

[1] For the table, the tips were translated from the original Dutch.

7.3.1.1 Generic knowledge

The test for generic knowledge of mathematical relations contained 32 items. Cronbach's alpha for this test was 0.71, which is considered to be reasonable. Test scores ranged from 11 to 28, with a mean score of 21.1 (SD = 4.21).

7.3.1.2 Optics domain-specific knowledge

The optics domain-specific knowledge pretest and posttest both consisted of 30 paper-and-pencil items. A reliability test showed no large inconsistencies in the given answers. Therefore, no items were discarded from both the pretest and the posttest. This resulted for the pretest in a Cronbach's alpha of 0.51, and for the posttest an alpha of 0.66, which are considered to be moderate. The mean test score on the test was 15.3 (SD = 3.7) on the pretest, and on the posttest 14.9 (SD = 4.2). No significant differences were found between the two school classes that were used in the experiment on either the pretest or the posttest. The range of scores on the pretest was 9-24, and on the posttest 4-21. This result indicates that there was a slight general drop in domain-specific knowledge from pretest to posttest. However, the difference between the pretest and posttest score is not statistically significant. Furthermore, when the results were split for the no-help condition and help condition, a different picture emerged. These results are shown in Figure 7-3.

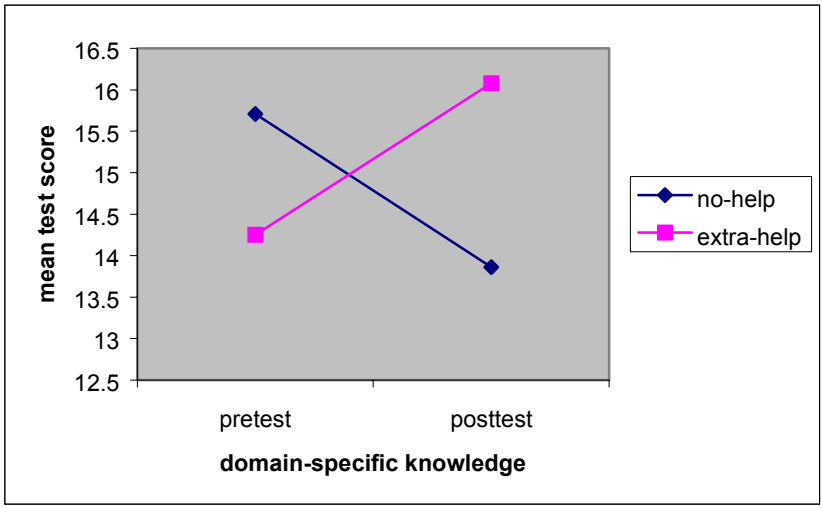


Figure 7-3. *Comparison of domain-specific knowledge test scores for the no-help and help conditions*

As can be seen from the figure, performance of students in the no-help condition deteriorated from pretest to posttest, in contrast to students in the help condition whose performance improved. A one-tailed t-test of independent samples shows that the difference between improvement from pretest to posttest in the two conditions is statistically significant; ($F_{1,24}$=3.21; p<.05).

Table 7-2. *Domain-specific knowledge, comparison of no-help condition and help condition*

Domain-specific knowledge	**Pretest score**	**Posttest score**	**Difference**	**Correlation**
No help condition	15.7 (4.5)	13.9 (4.6)	-1.86	0.52
Help condition	14.3 (3.3)	16.1 (3.7)	1.83	-0.47

Table 7-2 shows the mean test scores for the two conditions, and also shows the correlation between the domain-specific pretest and posttest. The overall correlation between pretest and posttest was 0.13, but from the table it can be seen that there was a difference between the two conditions. For the no-help condition, the correlation was positive. For the help condition, the correlation was negative. The result for the no-help condition is a replication of earlier findings, which showed a moderate correlation between pretest and posttest. The result for the help condition shows that making knowledge available to students has had the effect of reversing this relationship. Further analysis revealed that this is mainly caused by a large increase from pretest to posttest for some students in the help condition (the largest is an increase of 12 points, compared to a largest increase of 6 for the no-help condition).

7.3.1.3 Discovery skills

The Peter paper-and-pencil test was administered to students after they had indicated that they knew the effect of the five variables in the Peter task. Test items required students to state the effect of each of the five variables in the Peter task on the outcome. Based on the answers, students were classified in one of four different levels. Students were classified as level 1 when they did not find correct main effects, and no interaction effect (low performance); as level 2 when they did find correct main effects but no interaction effect; as level 3 when they found correct main effects and the correct interaction effect, but also one or more incorrect interaction effect; finally, as level 4 when they found correct main effects and the correct interaction effect, and no other effects (high performance). The category can be seen as a measure of performance in the Peter task, with level being the lowest, and level 4 the highest level of performance.

In this experiment, performance on the discovery skills test ranged from level 1 to level 4. Of the 32 subjects, 8 scored at level 1, 17 scored at level 2, 3 scored at level 3, and 4 scored at level 4. This indicates that most students (78 percent) in this experiment were not able to find an interaction in the data from the Peter task. This result replicates the finding from the Optics 99 experiment, in which students in the Vocational Education group performed worse at the test for discovery skills than students in the Pre-Scientific Education group. The correlation between performance on the generic knowledge test and the discovery skills test was 0.25, which is low.

7.3.2 Interaction results

This section gives results on the interaction with the Optics computer simulation. Differences between the no-help and the help conditions are especially focused on.

7.3.2.1 No-help versus Help condition

Loss of some subject data resulted in a total of 27 Optics interaction logfiles, 15 for the no-help condition and 12 for the help condition.

An important question is whether students in the help condition actually made use of the tips that were made available to them. The instruction made it clear that subjects were free to use the tips, but that they did not have to view every tip. Analysis of the help condition showed that subjects viewed a mean total of 5.5 tips (SD=3.0) out of 9 tips. No relation was found between the number of tips that were accessed by subjects and performance in the Optics simulation. The correlation between the number of tips and performance on the domain-specific knowledge posttest was –0.1. It can be concluded that subjects did make use of the knowledge tips that were offered to them. Reviewing of already available tips was used frequently by subjects.

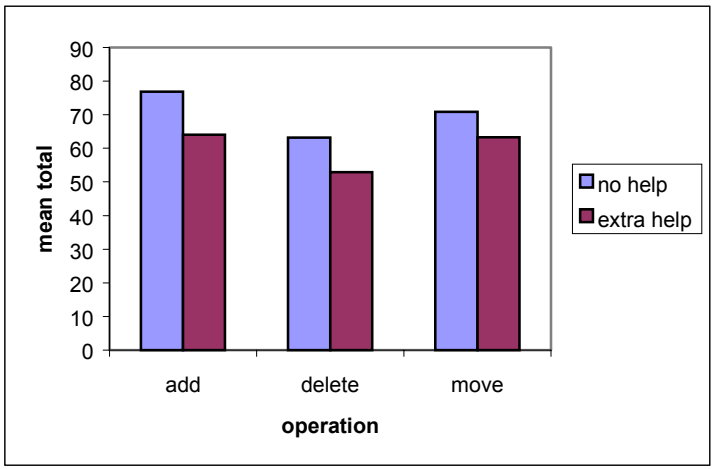


Figure 7-4. *Comparison of the no-help and help conditions for basic operations in Optics*

A comparison was made on the number of basic operations in Optics. The mean number of add, delete, and move operations that was performed by subjects in the no-help condition was 70.3 (SD=43.4), and in the help condition 60.1 (SD=37.5). The difference is not statistically significant. When we look at the three basic operations separately (as shown in Figure 7-4), it can be seen that for all three operations, differences between the conditions are small. Students in the help condition appear to be less active in general than students in the no-help condition.

7.3.2.2 Prior knowledge and Optics interaction

A division into poor and high prior knowledge was made on the basis of test scores on the pretests. The dividing line was set at subjects who scored lower than one standard deviation below the mean, and subjects who scored higher than one standard deviation above the mean. For generic knowledge, this meant that two groups were formed of subjects who scored below 17 (n=6) and above 25 out of 32 items (n=5). For domain-specific knowledge, two groups were formed of subjects who scored below 12 (n=6) and above 19 out of 30 items (n=5).

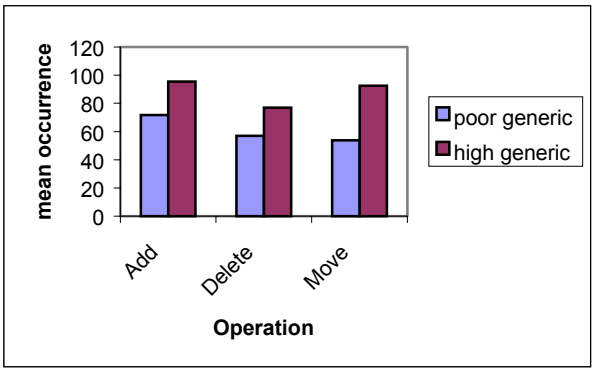


Figure 7-5. *Comparison on Add, Delete, and Move operations for poor and high generic knowledge*

In the analysis of prior knowledge and skills and interaction in the Optics simulation, there is a problem in that the number of subjects that was used in the experiment was not high enough to allow for a comparison of high and poor prior knowledge divided over the no-help and help conditions. It also turned out that not all logfiles could be processed, which resulted in further loss of information about interaction behavior. Below, we will present results for the combination of the groups in the help and in the extra help conditions. The availability of knowledge tips in the help condition may have interfered with the discovery learning process.

In Figure 7-5, a comparison on basic operations for students with poor generic knowledge (n=4) and high generic knowledge (n=4) is shown. The results replicate findings from the Optics 97 and Optics 99 experiments. Students with high generic knowledge are more active in working with the Optics simulation than students with poor generic knowledge.

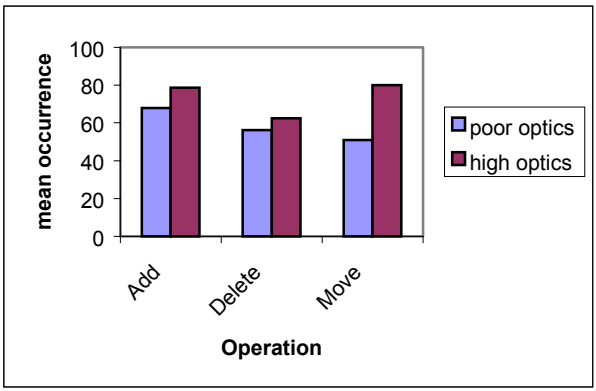


Figure 7-6. *Comparison on Add, Delete, and Move operations for poor and high domain-specific knowledge*

Figure 7-6 shows a comparison on basic operations for students with poor domain-specific knowledge (n=6) and high domain-specific knowledge (n=3). This time, results are similar to the findings for generic knowledge, although the low number of subjects in the high domain-specific knowledge group makes a comparison difficult.

Finally, a comparison between the groups with poor (n=6) and high (n=3) discovery skills is shown in Figure 7-7. Again, a comparison is made difficult

by the low number of subjects. Still, the comparison reveals a result that is similar to the result in the Optics 99 experiment in the comparison of students with poor and high generic knowledge. Both groups perform a similar amount of Add and Delete operations, but subjects with high discovery skills perform more Move operations.

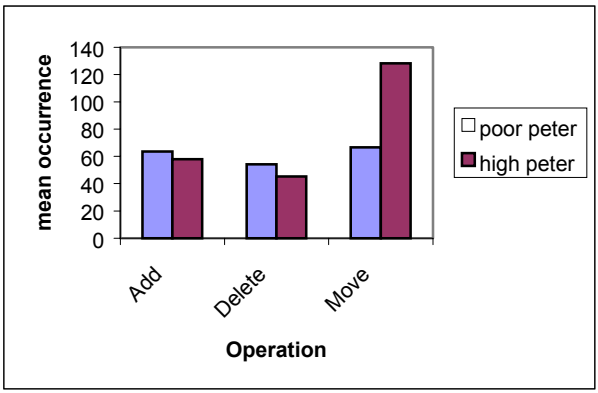


Figure 7-7. *Comparison on Add, Delete, and Move operations for poor and high discovery skills*

7.4 Discussion

The study that was described in this chapter was designed to examine whether or not discovery learning in the Optics computer simulation could be rendered more effective when learning was supported. The support that was available to a group of students in vocational education consisted of a small set of tips which gave domain-specific and generic background information on the subject of geometrical optics. This information was available on a 'just in time'-basis, which means that it could be accessed during discovery learning. The Optics simulation was configured for this experiment to be more limited in scope than had been the case in our previous studies. The topic of image projection through a lens was focused on. By randomly assigning students to either a no-help condition or an help condition, we were able to examine the effect of the extra information that was available during discovery learning.

The finding that test performance of subjects in the no-help condition deteriorated from pretest to posttest is in line with findings from the previous experiment. On the one hand, when discovery learning proceeds in an unsupported way, no gain in knowledge is found and performance can even get worse. On the other hand, when support is given by providing subjects with knowledge about the concepts and relations of the simulated domain, performance improves.

Although the knowledge tips were actively used by subjects in the help condition, and the effect on performance on the posttest was significant, it is difficult to find differences in the discovery learning behavior of subjects with or without tips.

It was found that students in the help condition appeared to be less active in general than students in the no-help condition. This finding replicates a finding in the Optics 99 experiment. In Figure 6-8, the mean activity level was shown for students with poor and high domain-specific knowledge. As can be seen from the figure, in the third situation students with high domain-specific knowledge are less active than subjects with poor domain-specific knowledge. The third situation in the Optics 99 experiment is very similar to the situation that was used in the current experiment. Combining these findings leads us to the conclusion that students who had access to domain-specific knowledge know more about the concepts and relations in a simulation. In the Optics 99 experiment, students with high domain-specific knowledge may have recognized the particular configuration that was used in the third situation, which led to their decreased activity in this situation. In the case of the present experiment, something similar may have happened to students who had knowledge available, causing them to be less active than students who did not.

8

Discussion: the implications of it all

> The Qanuc-folk of the snow-mantled Trollfells have a proverb. "He who is certain he knows the ending of things when he is only beginning them is either extremely wise or extremely foolish; no matter which is true, he is certainly an unhappy man, for he has put a knife in the heart of wonder."
>
> *Tad Williams, Memory Sorrow and Thorn part I (1988)*

8.1 Introduction

In this thesis, the main research question was: how does prior knowledge influence learning in a scientific discovery learning domain. The context in which discovery learning took place was computer simulation (Chapter 1). The SDDS model of scientific discovery learning (Klahr & Dunbar, 1988) was used as a theoretical framework for the research (Chapter 2). In three empirical studies, a simulation on the topic of geometrical optics ('Optics') was used (Chapter 3). Generic and domain-specific knowledge were operationalized through tests, and methods were developed to measure interaction with the computer simulation (Chapter 4). In the first study (Optics 97), generic knowledge and domain-specific knowledge were measured to examine differences between more and less proficient students in scientific discovery learning processes (Chapter 5). This study made use of the Optics simulation and also of the Bubbles simulation about which students did not have prior domain-specific knowledge. We did not find influence of domain-specific prior knowledge on discovery learning processes in the Bubbles and Optics simulations, and only a limited effect for generic knowledge. In the second study (Optics 99), differences in discovery learning behavior between more and less proficient students on generic and domain-specific knowledge were examined (Chapter 6). In addition to using revised tests for generic and domain-specific knowledge, this study used a test for discovery skills (the 'Peter'-test). Discovery learning behavior was studied in four situations (ranging from simple to complex) in the Optics simulation. This study showed a consistent effect of prior generic knowledge on the discovery learning process: students with high generic knowledge were more active in the Optics simulation than students with poor generic knowledge. Discovery skills were related to discovery learning behavior in the first two (simple) situa-

tions in the Optics simulation. Students with high discovery skills were more active in these situations than students with poor discovery skills. For domain-specific prior knowledge it was shown that students with high domain-specific knowledge were less active than students with poor domain-specific knowledge in one situation in the Optics simulation. A disadvantage of the Optics 99 experiment was that students spent only a short time in each situation of the Optics simulation. For the third and final study (Optics 2000) it was decided to focus on the situation in the Optics simulation for which a difference between groups of high and poor prior domain-specific knowledge had been found in the Optics 99 study. The Optics 2000 study looked again at prior generic knowledge, domain-specific knowledge, and discovery skills in relation to discovery learning in the Optics simulation (Chapter 7). In this study, one group of students had a set of 'knowledge tips' at their disposal. By letting students have access to knowledge tips, knowledge about optics was manipulated in this study. It was found in this study that students who had the tips available to them showed a learning gain from pretest to posttest, as opposed to a group who worked in a similar way in the Optics simulation, but without knowledge tips.

In this chapter, the overall results from the three empirical studies are discussed. First, the prior knowledge and discovery skills measures are closely examined. Possible limitations of the tests are discussed, and guidelines for improvements are provided. Second, the importance of using support to aid discovery learning with computer simulations is discussed. A comparison of the third experiment with the first and second reveals that even a relatively small support measure can lead to learning improvements. Implications for the SDDS framework of the results on the prior knowledge and skills tests on discovery learning behavior are discussed in the next section. Third, the method of studying interaction behavior as it was performed in all three studies is discussed. Limitations and possible extensions of the method that was used in the studies are taken into account. Finally, some of the general implications of the studies that were carried out are addressed. The studies have potential implications for education in general, and for further research in the field of scientific discovery learning.

8.2 Prior knowledge and skills measures

8.2.1 Generic knowledge

In all three experiments, a pretest for generic knowledge was used. The test was used to measure knowledge about mathematical relations. These rela-

tions could be of different types (e.g., qualitative or quantitative) and were presented in different formats (e.g., numerically or in a graph). The test that was administered in the Optics 99 and Optics 2000 studies differed from the test that was used in the Optics 97 study. In the first experiment, in which only students in pre-scientific education were used as subjects, the test was largely made up of items that used number sequences. Students had no problems with answering these items. A problem with measuring generic knowledge lies in specifying the border between generic and domain-specific knowledge. On the one hand, when different topics from mathematics are chosen to create a test, there is a risk of the questions being too specific. On the other hand, when only general questions are asked, the test runs the risk of becoming an intelligence test.[1] Therefore, it was decided to focus on the various ways in which mathematical relations can be represented. For scientific discovery learning to be successful, availability of all these types of generic knowledge was considered important. In Appendix A, the test that was used in the last experiment is shown. The Alpha level of 0.71 suggests a reasonable internal consistency of the test. The measure for generic knowledge as it was developed and used in the experiments can be said to lie in-between a pure mathematics test (which would be a domain-specific test) and a general test for number processing (which would be an intelligence test). In this sense, the test can be seen as a member of the family of general science tests such as TIPS (Burns, Okey, & Wise, 1985) and TOGS (McKenzie & Padilla, 1986), with an emphasis on knowledge of mathematical relations.

8.2.2 Domain-specific knowledge

All three experiments made use of tests for domain-specific knowledge about geometrical optics. Over the experiments, the design and way of administering the test was greatly changed. Based on results of the test, students were divided into two groups: one with poor and one with high prior domain-specific knowledge. All students who participated in the experiments had been following courses in which the topic of geometrical optics

[1] Indeed, many popular IQ tests make use of the type of number sequences that were used in the generic knowledge test that was used in the Optics 97 study; the actual sequences that were used were relatively easy for students with a normal level of intelligence.

was delivered. Therefore, it was expected that some students would remember more about the topic than others. It turned out that in general, students only knew little about the topic. In Table 8-1, an overview of prior domain-specific knowledge scores is given (with standard deviations between parentheses), together with the alpha level of the test and some comments on the test format.

Table 8-1. *Overview of domain-specific knowledge tests*

Study	Mean score / items	Alpha	Test format
Optics 97	6.8 (2.1) /18	0.29	Computer test – Intuitive 'what-if' knowledge test
Optics 99	12.9 (2.7) /24	0.34	Computer test – modified what-if test
Optics 2000	15.3 (3.7) /30	0.51	Paper-an-pencil test –what-if items and declarative items

In the first experiment, the fact that almost no students scored very high on the test had the consequence of reducing the variation in test scores, which made a division into groups of poor and high knowledge unreliable. Also, the what-if test that was used had a very low value of Cronbach's alpha. The test that was used in the first experiment has a number of characteristics that differentiate it from more standard types of tests. An important characteristic of the test is that learners are asked to balance their speed of answering and their accuracy. This is done to prevent learners from consciously solving the problem, and use their 'gut feeling' instead. The problem is that the request to balance speed and accuracy will be interpreted differently by different people. Some people will focus on answering as quickly as possible in disregard of the accuracy of their answers. Other people will find being accurate more important than being fast and answer more slowly. We feel that the instruction to balance speed and accuracy may influence test performance in a negative, or possibly in some other unpredictable way. One of the assumptions about intuitive knowledge is that this knowledge, though not verbalisable, can be quickly put to use by learners. In that case, imposing a restriction on the test by asking learners to balance accuracy and speed may introduce an undesired artifact. Since learners who have intuitive knowledge available are expected to be quick in their answering, asking them to trade off speed and accuracy may lead some of them to favor accuracy, in which case the effect of intuitive knowledge may be lost. A solution to the problem may be, for example, to impose no restrictions on the answering behavior of learners, but to have them instead answer within a certain period of time (shown next to the question as a decreasing counter). Unfortunately, no studies have been carried out yet in which different types of in-

structions are compared in the context of an intuitive knowledge test. The possible ambiguity that the speed/accuracy demand imposed, combined with the meager test results on the prior domain-specific knowledge test, led to the decision to discard this instruction for the second experiment. As can be seen from Table 8-1, the results were not completely satisfactory. The variation in test scores and the alpha level still were quite low. However, students performed better on the test (about 54 percent correct on the average, as opposed to 38 percent correct in the first experiment). To make a further comparison between the tests, we looked at the time taken to answer the test. In the first experiment, the mean time taken to answer all test items was 672 seconds (as reported in Chapter 5), or 37 seconds for each question. Although, in the second experiment, students were not asked to trade off speed and accuracy, answer times were still registered. The mean time taken to answer all items was 776 seconds, or 26 seconds for each question.[2] This means that, although no instruction was given about trading off speed and accuracy, students were faster in answering test items in the second experiment. It may be that the revised items that were used in the domain-specific knowledge test in the second experiment were easier to answer than the items in the first experiment (an explanation which is corroborated by the higher percentage of correctly answered items). Still, we can conclude that when the instruction to trade off speed and accuracy is left out of the instruction of the intuitive knowledge test, this does not negatively affect answering times.

The poor test results and low reliability of the test led us to use a more 'classical' test in the final experiment. A paper-and-pencil format was used, and students were allowed to review their answers. Also, a number of test items that asked more directly for declarative knowledge about geometrical optics (such as the lens formula) were introduced in the test. The percentage correct on this test (51 percent) was similar to the test for the second experiment. However, there was a larger variation in scores, and the alpha level of the test was moderate. The declarative knowledge items were mostly answered correctly, percentages of correctness for these items were higher than for the other items.

---

[2] The students who were in technical vocational education were a little slower in answering each question (mean time 27.2 seconds) than students in pre-scientific education (mean time 24.8 seconds). However, the result was not significant.

The question may be raised why the what-if test format was used in the domain-specific prior knowledge tests. If intuitive knowledge is gained from working with a computer simulation, why use a prior measure for it? This is really a question about whether or not prior intuitive knowledge can exist. Answering an intuitive knowledge test means that knowledge of concepts and relations in a domain have to be combined. The poor performance that was found on the prior domain-specific knowledge tests may indicate, that although students had knowledge about the concepts and relations of geometrical optics, they had trouble in combining this knowledge to answer the intuitive knowledge test items.

8.2.3 Discovery skills

In the second experiment, the Peter test for discovery skills was introduced. The test used a combination of a computer task and a paper-and-pencil test. To gain a high score on the discovery skills test, the results of experimenting with the Peter computer task environment had to be successful. What was tested for in the task was the skill level in performing relevant experiments and drawing the right conclusions from them. It was shown that performance on the Peter test was positively related to the number of experiments that students performed in the computer task. Also, the strategy followed in the computer task varied among students with different discovery skills. Students with high discovery skills showed more VOTAT-oriented behavior than students with poor discovery skills.[3]

In the Optics 99 study, a relation was found between level of discovery skills and activity in the first two situations in the Optics simulation. It was concluded that discovery skills affected discovery learning behavior in the more simple situations in the simulation, but not in the more complex situations.

In summary, what answers can be given to the main research questions that were asked at the beginning of Chapter 4? The first research question asked about differences between the influence of domain-specific and generic knowledge on discovery learning processes. For generic knowledge, a general effect was found: students with high generic knowledge are more active in discovery learning than students with poor generic knowledge. For domain-specific knowledge, only a limited influence on discovery learning was found. However, when knowledge about optics is available during discov-

[3] VOTAT means Vary One Thing At A Time (Tschirgi, 1980). This refers to the strategy to change only one variable from one experiment to the next.

ery learning, it supports students in activating their prior knowledge, which has a positive influence on the learning result. This leads us to the second research question, which dealt with the effectiveness of discovery learning. In the Optics 97 and Optics 99 studies, a positive relation was found between domain-specific prior knowledge and performance on a domain-specific knowledge posttest. This means that pre- and posttest performance was similar for most students. In the Optics 2000 experiment, this result was reversed for students in the help condition. This was caused by students with poor domain-specific prior knowledge who showed a large learning gain after working with the Optics simulation. The conclusion is that enabling students to access information during discovery learning can change the relation between prior knowledge and posttest performance. The third research question dealt with the strategy that students followed during discovery learning. What changes could be detected over time? In the Optics 97 and Optics 99 studies, discovery learning processes were analysed over time. When the level of activity was analysed, differences between groups with poor and high knowledge were found. However, it turned out to be difficult to unequivocally interpret the results. In general, the discovery learning process does change over time. The beginning and end part of a discovery learning session are different from the middle part. In what way this is dependent on prior domain-specific and generic knowledge is still an unresolved question.

8.3 Implications for the SDDS framework

The empirical studies that are described in this thesis were carried out in the context of the SDDS framework by Klahr and Dunbar (1988). The extended version of this framework by Van Joolingen and De Jong (1997) was used to explain the (potential) role of prior knowledge on scientific discovery learning processes. An important contribution by the extended SDDS framework is the division of hypothesis space into a number of subspaces. In Chapter 2, it was argued that domain-specific knowledge is represented in hypothesis space in the *learner hypothesis space* (all statements about a domain that a learner has knowledge of). The process of scientific discovery learning entails searching experiment space and hypothesis space, with the goal of moving the learner hypothesis space closer to the *target conceptual model* (all true statements about a domain). The absence or presence of domain-specific prior knowledge implies a different configuration of the learner hypothesis

space, which affects the distance between the learner hypothesis space and the target conceptual model. As it was used in this thesis, generic knowledge referred to the level of knowledge about mathematical relations. Hypothesis space can be broadly divided in a space of variables and a space of relations (Van Joolingen & De Jong, 1997). Generic knowledge was expected to influence the type of mathematical relations a student would have knowledge about, which should have an effect on the discovery learning process.

In the Optics 97 and Optics 99 studies, the effect of domain-specific knowledge on discovery learning in the Optics simulation was limited. This can be explained by the poor knowledge about geometrical optics that students had in general. In simple configurations of the Optics simulation, students may have had trouble in connecting information from the simulation with their background knowledge. In complex configurations, their knowledge probably was simply too limited to have an influence on the discovery learning process. Still, in the Optics 99 study it was found that in one specific situation in the optics simulation (a configuration in which an object was projected on a screen), students who had been classified as having high domain-specific knowledge showed less activity than students with poor domain-specific knowledge. The hypothesis was stated that the particular situation that was used was recognized by students with high domain-specific knowledge. Because they had knowledge already available, they were less active in the situation than students with poor domain-specific prior knowledge. For the SDDS framework, this means that when the learner hypothesis space is closer to the target conceptual model, students make less 'moves' through experiment space to gain information.

From the empirical studies, it appears that generic knowledge influences the general level of activity in the Optics simulation. Students with a high level of generic knowledge showed more activity than students with poor generic knowledge. This effect was similar for simple and complex configurations of the computer simulation, and was found in both the Optics 97 and the Optics 99 experiments. This result can be explained by the extended SDDS framework. High generic knowledge means more knowledge of the type of mathematical relations that exist. This means that the effective learner space is larger, which gives more freedom in performing experiments.

Discovery skills, which were measured in the Optics 99 and Optics 2000 study, influenced learning behavior in simple configurations of the Optics simulation. In these situations, students with high discovery skills were more active than students with poor discovery skills. In the computer task that was related to the Peter test, successful students were also more active than students with poor discovery skills. The level of discovery skills was also moderately correlated with the level of generic knowledge. One conclu-

sion that may be drawn is that students with high generic knowledge have more knowledge about types of (interesting) relations, and know better than students with poor generic knowledge what different types of experiments they can perform. In the Peter computer task, they also know better what type of experiments to perform, which is reflected in better performance on the Peter test.

The relation between discovery skills and generic knowledge could be the subject of further research. It may be that both are really measures of generic inquiry skills: knowledge of mathematical relations is needed to perform effective experiments, and by performing effective experiments more (generic) knowledge is gathered. Also, both discovery skills and generic knowledge may be related to a more general measure of intelligence, an issue that was not explored in these studies.

We have come to the following general conclusions with respect to the influence of prior knowledge and skills on scientific discovery learning processes.

Generic knowledge. The knowledge of types of mathematical relations that can exist between variables, as expressed in generic knowledge, has a general influence on discovery learning, regardless of the complexity of the learning environment. Students with high generic knowledge have more knowledge about the type of mathematical relations that exist, which means that they can perform more different types of experiments (their effective learner space is larger). From this, it follows that they are more active than students with poor generic knowledge.

Domain-specific knowledge. Knowledge about the variables and relations that exist within a certain domain, as expressed in domain-specific knowledge, only influences discovery learning behavior when it can be accessed by students. It may be that prior domain-specific knowledge can only be accessed when relevant variables in a situation are recognized by students. When this is realized, high domain-specific knowledge means that the learner hypothesis space is larger and closer to the target conceptual model. This will lead students to be less active than students with poor domain-specific knowledge. In a complex situation, for which no domain-specific knowledge is available, domain-specific knowledge does not influence discovery learning processes.

Discovery skills. The skill with which relations between variables are discovered in a domain influences discovery learning behavior when a situation is

simple enough. In a complex situation, the absence or presence of other types of knowledge becomes more important.

8.4 The use of support

To study the influence of prior knowledge on scientific discovery learning in a computer simulation, it was decided to keep support at a minimum. We did not want active support measures to interfere with the basic discovery learning task. The absence of active support had two side effects. Firstly, the discovery learning task was not optimized to be as effective as possible. This means that we did not expect a positive learning effect from working with the Optics simulation. In the experiments, even negative learning effects were found. Secondly, in combination with the low level of domain-specific prior knowledge that was found in general, working with the Optics simulation was very complicated for students. Many students reported having trouble in making sense of the general assignment that was given to them. In the Optics 99 experiment, the Optics simulation was worked with in a fixed order, through four situations that built up in complexity. Because of time limitations, this meant that each situation was only worked with for a short time, which means that a large part of working with the simulation was spent on getting used to the new situation. It was decided to focus on one situation only in an experiment that was designed to test the effectivity of knowledge support. In the Optics 2000 study, domain-specific knowledge was manipulated by delivering information to students on a just-in-time basis: domain-specific information was delivered while students worked with the Optics simulation. The results of the Optics 2000 study showed that when students had access to information about geometrical optics, they did make use of it. Also, the information that was provided by the 'knowledge tips' helped in manipulating domain-specific knowledge, which in turn led to an improvement on domain-specific knowledge from pretest to posttest. Our conclusion is that discovery learning in the Optics simulation can be made effective, even by using the relatively minor support measure of knowledge support.

8.5 Interaction analysis

The method that was chosen for analyzing the interaction between students and the Optics computer simulation was outlined in Chapter 4. The operations that students performed while they worked with the computer simulation were registered in a log file. The log file was subjected to an analysis, for

example the frequency of different types of operations was calculated. This method of analysis offered several advantages over other methods, such as thinking aloud. We were able to experiment with a group of students who worked simultaneously with the Optics simulation. Also, raw log data did not have to be transcribed before it could be analyzed (although the data was converted to become more easily processed). A disadvantage of the method was that students' reasoning behind the operations they performed in the simulation could not be ascertained this way. In the Optics 99 experiment, students were asked in-between situations in the Optics simulation about their reasons for working in the way they had done. However, use of this method did not yield much information, because most students merely stated what operations they had performed. To get insight in the discovery learning process, a distinction between different types of operations was made. In the Optics 97 experiment, a distinction between Type I (core), Type II (peripheral), and Type III (irrelevant) operations was made. Type I operations were necessary to experiment in the computer simulation, and referred to object manipulation. Type II operations helped students get more information from the simulation than was visible from the objects alone. In the Optics 99 experiment, a similar distinction was made between manipulation operations and guiding operations.

In the analysis, interaction behavior was examined both as an aggregate of all operations in a situation in the Optics simulation, and over time. In the analysis over time, we defined a number of short sequences of operations a priori. These sequences consisted of operations that yielded either qualitative information or quantitative information. As was discussed in Chapter 4, this approach was chosen in favor of a purely sequential analysis, in which occurrences of any sequential patterns are explored. The approach that was chosen was not entirely successful. Firstly, it turned out to be difficult to define meaningful short sequences of operations. Secondly, in many cases a particular sequence can deviate from the defined pattern, for example when another (irrelevant) operation is performed in-between the other operations. Which operations should be left out and which operations should not is an open question, which is not answered by the research this thesis. Thirdly, it is probably not possible to capture every instance of behavior that is meaningful to a student by pre-defining sequences of operations.

Overall, we conclude that the frequency-based analysis of a priori defined sequences of operations that was used here can be used in general in the analysis of discovery learning processes in computer-based simulations. At

the moment, no general method of analyzing learner-computer interaction in these situations exist, with each study following its own guidelines. The definition of a generic set of sequences of operations can help in creating a stronger foundation for the analysis of discovery learning behavior.

8.6 General implications

A conclusion that can be drawn from this thesis as a whole, is that discovery learning is an intricate and complex phenomenon. It was shown that prior generic knowledge has a general influence on discovery learning processes. The effect of prior domain-specific knowledge was shown to be limited. Domain-specific knowledge only has influence when it can be accessed during discovery learning. For instructional designers, it is important to take these findings into account when creating situations in which discovery learning should yield optimal results. For students, it is a complex and (still) not familiar way to learn. For researchers, the discovery learning process is difficult to study. What are the more general implications that can be drawn from the studies that have been described in this thesis? In this final section, we will list a number of possible implications for the educational field and for further research in the area of scientific discovery learning.

8.6.1 Implications for the educational field

In Chapter 1, the changes that the educational field has been experiencing for some time were discussed. A distinction was made between influences that originate from technological advances which have found their way into the classroom, and theoretical advances in views on how a learning situation should be designed. The use of computer simulations that foster discovery learning processes will in all probability continue to increase. This makes research into the way different learners react to a discovery learning situation, and into the way learning can be supported, a useful enterprise.

In the studies that were conducted as part of this thesis, it was found that students scored poorly on the domain-specific knowledge tests. It was observed that some of the teachers who were involved in the design of the experiments expressed surprise at this finding, because they had expected knowledge about geometrical optics to be much higher than it turned out to be. This indicates that there is a discrepancy between knowledge that is measured by a standard test for declarative knowledge (on which students can show normal performance) and an intuitive knowledge test (on which students showed poor performance). The implication for education is, that when a simulation is used to foster discovery learning, special care has to be

taken to use a domain-specific knowledge test that is appropriate for the learning situation.

In the Optics 97 study, it was found that students were more motivated to work with the Bubbles computer simulation than with the Optics simulation. The game-like task that the Bubbles simulation provided, combined with the fact that students had no prior knowledge about the Bubbles domain, probably had a positive effect on motivation. In contrast, the subject of geometrical optics was recognized by students as a typical school subject, which had a negative effect on motivation. The implication for education is that simulation-based discovery learning environments may be best employed with topics that have not been treated in the normal curriculum yet. Students appear to be sensitive to the difference between learning something new and learning something that has been learned before.

8.6.2 Implications for further research

In this thesis, the process of scientific discovery learning was studied in the context of a computer simulation. For the experiments a number of tests and tools were developed. For domain-specific knowledge, originally a what-if test format was used. In later experiments, some of the characteristics of a what-if test were removed from the test, which had a positive effect on test scores and reliability. It is uncertain why the what-if test format failed for the domain of geometrical optics. We did find that students had poor knowledge on the subject, and it was hypothesized that students may have had trouble in trading off speed and accuracy in answering. Still, research is needed to investigate under what conditions a what-if test format can be used and what changes can be made to the format. The test format that was used in the different studies in this thesis can serve as a starting point for further study, in which different formats are compared with each other. Also, the relation between intuitive knowledge and explicit knowledge can be further explored. Using a combination of both types of test could be more useful in measuring prior knowledge than the use of either type on its own.

The Optics 99 and Optics 2000 studies made use of a test for generic knowledge and a test for discovery skills. It was found that performance on these tests was moderately correlated. The exact relation between generic knowledge and discovery skills is not known, and it was suggested that both measures might be related to another construct, such as general level of intelligence.

The set of measures for interaction between students and the computer simulation that were developed in the context of the empirical studies in this thesis have proven to be useful in analyzing discovery learning processes. We have argued that a frequency-based analysis of a prior defined sequences of operations can provide a general framework for studying scientific discovery learning processes. In further research, this approach should be refined and tested in other simulation-based environments.

References

Alessi, S. M. (1988). Fidelity in the design of instructional simulations. *Journal of Computer-based Instruction, 15*, 40-47.

Alexander, P. A. (1992). Domain knowledge: evolving themes and emerging concerns. *Educational Psychologist, 27*, 33-51.

Alexander, P. A., & Judy, J. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research*, *58*, 375-405.

Amarel, M. (1983). The classroom: An instructional setting for teachers, students, and the computer. In A. C. W. Wilkinson, *Classroom Computers and Cognitive Science* (pp. 15-29). New York: Academic Press.

Anderson, J. R. (1990). *Cognitive Psychology and its Implications* (3rd ed.). New York: Freeman.

Anderson, J. R. (1995). *Learning and Memory: An Integrated Approach*. New York: Wiley.

Ausubel, D. P. (1963). *The Psychology of Meaningful Verbal Learning*. New York, London: Grune and Stratton.

Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). *Educational Psychology: A Cognitive View*. (2nd ed.). New York: Holt, Rinehart, & Winston.

Bakeman, R., & Gottman, J. M. (1997). *Observing Interaction: An Introduction to Sequential Analysis*. (2nd ed.). New York: Cambridge University Press.

Bakeman, R., & Quera, V. (1995). *Analyzing Interaction: Sequential Analysis with SDIS and GSEQ*. New York: Cambridge University Press.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C-L. C. (1985). Effectiveness of computer-based education in secondary schools. *Journal of Computer-based Instruction, 12*, 59-68.

Berkum, J. J. A., & De Jong, T. (1991). Instructional environments for simulations. *Education & Computing, 6*, 305-358.

Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *The Quarterly Journal of Experimental Psychology*, *36*, 209-231.

Betz, J. A. (1995). Computer games: Increase learning in an interactive multidisciplinary environment. *Journal of Educational Technology Systems, 24*, 195-205.

Biemans, H. J. A. (1997). *Fostering Activation of Prior Knowledge and Conceptual Change*. PhD Thesis, University of Nijmegen, The Netherlands.

Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, *9*, 171-184.

Van Breukelen, G. J. P. (1989). *Concentration, Speed and Mental Precision in Mental Tests*. PhD Thesis, University of Nijmegen, The Netherlands.

Van Breukelen, G. J. P. (1997). Separability of item and person parameters in response time models. *Psychometrika, 4*, 525-544.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, *31*, 21-32.

Bruner, J. S. (1974). *Beyond the Information Given: Studies in the Psychology of Knowing*. London: Allen and Unwin.

Burns, J. C., Okey, J. R., & Wise, K. C. (1985). Development of an integrated process skill test: TIPS II. *Journal of Research in Science Teaching*, *22*, 169-177.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*, 1-49.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research & Development, 42*, 21-29.

Driver, R., & Easley, J. (1978). Pupils and paradigms: A review of literature related to concept development in adolescent science studies. *Studies in Science Education*, Vol. 5, 61-84.

Duffy, T. M., & Cunningham, D. J. (1996). Constructivism: Implications for the design and delivery of instruction. In D. H. Jonassen, *Handbook for Educational Communications and Technology* (pp. 170-198). New York: Simon & Schuster.

Dugdale, S., & Kibbey, D. (1990). Beyond the evident content goals: Part I, tapping the depth and flow of the educational undercurrent. *Journal of Mathematical Behavior, 9*, 201-228.

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science, 17*, 397-434.

Dunbar, K., & Klahr, D. (1989). Developmental differences in scientific discovery processes. In D. Klahr, & K. Kotovsky, *Complex Information Processing: The Impact of Herbert A. Simon* (pp. 109-143). Hillsdale, New Jersey: Erlbaum.

Edwards, L. D. (1993). Microworlds as representations. In A.A. di Sessa, C. Hoyles, & R. Noss, *Computers and Exploratory Learning* (pp. 155-173). Berlin, Germany: Springer-Verlag (NATO ASI Series F, vol. 146).

Edwards, D., & Mercer, N. (1987). *Common Knowledge: The Development of Understanding in the Classroom*. London, New York: Methuen.

Elmore, R. L., & Tennyson, R. D. (1996). Psychological foundations and the automation of instructional systems. *Journal of Structural Learning*, *12*, 253-267.

Fischbein, E. (1987). *Intuition in Science and Mathematics*. Dordrecht, The Netherlands: Reidel.

Friedler, Y., Nachmias, R., & Linn, M. C. (1990). Learning scientific reasoning skills in microcomputer-based laboratories. *Journal of Research in Science Teaching, 27*, 173-191.

Furneaux, W. D. (1973). Intellectual abilities and problem-solving behavior. In H. J. Eysenck, *The Measurement of Intelligence* (pp 212-237). Lancaster, England: MTP.

Gagné, R. M. (1966). Varieties of learning and the concept of discovery. In L. S. Shulman, & E. R. Keislar, *Learning by Discovery, a Critical Appraisal* (pp. 135-150). Chicago, Illinois: Rand McNally.

Galili, I., & Hazan, A. (2000). Learners' knowledge in optics: Interpretation, structure and analysis. *International Journal of Science Education, 22,* 57-88.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review, 98*, 254-267.

Gilbert, J.K., & Watts, M. (1983). Concepts, misconceptions and alternative conceptions: Changing perspectives in science education. *Studies in Science Education*, 10, 61-98.

Glaser, R., Schauble, L., Raghavan, K., & Zeits, C. (1992). Scientific reasoning across different domains. In E. de Corte, M. Linn, H. Mandl, & L. Verschaffel, *Computer-based Learning Environments and Problem Solving*. Berlin, Germany: Springer-Verlag.

Goldenberg, E. P. (1988). Mathematics, metaphors, and human factors: Technical, and pedagogical challenges in the educational use of graphical representation of functions. *Journal of Mathematical Behavior, 7*, 135-173.

Grabinger, S. R. (1996). Rich environments for active learning. In D. H. Jonassen, *Handbook for Educational Communications and Technology* (pp. 665-693). New York: Simon & Schuster.

De Groot, A. D. (1969). *Methodology: Foundations of Inference and Research in the Behavioral Sciences*. The Hague: Mouton.

Hasselerharm, E., & Leemkuil, H. (1990). The relation between instructional control strategies and performance and attitudes in computer-based instruction. In J. M. Pieters, P. R. J. Simons, & L. De Leeuw, *Research on Computer-based Instruction*. Amsterdam: Swets & Zeitlinger.

Hawkridge, D. (1990). Who needs computers in schools and why? *Computers and Education, 15*, 1-6.

Hecht, E. (1998). *Optics* (3rd ed.). Reading, Massachusetts: Addison-Wesley.

Hewson, P. W., & Hewson, M. G. A. (1984). The role of conceptual conflict in conceptual change and the design of science instruction. *Instructional Science*, *13*, 1-13.

Holland, J. H., Holyoak, K., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of Inference, Learning, and Discovery*. Cambridge, Massachusetts: MIT Press.

Hulshof, C. D., De Jong, T., & Van Joolingen, W. R. (1998). *The influence of different types of prior knowledge on the learning process in two discovery learning environments*. Internal Report. University of Twente: IST.

Van Der Hulst, A. H. (1996). *Cognitive Tools: Two Exercises in Non-directive Support for Exploratory Learning*. PhD Thesis, University of Amsterdam, The Netherlands.

Johnson, M. A., & Lawson, A. E. (1998). What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes? *Journal of Research in Science Teaching, 1*, 89-103.

Jonassen, D. H. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm? *Educational Technology Research & Development*, *39*, 5-14.

Jonassen, D. H., Campbell, J. P., & Davidson, M. E. (1994). Learning *with* media: Restructuring the debate. *Educational Technology Research & Development, 42*, 31-39.

Jonassen, D. H., & Reeves, T. C. (1996). Learning *with* technology: Using computers as cognitive tools. In D. H. Jonassen, *Handbook for Educational Communications and Technology* (pp. 46-73). New York: Simon & Schuster.

Jones, M. G., & Okey, J. R. (1995). *Interface Design for Computer-based Learning Environments*. [Online] Available at http://www.gsu.edu/~wwwitr/docs/idguide/ .

De Jong, T. (1991). Learning and instruction with computer simulations. *Education & Computing, 6*, 217-229.

De Jong, T. (1999). *De Proef de Som [The Proof or the Pudding]*. Inaugural Lecture, University of Twente, The Netherlands.

De Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, *31*, 105-113.

De Jong, T., De Hoog, R., & de Vries, F. Coping with complex environments: The effects of providing overviews and a transparent interface on learning with a computer simulation. *International Journal of Man-Machine Studies, 39*, 621-639.

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

De Jong, T., & Njoo, M. K. H. (1992). Learning and instruction with computer simulations: learning processes involved. In E. de Corte, M. Linn, H. Mandl, & L. Verschaffel, *Computer-based Learning Environments and Problem Solving*. Berlin, Germany: Springer-Verlag.

Van Joolingen, W. R. (1993). *Understanding and Facilitating Discovery Learning in Computer-based Simulation Environments.* PhD Thesis, University of Eindhoven, The Netherlands.

Van Joolingen, W. R., & Jong, T. de (1991). Characteristics of simulations for instructional settings. *Education & Computing, 6*, 241-262.

Van Joolingen, W. R., & De Jong (1997). An extended dual search space model of scientific discovery learning. *Instructional Science, 25*, 307-346.

Karassavidis, I. (1999). *Learning to Solve Correlational Problems: A Study of the Social and Material Distribution of Cognition*. PhD Thesis, University of Twente, The Netherlands.

Klahr, D. (2000). *Exploring Science: The Cognition and Development of Discovery Processes*. Cambridge, London: MIT Press.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12,* 1-48.

Klahr, D., Dunbar, K., & Fay, A. L. (1990). Designing good experiments to test bad hypotheses. In J. Shrager, & P. Langley, *Computational Models of Scientific Discovery and Theory Formation* (pp. 355-402). San Mateo, California: Morgan Kaufman.

Klahr, D., Fay, A. L., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology, 25*, 111-146.

Klahr, D., & Simon, H. A. (1999). Studies of scientific discovery: Complementary approaches and convergent findings. *Psychological Bulletin, 125*, 524-541.

Koschmann, T. (1996). Paradigm shifts and instructional technology: An introduction. In T. Koschmann, *CSCL: Theory and Practice of an Emerging Paradigm*. Mahwah, New Jersey: Erlbaum.

Koschmann, T., Kelson, A. C., Feltovich, P. J., & Barrows, H. S. (1996). Computer-supported problem-based learning: A principled approach to the use of computers in collaborative learning. In T. Koschmann, *CSCL: Theory and Practice of an Emerging Paradigm*. Mahwah, New Jersey: Erlbaum.

Kozma, R. B. (1994). *Will* media influence learning? Reframing the debate. *Educational Technology Research & Development, 42*, 7-19.

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

Kulkarni, D., & Simon, H. A. (1988). The processes of scientific discovery: The strategy of experimentation. *Cognitive Science*, *12*, 139-175.

Kulkarni, D., & Simon, H. A. (1990). Experimentation in machine discovery. In J. Shrager, & P. Langley, *Computational Models of Scientific Discovery and Theory Formation* (pp. 255-273). San Mateo, California: Morgan Kaufman.

Langley, P. (2000). The computational support of scientific discovery. *International Journal for Human-Computer Studies, 53*, 393-410.

Langley, P., Simon, H. A., Bradshaw, G. L., & Zytkow, J. M. (1987). *Scientific Discovery: Computational Explorations of the Creative Processes*. Cambridge, Massachusetts: MIT Press.

Langley, D., Ronen, M., & Eylon, B. (1997). Light propagation and visual patterns: Preinstruction learners' conceptions. *Journal of Research in Science Teaching, 34*, 399-424.

Lavoie, D. R., & Good, R. (1988). The nature and use of prediction skills in a biological computer simulation. *Journal of Research in Science Teaching, 25*, 335-360.

Leemkuil, H., De Jong, T., & Ootes, S. (2000). *Review of Educational Use of Games and Simulations*. KITS Consortium, Deliverable D1.

Leutner, D. (1993). Guided discovery learning with computer-based simulation games: Effects of adaptive and non-adaptive instructional support. *Learning and Instruction*, *3*, 113-132.

Liao, Y. C. (1998). Effects of hypermedia versus traditional instruction on Students' achievement: A meta-analysis. *Journal of Research on Computing in Education, 30*, 341-359.

Lindström, B., Marton, F., Ottosson, T., & Laurillard, D. (1993). Computer simulations as a tool for developing intuitive and conceptual understanding in mechanics. *Computers in Human Behavior*, *9*, 263-281.

Macaulay, L. (1995). *Human-Computer Interaction for Software Designers*. London: International Thomson Computer Press.

McKenzie, D. L., & Padilla, M. J. (1986). The construction and validation of the test of graphing in science (TOGS). *Journal of Research in Science Teaching*, *23*, 571-579.

Miller, C. S., Lehman, J. F., & Koedinger, K. R. (1999). Goals and learning in microworlds. *Cognitive Science, 23*, 305-336.

De Mul, S., & Van Oostendorp, H. (1996). Learning user interfaces by exploration. *Acta Psychologica, 91*, 325-344.

Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, New Jersey: Prentice-Hall.

Njoo, M. K. H. (1994). *Exploratory Learning with a Computer Simulation: Learning Processes and Instructional Support*. PhD Thesis, University of Eindhoven, The Netherlands.

Njoo, M. K. H., & De Jong, T. (1993). Exploratory learning with a computer simulation for control theory: Learning processes and instructional support. *Journal of Research in Science Teaching, 30*, 821-844.

Norman, D. A. (1993). *Things that Make Us Smart: Defending Human Attributes in the Age of the Machine*. New York: Voyager.

Novak, J.D. (1980). Methodological issues in investigating meaningful learning. In W.F. Archenhold, *Cognitive Development Research in Science and Mathematics: Proceedings of an International Seminar* (pp. 129-148). Leeds: Centre for Studies in Science Education, School of Education, University of Leeds.

Papert, S. (1980) *Mindstorms: Children, Computers and Powerful Ideas.* New York: Basic Books

Papert, S. (1987). Computer criticism versus technocentric thinking. *Educational Researcher*, january/february, 22-30.

Park, D. (1997). *The Fire within the Eye: A Hstorical Essay on the Nature and Meaning of Light*. Princeton, New Jersey: Princeton University Press.

Pea, R. D. (1992). Augmenting the discourse of learning with computer-based learning environments. in E. de Corte, M. C. Limm, H. Mandl, & L. Verschaffel, *Computer-based Learning Environments and Problem Solving*. Bonn, Germany: Springer-Verlag.

Peterson, N. S., Junck, J. R., Sharpe, D. M., & Finzer, W. F. (1987). A design approach to science. Simulated laboratories: learning via the construction of meaning. *Machine-mediated Learning, 2*, 111-127.

Pieters, J. M., & Van der Meij, H. (1994). *When support hampers: Constructivism and minimalism in supporting knowledge acquisition*. Paper presented at the AERA Annual Meeting, New Orleans, April 4-9.

Plötzner, R., & Spada, H. (1992). Analysis-based learning on multiple levels of mental domain representation. In E. De Corte, M.C. Linn, H. Mandl, & L. Verschaffel, *Computer-based Learning Environments and Problem Solving* (pp. 249-263). Berlin, Germany: Springer-Verlag (NATO ASI Series F, Vol. 84).

Popper, K. R. (1959). *The Logic of Scientific Discovery*. London: Hutchinson.

Qin, Y., & Simon, H. A. (1990). Laboratory replication of scientific discovery processes. *Cognitive Science, 14*, 281-312.

Reimann, P. (1990). *Problem Solving Models of Scientific Discovery Learning Processes*. Frankfurt, Germany: Peter Lang.

Reimann, P. (1991). Detecting functional relations in a computerized discovery environment. *Learning and Instruction, 1*, 45-65.

Rieber, L. P., & Parmley, M. W. (1995). To teach or not to teach? Comparing the use of computer-based simulations in deductive versus inductive approaches to learning with adults in science. *Journal of Educational and Computing Research, 13*, 359-374.

Salomon, G. (1992a). New challenges for educational research: Studying the individual within learning environments. *Scandinavian Journal of Educational Research, 3*, 167-182.

Salomon, G. (1992b). Effects *with* and *of* computers and the study of computer-based learning environments. In E. De Corte, M.C. Linn, H. Mandl, & L. Verschaffel, *Computer-based Learning Environments and Problem Solving* (pp. 249-263). Berlin, Germany: Springer-Verlag (NATO ASI Series F, Vol. 84).

Salomon, G., & Perkins, D. (1996). Learning in wonderland: What do computers really offer education? In S. T. Kerr, *Technology and the Future of Schooling. 95th Yearbook of the National Society for the Study of Education, Part II* (pp. 111-130). Chicago, Illinois: The University of Chicago Press.

Säljo, R. (1996). Mental and physical artifacts in cognitive practices. In P. Reimann, & H. Spada, *Learning in Humans and Machines: Towards an Interdisciplinary Learning Science* (pp. 83-96). New York: Pergamon Press.

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

Schauble, L., Glaser, R., Raghavan, K., & Reiner, M. (1991). Causal models and experimentation strategies in scientific reasoning. *The Journal of the Learning Sciences, 1*, 201-239.

Schoenfeld, A. H. (1992). On paradigms and methods: What do you do when the ones you know don't do what you want them to? Issues in the analysis of data in the form of videotapes. *The Journal of the Learning Sciences, 2*, 179-214.

Schunn, C. D., & Klahr, D. (1995). A 4-space model of scientific discovery. In J. D. Moore, & L. F. Lehman, *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 106-111). Mahwah, New Jersey: Erlbaum.

Schunn, C. D., & Klahr, D. (1996). Integrated yet different: Logical, empirical, and implementational arguments for a 4-space model of inductive problem solving. In G. Cottrell, *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 25-26). Mahwah, New Jersey: Erlbaum.

Shrager, J., & Klahr, D. (1986). Instructionless learning about a complex device: The paradigm and observations. *International Journal of Man-Machine Studies, 23*, 153-189.

Shrager, J., & Langley, P. (1990), *Computational Models of Scientific Discovery and Theory Formation*. San Mateo, California: Morgan Kaufmann.

Shute, V. J. (1991). A comparison of learning environments: All that glitters… In S. P. Lavoie, & S. J. Derry, *Computers as Cognitive Tools* (pp. 279-299). Hillsdale, New Jersey: Erlbaum.

Shute, V. J., & Gawlick-Grendell, L. A. (1994). What does the computer contribute to learning? *Computers and Education, 23*, 177-186.

Shute, V. J., & Glaser, R. (1990). A large-scale evaluation of an intelligent discovery world: Smithtown. *Interactive Learning Environments, 1*, 51-77.

Shute, V. J., Glaser, R., & Raghavan, K. (1989). Inference and discovery in an exploratory laboratory. In P. L. Ackerman, R. J. Sternberg, & R. Glaser, *Learning and Individual Differences: Advances in Theory and Research* (pp. 279-326). New York: Freeman.

Simon, H. A. (1973). Does scientific discovery have a logic? *Philosophy of Science, 12*, 471-480.

Simon, H. A. (1996). *The Sciences of the Artificial* (Rev. ed.). Cambridge, London: MIT Press.

Simon, H. A., & Lea, G. (1974). Problem solving and rule-induction: A unified view. In L.W. Gregg, *Knowledge and Cognition*. Hillsdale, New Jersey: Erlbaum.

Simon, H. A., Langley, P., & Bradshaw, G. L. (1981). Scientific discovery as problem solving. *Synthese, 47*, 1-27.

Stark, R., Renkl, A., Gruber, H., & Mandl, H. (1998). Indeed, sometimes knowledge does not help: A replication study. *Instructional Science, 26*, 391-407.

Swaak, J. (1998). *What-If: Discovery Simulations and Assessment of Intuitive Knowledge*. PhD Thesis, University of Twente, The Netherlands.

Swaak, J., & De Jong, T. (1996). Measuring intuitive knowledge in science: The development of the What-if test. *Studies in Educational Evaluation, 22*, 341-362.

Swaak, J., van Joolingen, W.R., & de Jong, T. (1998). Supporting simulation-based learning: The effects of model progression and assignments on definitional and intuitive knowledge. *Learning and Instruction, 8*, 235-253.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295-312.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction, 8*, 351-362.

Taylor, R. P. (1980). *The Computer in the School: Tutor, Tool, Tutee*. New York: Teachers College Press.

Tennyson, R. D. (1990). A proposed cognitive paradigm of learning for educational technology. *Educational Technology*, *30*, 16-19.

Tikhomirov, O.K. (1974). Man and computer: The impact of computer technology on the development of psychological processes. In D. R. Olson, *Media and Symbols: the Forms of Expression, Communication and Education*. The 73rd Yearbook of the National Society for the Study of Education, Part I (pp. 357-382). Chicago, Illinois: The University of Chicago Press.

Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development, 51*, 1-10.

Veenman, M. V. J. (1993). *Intellectual Ability and Metacognitive Skill: Determinants of Discovery Learning in Computerized Learning Environments*. PhD Thesis, University of Amsterdam, The Netherlands.

Veenman, M. V. J., Elshout, J. J., & Hoeks, J. C. J. (1993). Determinants of learning in simulation environments across domains. In D. Towne, T. de Jong, & H. Spada, *Simulation-based Experiential Learning* (pp. 235-249). Berlin, Germany: Springer-Verlag.

Veenman, M. V. J., Elshout, J. J., & Meijer, J. (1997). The generality vs domain-specifity of metacognitive skills in novice learning across domains. *Learning and Instruction, 7*, 187-209.

Veenman, M. V., Prins, F., & Elshout, J. J. (submitted). Initial inductive learning in a complex computer simulated environment: The role of metacognitive skills and intellectual ability. Article submitted for publication.

Vygotsky, L. S. (1987). Thinking and speech. In R. S. Rieber, & A. S. Carton, *The Collected Works of L.S. Vygotsky. Vol. 1. Problems of General Psychology* (Transl. by N. Minick). New York: Plenum Press.

Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology*, *49*, 436-458.

Wagner, R. K., & Sternberg, R. J. (1986). Tacit knowledge and intelligence in the everyday world. In: R. J. Sternberg, & R. K. Wagner, *The Nature and Origin of Competence in the Everyday World*. Cambridge, England: Cambridge University Press.

Wilhelm, P., Van Rijn, H., Beishuizen, J. J., & Niewold, P. (2001). *Studying Self-directed Inductive Learning*. Manuscript in preparation.

Wittrock, M. C. (1966). The learning by discovery hypothesis. In L. S. Shulman, & E. R. Keislar, *Learning by Discovery, a Critical Appraisal* (pp. 33-75). Chicago, Illinois: Rand McNally.

Nederlandse samenvatting

Een belangrijk onderdeel van iets ontdekken is dat er sprake is van een verrassingselement: er treedt een gebeurtenis op die men niet had zien aankomen. Het ontdekken van nieuwe informatie leidt tot een kwalitatieve verandering in kennis. Van dit kenmerk kan gebruik worden gemaakt in het onderwijs. Dit is wat plaatsvindt bij de manier van leren die *ontdekkend leren* wordt genoemd. Ontdekkend leren is een leervorm waarvan de populariteit sinds een aantal jaar sterk is toegenomen. Een reden hiervoor is dat nieuwe toepassingen zijn ontwikkeld waarmee ontdekkend leren gestimuleerd kan worden. Een van die toepassingen is een computersimulatie. Een computersimulatie is een programma dat een systeem simuleert door middel van een model. Tijdens het werken met een simulatie kan een leerling variabelen manipuleren en het effect hiervan waarnemen. De uitkomsten van manipulaties worden bepaald door het model van de simulatie, dat niet zichtbaar is voor de leerling. Het doel van ontdekkend leren met een computersimulatie is om door middel van interactie met de computer de regels waaraan het model voldoet te 'ontdekken'.

Het onderwerp van dit proefschrift is de invloed van voorkennis op ontdekkend leren. De manier waarop ontdekkend leren wordt onderzocht is met behulp van een computersimulatie. In het kader van dit onderzoek werden tests ontworpen om voorkennis op een adequate wijze te meten. Tevens werden methoden ontwikkeld om het gedrag van leerlingen tijdens het werken met een computersimulatie te registreren. In het kader van het onderzoeksproject 'Inductief Leren' werd onder andere een computersimulatie ontwikkeld die als onderwerp geometrische optica heeft. De simulatie (genaamd 'Optica') werd in alle experimenten die voor dit onderzoek zijn uitgevoerd gebruikt. Drie experimenten werden uitgevoerd (in 1997, 1999, en in 2000).

In hoofdstuk een worden de beweegredenen die ten grondslag liggen aan het onderzoek besproken. De veranderingen die in het onderwijs gaande zijn kunnen verklaard worden vanuit een praktisch en een theoretisch perspectief. In praktische zin speelt de enorme technologische vooruitgang van de laatste jaren een belangrijke rol. Computers zijn niet meer weg te denken uit het onderwijs, en verantwoord gebruik ervan vereist een andere opstelling van zowel leraren als leerlingen ten opzichte van meer traditionele leermiddelen. In theoretische zin speelt de veranderde kijk op het leerproces een rol. Een groep theorieën die hierbij centraal staat kan worden samengenomen onder de naam *constructivisme*. Vanuit het constructivistische perspectief worden leerlingen niet gezien als passieve ontvangers van informa-

tie, maar als actieve deelnemers aan het leerproces. De rol van de leraar is hierbij niet slechts het overbrengen van informatie naar leerlingen toe, maar het faciliteren van het leerproces. Het idee van ontdekkend leren sluit zeer nauw aan bij deze ideeën. Met behulp van een computersimulatie kan deze leervorm bij leerlingen worden 'uitgelokt'. Er is veel onderzoek verricht naar ontdekkend leren. Veel onderzoek beperkt zich echter tot het bekijken van de effectiviteit van ontdekkend leren in relatie tot andere leervormen. Het blijkt dat een groot aantal factoren het succes van ontdekkend leren bepalen. In meerdere studies is gevonden dat de kennis die leerlingen bezitten voordat ze aan een leertaak beginnen een belangrijke factor is.

In het in dit proefschrift beschreven onderzoek staat het leer*proces* tijdens ontdekkend leren centraal. De algemene onderzoeksvraag was: op welke wijze beïnvloedt reeds aanwezige kennis het leerproces tijdens ontdekkend leren met computersimulaties? Meer specifiek geformuleerd werden de volgende drie onderzoeksvragen gesteld:

Op welke manier beïnvloeden domeinspecifieke en algemene kennis de processen die tijdens ontdekkend leren plaatsvinden? Verschilt de wijze waarop beide typen kennis het leren beïnvloeden?

Beïnvloeden domeinspecifieke en algemene kennis de effectiviteit van ontdekkend leren? Leidt meer voorkennis tot een beter leerresultaat?

Treden veranderingen in het ontdekkend leerproces op over de tijd? Is de leerstrategie die wordt gevolgd afhankelijk van domeinspecifieke en algemene kennis?

In de empirische studies die in het kader van dit proefschrift werden uitgevoerd werd getracht tot een antwoord op deze vragen te komen.

In hoofdstuk twee wordt de theoretische achtergrond van het onderzoek geschetst. Ontdekkend leren kan worden beschreven als een speciale vorm van probleemoplossen. In de theorie die door Newell en Simon in 1972 werd gepostuleerd wordt probleemoplossen gezien als een proces waarbij een 'probleemruimte' wordt doorlopen. Klahr en Dunbar (1988) hebben een model opgesteld om processen die tijdens ontdekken plaatsvinden op eenzelfde manier te verklaren. In dit 'sdds'-model wordt ontdekkend leren beschreven als het doorzoeken van twee probleemruimtes: een hypotheseruimte (die bestaat uit alle mogelijke relaties tussen variabelen binnen een domein) en een experimentruimte (die bestaat uit alle experimenten die mogelijk zijn binnen een domein). Van Joolingen en De Jong (1997) hebben dit model verder verfijnd. In de hypotheseruimte onderscheiden zij de *leerling hypotheseruimte*, de *effectieve leerruimte*, en het *conceptuele doelmodel*. De leerling hypotheseruimte is opgebouwd uit de relaties en variabelen in een domein waar

een leerling kennis over heeft. Domeinspecifieke kennis is dus bepalend voor de opbouw van de leerling hypotheseruimte. Algemene kennis bepaalt welke wiskundige relaties iemand wel en niet kent. Het bezitten van meer algemene kennis betekent kennis over meer typen relaties die kunnen voorkomen, wat impliceert dat een leerling met meer algemene kennis meer mogelijkheden zal zien om relaties in een leeromgeving te onderzoeken.

De beschrijving van voorkennis in termen van het sdds-model leidt tot de voorspelling dat deze kennis het ontdekkend leerproces zal beïnvloeden. In een aantal experimenten werd onderzocht op welke wijze dit plaatsvindt.

De computersimulatie die in elk van deze experimenten werd gebruikt om ontdekkend leerprocessen te stimuleren was de Optica simulatie. Deze simulatie, die voor het Inductief Leren-project werd ontworpen, wordt beschreven in hoofdstuk drie. Als onderwerp voor de simulatie werd gekozen voor het onderwerp geometrische optica. Dat wil zeggen dat het model van de simulatie formules bevat die in de geometrische optica een rol spelen (zoals bijvoorbeeld de lenzenformule). De interface van de Optica simulatie is zo ontworpen dat met de muis van de computer experimenten kunnen worden uitgevoerd. Lenzen en lampjes kunnen in een werkgebied worden gepositioneerd, waarbij lichtbreking van een lichtstraal door een lens door de simulatie wordt weergegeven. Het updaten van de simulatie gebeurt in real-time, wat als consequentie heeft dat er geen duidelijke scheiding tussen experimenten in de simulatie aan te wijzen is. Ook kwantitatieve metingen kunnen worden verricht in de Optica simulatie, door middel van het uitvoeren van afstandsmetingen. Om het gedrag van leerlingen tijdens het werken met de Opticasimulatie te meten werd gebruik gemaakt van een methode waarbij alle operaties ('muisklikken') die ze in de simulatie uitvoerden werden geregistreerd.

In hoofdstuk vier wordt de manier waarop de frequentie en volgorde van operaties beschreven. Tevens wordt een methode geïntroduceerd om frequentie en volgorde van operaties in een maat te combineren. In hoofdstuk vier worden ook de verschillende tests die in de experimenten worden gebruikt onder de loep genomen. Domeinspecifieke kennis en algemene kennis worden in twee tests geoperationaliseerd. De items van de test voor algemene kennis vragen naar kennis over verschillende typen wiskundige relaties. Om domeinspecifieke voorkennis te testen werd gebruik gemaakt van de zogenaamde 'what-if'-test, een test voor intuïtieve kennis. Deze test combineert snelheid en accuraatheid door leerlingen te vragen om deze twee fac-

toren als even belangrijk te beschouwen tijdens het maken van de test. In de experimenten worden varianten van de beschreven tests gebruikt.

De hoofdstukken vijf, zes, en zeven zijn gewijd aan de empirische studies die in het kader van het proefschrift werden uitgevoerd.

In hoofdstuk vijf wordt het Optica 97 experiment besproken. Het doel van dit experiment was om licht te werpen op de eerder gestelde onderzoeksvragen. Hierom werd in dit experiment gebruik gemaakt van twee computersimulaties: de Optica simulatie (waarover variatie in voorkennis werd verwacht) en de Bubbles simulatie (een simulatie van een door de onderzoekers bedacht systeem, waarover geen domeinspecifieke voorkennis aanwezig kon zijn bij leerlingen). Algemene kennis werd vooraf gemeten, domeinspecifieke kennis werd zowel vooraf gemeten als gemanipuleerd: kennis over geometrische optica werd gemeten, kennis over Bubbles werd gemanipuleerd. De resultaten van het experiment lieten zien dat de invloed van domeinspecifieke voorkennis op het ontdekkend leerproces nihil was, terwijl slechts een geringe invloed van algemene kennis werd gevonden. Het bleek dat leerlingen zeer weinig voorkennis over optica bezaten, en dat de betrouwbaarheid van de gebruikte tests laag was. Ook de variatie in testscores was laag, wat indeling in groepen met veel en weinig voorkennis bemoeilijkte. Analyse van het aantal verrichte operaties over de tijd liet wel verschillen zien tussen leerlingen met weinig en leerlingen met veel voorkennis, maar deze verschillen waren niet eenduidig te verklaren.

Om meer duidelijkheid over de gevonden effecten te krijgen (en over de afwezigheid van een effect van domeinspecifieke voorkennis) werd het Optica 99 experiment uitgevoerd. Dit experiment wordt in hoofdstuk zes beschreven. Net als in het Optica 97 experiment werden in dit experiment algemene kennis en domeinspecifieke voorkennis (over optica) gemeten. Ook werd in dit experiment een test gebruikt om ontdekkingsvaardigheden te meten (de 'Petertest'). Door twee groepen proefpersonen (afkomstig uit VWO en MBO) te combineren werd getracht meer variatie in testscores te verkrijgen. In dit experiment werd alleen de Opticasimulatie gebruikt. De simulatie werd in vier verschillende 'situaties' gedeeld (verschillend in opbouw en beschikbare functies), waarmee elke leerling korte tijd werkte. De resultaten lieten zien dat leerlingen met veel algemene kennis consistent actiever waren in het leerproces dan leerlingen met weinig algemene kennis. Wat betreft domeinspecifieke voorkennis werd alleen een verschil gevonden in een van de vier situaties: leerlingen met veel domeinspecifieke voorkennis waren minder actief in deze situatie dan leerlingen met weinig voorkennis. Voor ontdekkingsvaardigheden werd alleen een verschil gevonden in de

eerste (eenvoudige) situaties, waarbij leerlingen met goede vaardigheden actiever waren dan leerlingen met minder goede vaardigheden.

Het verschil dat in het Optica 99 experiment werd gevonden voor domeinspecifieke voorkennis werd nader bekeken in het laatste experiment, Optica 2000. Dit experiment wordt beschreven in hoofdstuk zeven. Ook nu werden domeinspecifieke en algemene kennis vooraf gemeten, alsmede ontdekkingsvaardigheden. Voor het ontdekkend leren werd gebruik gemaakt van een situatie in de Opticasimulatie. Het experiment verschilde van de voorgaande experimenten, omdat kennis over geometrische optica werd gemanipuleerd. De manier waarop dat gebeurde was door middel van beschikbaar stellen van informatie ('kennistips'). De informatie was tijdens het werken met de Opticasimulatie beschikbaar, en kon door leerlingen (in dit experiment allen afkomstig uit het MBO) worden aangeroepen. Een deel van de leerlingen had de informatie tot haar beschikking, het andere deel kreeg alleen de standaardinstructie vooraf.

De resultaten lieten zien dat het beschikbaar hebben van informatie tijdens de ontdekkend leertaak een positief leereffect tot gevolg had. Dit in tegenstelling tot de conditie zonder hulp, die een negatief leereffect liet zien. Overeenkomstig met het Optica 99 experiment werd gevonden dat leerlingen met veel domeinspecifieke voorkennis minder actief waren in de simulatie dan leerlingen met weinig voorkennis. De manipulatie door middel van het aanbieden van kennis, gekoppeld aan het kleine aantal proefpersonen, belemmerde analyse van de invloed van algemene kennis en ontdekkingsvaardigheden.

In hoofdstuk acht worden een aantal algemene conclusies getrokken met betrekking tot datgene wat in het voorgaande werd besproken en wat in de uitgevoerde experimenten werd gevonden. De discussie spitst zich toe op de tests voor domeinspecifieke kennis, algemene kennis, en ontdekkingsvaardigheden, op de gevolgen van de onderzoeksresultaten voor het sdds-model van ontdekkend leren, het effect van het geven van ondersteuning tijdens ontdekkend leren, en conclusies naar aanleiding van de analyse van leerling-computer interactiegedrag. Een nadeel van de analyse van interactiedata lijkt te zijn dat de grote variatie in individueel ontdekkend leergedrag veel verschillen doet wegvallen bij het aggregeren van gegevens. Om verschillen in gedrag te vinden tussen leerlingen met veel en weinig voorkennis moeten de scores ver uiteen liggen. Wat betreft domeinspecifieke kennis lijkt het erop dat deze alleen van invloed is op het ontdekkend leerproces als de kennis

tijdens het werken met een computersimulatie kan worden aangesproken. Ondersteuning door middel van het aanbieden van kennis tijdens het leren kan hierbij helpen.

Het uitgevoerde onderzoek heeft implicaties voor het onderwijsveld en voor de uitvoering van verder onderzoek. Het onderzoek laat zien dat het belangrijk is om het verband tussen de leertaak die wordt gebruikt (in dit geval ontdekkend leren met een computersimulatie) en het type test dat wordt gebruikt in de gaten te houden. Het onderzoek laat ook zien dat leerlingen minder gemotiveerd zijn om over een 'school'-onderwerp als optica te leren dan over een 'spel'-onderwerp als Bubbles. In verder onderzoek zou de relatie tussen het onderwerp van een simulatie en de motivatie van leerlingen kunnen worden onderzocht. Ook kan de analyse van leerling-computer interactie verder worden verfijnd, iets wat uiteindelijk kan resulteren in een algemene manier om het gedrag van leerlingen in een computersimulatie te meten.

Appendix A

Number and graph test (generic knowledge test)

This test is meant to determine your knowledge about mathematical relations in graphs and numbers. There are **32** questions in total. Read each question carefully, and choose what you think is the right answer. For most questions, the answer can be chosen from a list of alternative answers, by marking the square in front of the right answer. For some questions, a short mathematical relation has to be filled in. *Only one answer is correct for each question*.

If you are not sure about the answer to a problem, skip it for the time being, and continue with the other questions. In any case, try to answer *as many questions as possible* within the time given.

1. Consider the following relation:

$$A^2 + B^2 = C^2$$

Under what name is this relation also known?

- □ The lens law
- □ The law of Snellius
- ■ The rule of Pythagoras
- □ None of the relations above

2. In the questions below you will see two columns with numbers, one marked X and another marked Y (or A and B). The numbers under X and Y are related to each other. On the right you see some relations between the numbers, only one of which is correct. Your task is to mark the correct relation.

What type of relation is represented in the table below?

X	Y	*Answer*
1	1	□ $Y = 2X - 1$
2	.5	□ $Y = \frac{1}{X^2}$
3	.3333…	
4	.25	■ $Y = \frac{1}{X}$
		□ $Y = \sqrt{X}$

3. What type of relation is represented in the table below?

X	Y	*Answer*	
2	6	□	geometrical relation
4	12	■	linear relation
6	18	□	coincidental relation
8	24	□	exponential relation

4. What type of relation is represented in the table below?

X	Y	*Answer*	
1	1	□	$Y = \frac{1}{2}X + 1$
2	1.4142		
3	1.7321…	□	$Y = \frac{1}{X^2}$
4	2		
		□	$Y = \frac{1}{X}$
		■	$Y = \sqrt{X}$

5. The table below shows a relation between the variables X, Y, and Z. In the answer column, write in one formula how X can be computed from Y and Z.

X	Y	Z	*answer*
.5	1	2	$X = \frac{Y}{Z}$
.666…	2	3	
.8	4	5	
.857…	6	7	

6. How can the relation that is represented in the table below be described?

A	B	*Answer*	
10	4	□	The more A increases, the slower B decreases
20	8		
30	16	□	When A approaches zero, B also approaches zero
60	64	■	The more A decreases, the faster B decreases
100	1024		
		□	The more A increases, the slower B increases

7. How can the relation that is represented in the table below be described?

A	B	*Answer*	
10	10	■	The more A increases, the slower B decreases
20	5		
30	3	□	When A approaches zero, B approaches the infi nite
60	2.5	□	The more A increases, the slower B increases
100	2.25		
		□	The more A increases, the faster B decreases

8. The table below shows a relation between the variables A and B. In the answer column, write in one formula how B can be computed from A.

A	B	*Answer*
2	.5	
4	.25	$B = \frac{1}{A}$
6	.1666…	
8	.125	

9. How can the relation that is represented in the table below be described?

A	B	*Answer*	
4	.0625	□	$B = \sqrt{A} - 2$
6	.0277…	□	$B = \frac{1}{A}$
8	.0156…		
10	.01	■	$B = \frac{1}{A^2}$
		□	$B = \frac{2A}{7}$

10. What type of relation is represented in the table below?

A	B	*Answer*	
4	2	■	$B = \sqrt{A}$
6	2.4494…	□	$B = \frac{A}{2\frac{1}{2}}$
8	2.8284…		
10	3.1622…	□	$B = \frac{1}{A^2}$
		□	$B = 2A$

11. The table below shows a relation between the variables X, Y, and Z. In the answer column, write in one formula how X can be computed from Y and Z.

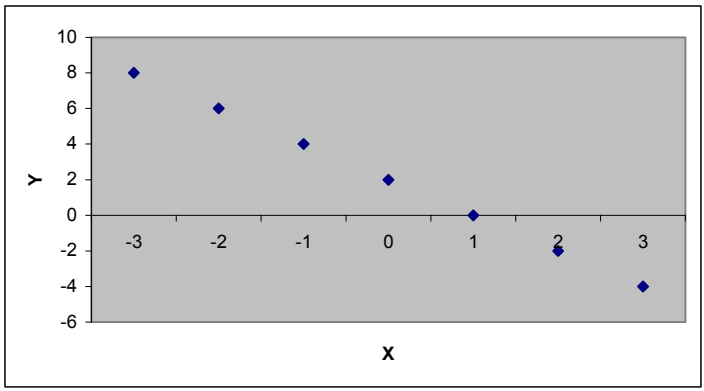
X	Y	Z	*Answer*
0	10	20	□ X only influences Y, not Z
5	10	20	□ X only influences Z, not Y
10	10	20	□ X influences both Y and Z
15	10	20	■ X does not influence Y and Z
20	10	20	

12. If X, Y, and Z from the previous question would represent amounts of fluid, how could the influence of fluid X on Y and Z then be called?

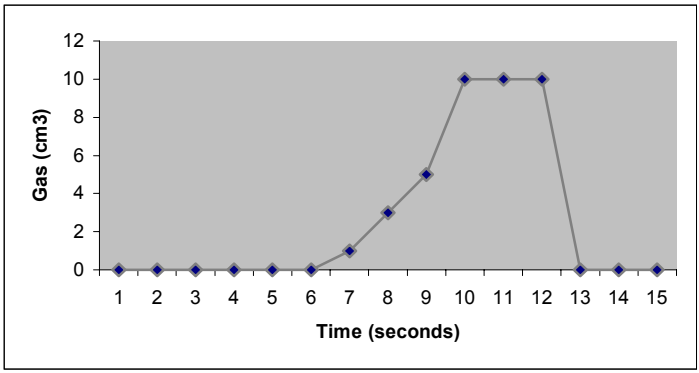
□ X is a moderator
□ X is a katalyst
□ X is an oxydizer
■ X does not influence Y and Z

13. What type of relation is represented in the graph below?

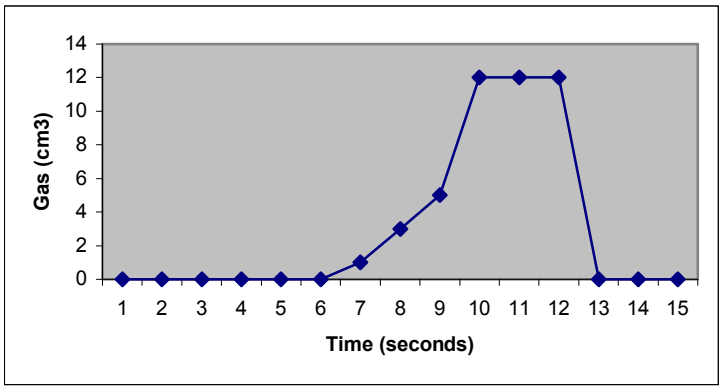
□ A quadratic function
□ A function with an aymptote
■ A monotonic decreasing function
□ A logaritmic function



14. In an experiment two fluids A and B are mixed; a certain amount of gas appears from this reaction. When the amount of gas that appears in time is checked, the following graph can be drawn:



Now, the experiment is repeated, but also a fluid C is added to the mix. Again, gas appears, giving the following graph:



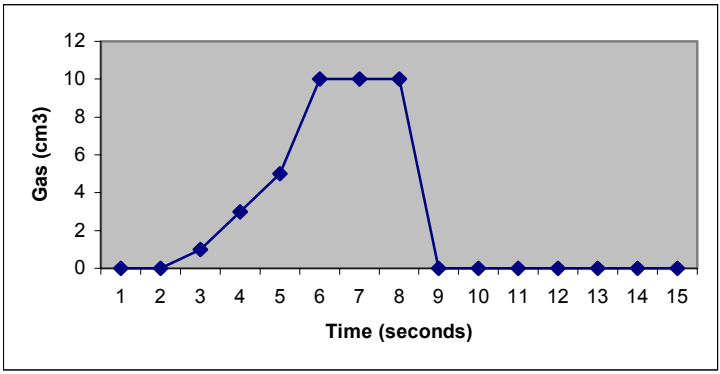
What influence does adding fluid C have for the amount of gas?

- ☐ The reaction happens with more difficulty
- ■ More gas appears during the reaction
- ☐ The reaction starts earlier
- ☐ Nothing changes about the reaction

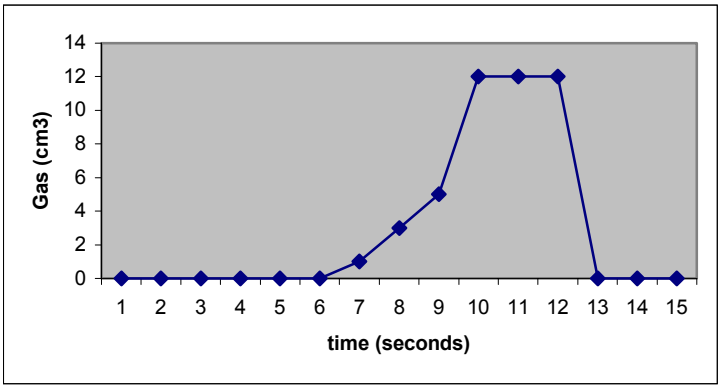
15. See the graphs in the previous question. Instead of fluid C, a fluid D is added to the mixture of A and B. Checking the amount of gas that appears, the following graph is drawn:

What influence does adding fluid D have for the amount of gas?

- ☐ The reaction happens with more difficulty
- ☐ More gas appears during the reaction
- ■ The reaction starts earlier
- ☐ Nothing changes about the reaction



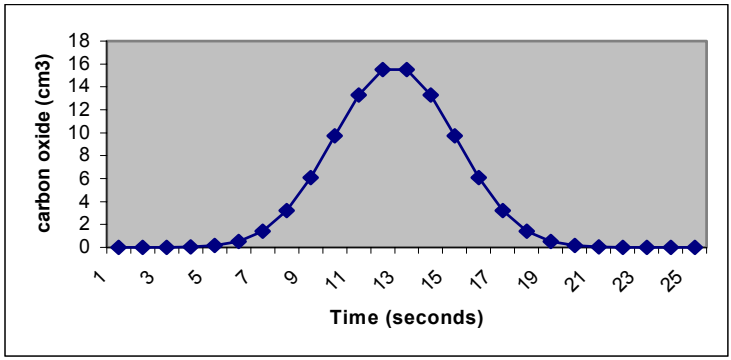
16. Again to the mixture of A and B fluid C is added, but twice as much as in question 14. The amount of gas that appears gives the following graph:



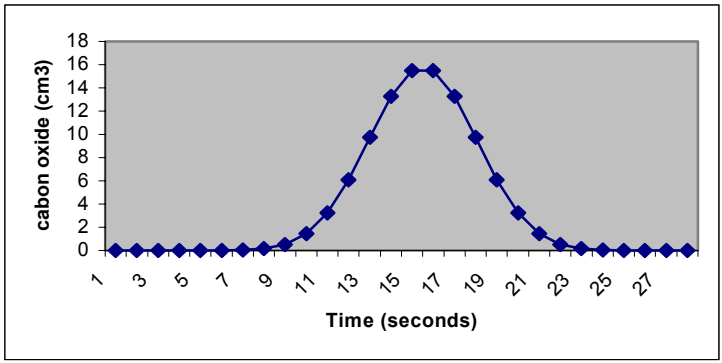
What conclusion can be drawn about adding fluid C to the mixture?

- ■ C sometimes does, and sometimes doesn't influence the reaction
- ☐ The more there is of C, the faster the reaction takes place
- ☐ The more there is of C, the more gas appears in the reaction
- ☐ C has only a limited influence on the reaction

17. In an experiment two fluids A and B are mixed; a certain amount of carbon oxide appears from this reaction. When the amount of gas that appears in time is checked, the following graph can be drawn:



Now this experiment is repeated, but at the same time the mixture is heated. Again, carbon oxide appears, and the following graph can be drawn:



What conclusion can be drawn about the influence of heating the mixture?

- □ The reaction takes a shorter time
- □ More carbon oxide appears during the reaction
- □ The reaction takes a longer time
- ■ The reaction starts later

18. Consider the statement: the more A increases, the faster B increases. What relation is described?

- ■ $B = 2^A$
- □ $A = 2^B$
- □ $B = 2\ A^2$
- □ $A = 2\ B^2$

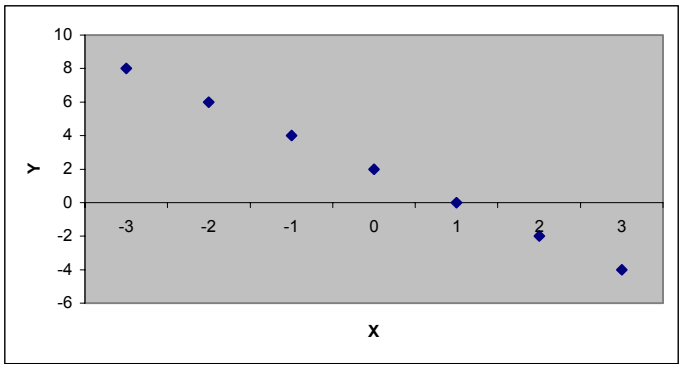
19. Consider the statement: the more A increases, the slower B decreases. What relation is described?

- □ $B = \sqrt{A}$
- ■ $B = \frac{1}{A}$
- □ $B = -A^2$
- □ $B = \frac{A}{5}$

20. Consider the statement: the more A decreases, the faster B increases. What relation is described?

- □ $B = -A^2$
- □ $B = \sqrt{A}$
- □ $A = \frac{1}{B}$
- ■ $B = 2^{-A}$

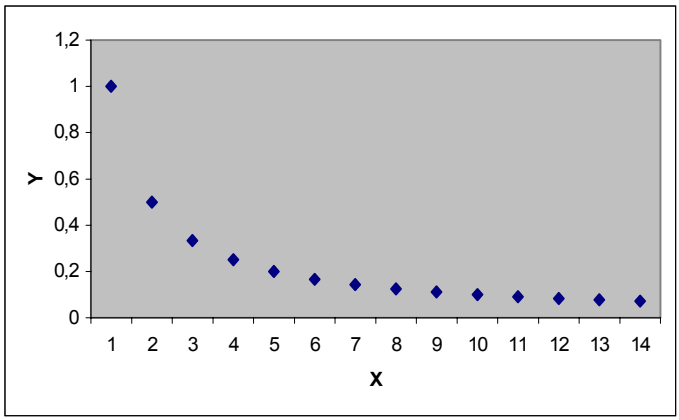
21. What type of relation is represented in the graph below?



- □ $Y = -X$
- ■ $Y = -2X + 2$
- □ $Y = X + 2$
- □ $Y = -X - 2$

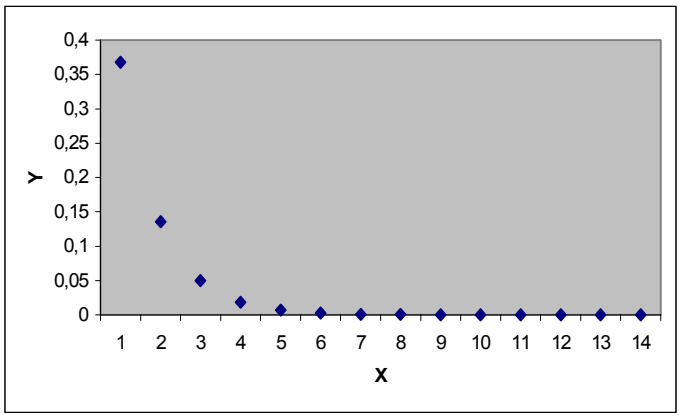
22. What type of relation is represented in the graph below?

- ☐ $Y = -X^2$
- ☐ $Y = e^{-X}$
- ■ $Y = \frac{1}{X}$
- ☐ $Y = \frac{1}{X^2}$



23. What type of relation is represented in the graph below?

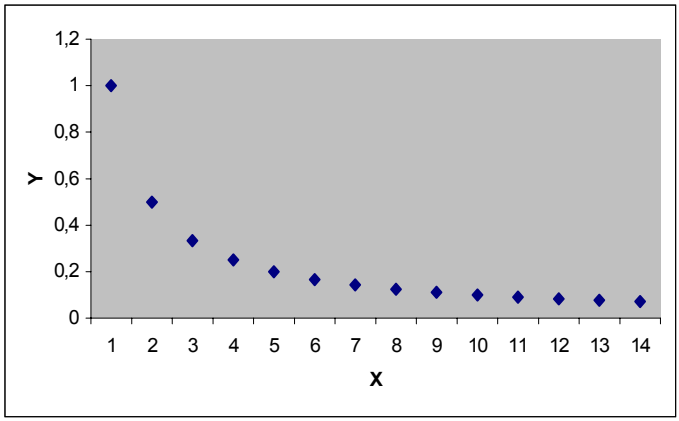
- ■ $Y = e^{-X}$
- ☐ $Y = -X^2$
- ☐ $Y = \frac{1}{X}$
- ☐ $Y = \frac{1}{X^2}$



24. What type of relation is represented in the table below?

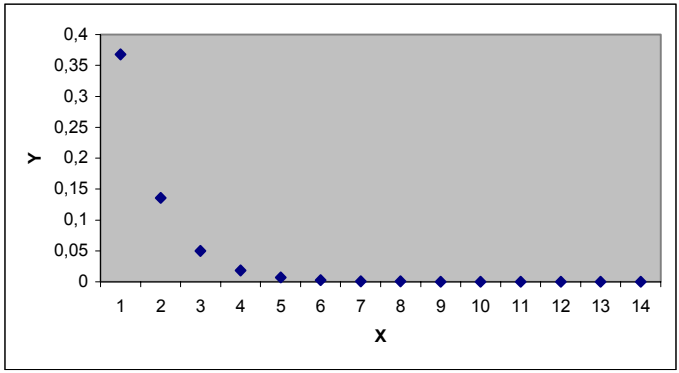
A	B	*Answer*	
3	8	□	linear relation
5	24	■	quadratic relation
7	48	□	exponential relation
9	80	□	none of the relations above

25. What type of relation is represented in the graph below?



■ A function with two asymptotes
□ A monotonous increasing function
□ A logaritmic function
□ A quadratic function

26. What type of relation is represented in the graph below?



- □ A quadratic function
- □ A function with asymptotes
- □ A monotonous increasing function
- ■ A logaritmic function

27. What type of relation is represented in the table below?

X	Y	*Answer*
1	1	□ $Y = 2X$
2	.25	■ $Y = \frac{1}{X^2}$
3	.1111…	□ $Y = \frac{1}{X}$
4	.0625	□ $Y = \sqrt{X}$

28. The table below shows a relation between the variables X, Y, and Z. In the answer column, write in one formula how X can be computed from Y and Z.

X	Y	Z	*answer*
3	1	1	
6	2	2	$X = Y + 2Z$
9	3	3	
12	4	4	

29. What type of relation is represented in the table below?

X	Y	*Answer*	
1	2	□	quadratic relation
2	4	■	exponential relation
3	8	□	geometrical relation
4	16	□	linear relation

30. What type of relation is represented in the table below?

A	B	*Answer*	
-5	32	□	The more A increases, the faster B increases
-3	8	□	When A approaches zero, B approaches the infinite
-1	2	■	The more A decreases, the faster B increases
1	.5	□	When A approaches zero, B also approaches zero
3	.125		

31. The table below shows a relation between the variables X, Y, and Z. In the answer column, write in one formula how X can be computed from Y and Z.

X	Y	Z	*Answer*	
5	10	10	□	X only influences Y, not Z
10	10	15	■	X only influences Z, not Y
15	10	20	□	X influences both Y and Z
20	10	25	□	X does not influence Y and Z
25	10	30		

32. What type of relation is represented in the table below?

A	B	*Answer*	
4	10	□	linear relation
6	10	□	quadratic relation
8	20	□	exponential relation
10	20	■	none of the relations above

Appendix B

Optics test (domain-specific knowledge test)

This test is meant to determine your knowledge about lenses and light. Read each question carefully, and choose what you think is the right answer. For most questions, the answer can be chosen from a list of alternative answers, by marking the square in front of the right answer. *Only one answer is correct for each question*.

If you are not sure about the answer to a problem, skip it for the time being, and continue with the other questions. In any case, try to answer *as many questions as possible* within the time given.

During the test, you are allowed to write down notes.

Most of the questions refer to a numbered figure. Sometimes, distances in centimeters are presented in a figure. Also, the following objects are used in the figures:

- lamps

A lamp with one lightbeam. The lightbeam is represented as a black line.

A set of three lamps. All three lightbeams are shining in the same direction.

A set of three lamps. All three lightbeams are shining in the different directions.

- lenses

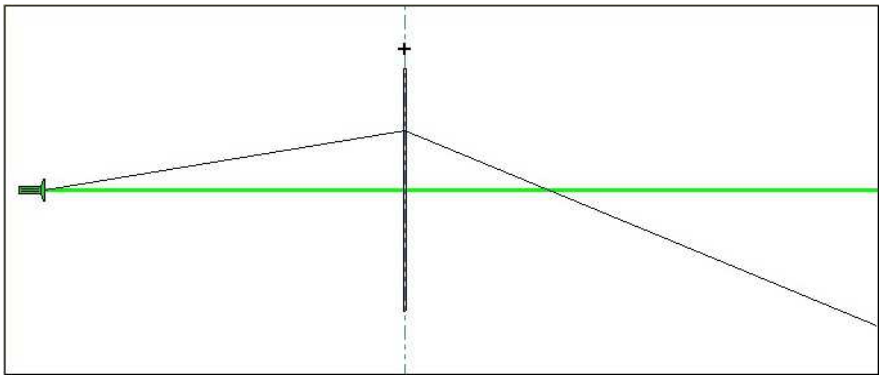
Different types of lenses have different properties

- big lamp and screens

The big lamp sends lightbeams in all directions, through a plate with holes (L-shaped). Light that goes through the holes is projected through a lens on a screen.

Note: The word 'baseline' that is used in most of the questions refers to the thick horizontal line that is displayed in all the figures.

Figure 1



1. See Figure 1. If the lens is moved a little to the left, then where will the lightbeam to the right of the lens strike the baseline?

- ❑ Closer to the lens
- ❑ The same distance to the lens
- ❑ Further away from the lens

2. See Figure 1. If the lightbeam is pointed up a little bit more, then where will the lightbeam to the right of the lens strike the baseline?

- ❑ Closer to the lens
- ❑ The same distance to the lens
- ❑ Further away from the lens

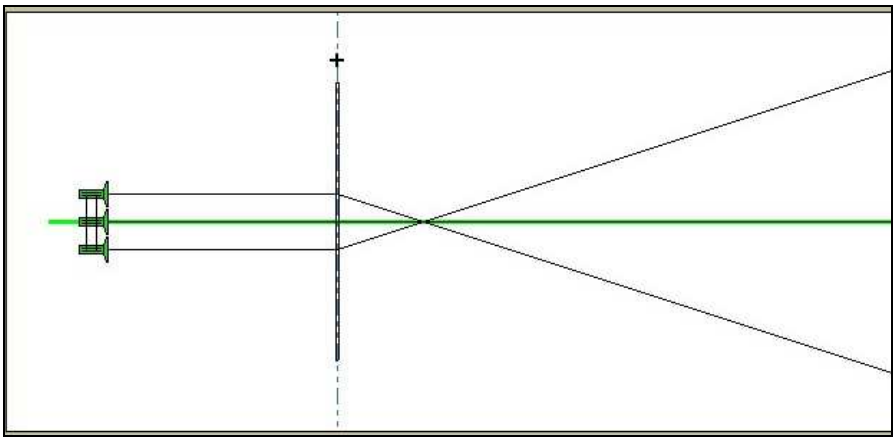
3. See Figure 2. If the lamps are moved a little to the right, then where will the light-beams to the right of the lens strike the baseline?

- ❑ Closer to the lens
- ❑ The same distance to the lens
- ❑ Further away from the lens

4. See Figure 2. If the lamps are moved a little to the left, then where will the lightbeams to the right of the lens strike the baseline?

- ❑ Above the baseline
- ❑ On the baseline
- ❑ Below the baseline

Figure 2



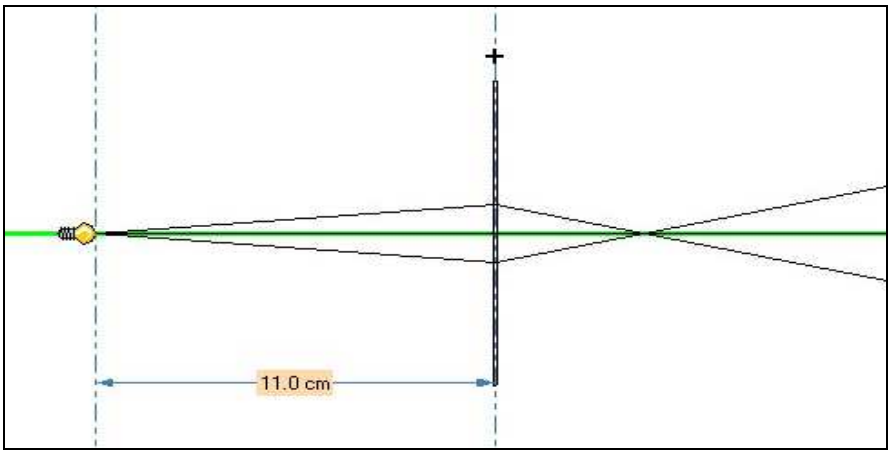
5. See Figure 2. If the lightbeams are pointed up a little bit more, then where will the lightbeams to the right of the lens strike the baseline?

- ❑ Above the baseline
- ❑ On the baseline
- ❑ Below the baseline

6. See Figure 2. If the lightbeams are pointed up a little bit more, then where will the lightbeams to the right of the lens strike the baseline?

- ❑ Closer to the lens
- ❑ The same distance to the lens
- ❑ Further away from the lens

Figure 3



7. See Figure 3. If the lamps are moved a little to the left, then what happens to the point where the lightbeams converge?

- ❑ The convergence point gets closer to the lens
- ❑ The convergence point gets further away from the lens
- ❑ The convergence point stays the same distance from the lens

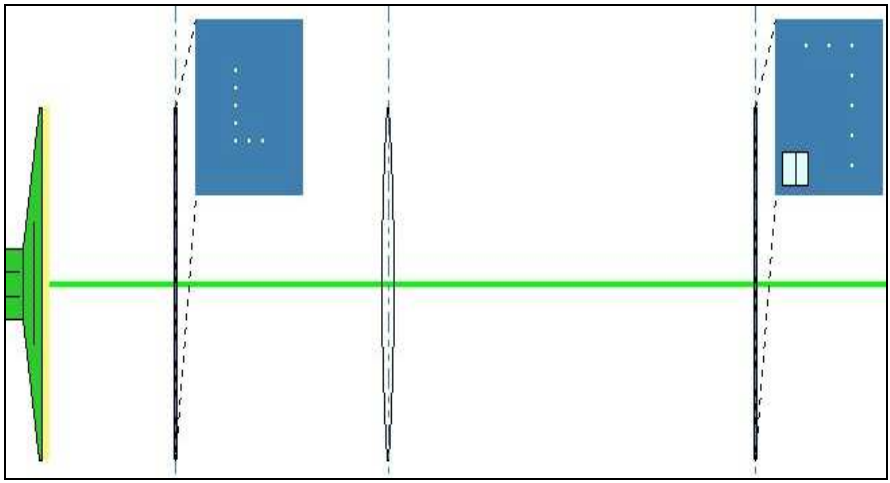
8. See Figure 3. If the lightbeams are pointed up a little bit more, then what happens to the point where the lightbeams converge?

- ❑ The convergence point moves above the baseline
- ❑ The convergence point stays on the baseline
- ❑ The convergence point moves below the baseline

9. See Figure 3. If the lens in the figure is replaced by a lens with a focal distance which is twiece as large, what happens to the convergence point?

- ❑ The convergence point gets closer to the lens
- ❑ The convergence point gets further away from the lens
- ❑ The convergence point becomes a virtual point

Figure 4



10. See Figure 4. What is the distance of the plate with holes (on the left) to the lens called?

- ❑ Object distance (v)
- ❑ Focal distance (f)
- ❑ Image distance (b)

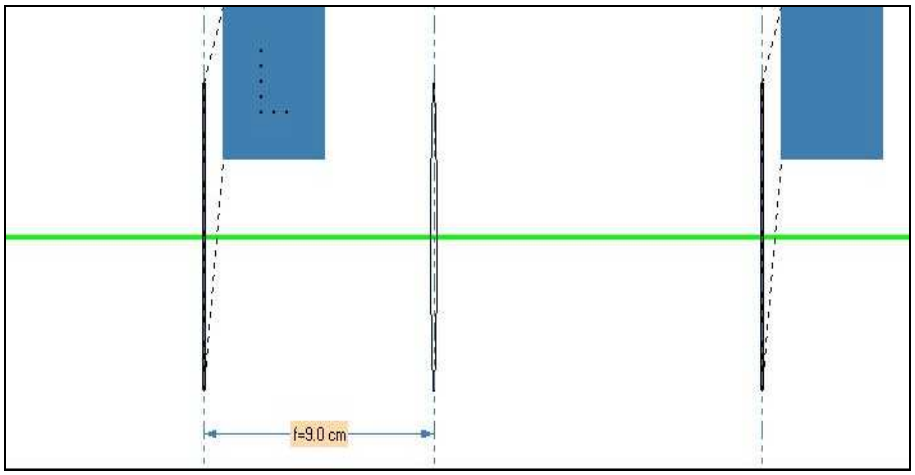
11. See Figure 4. What is the distance of the screen (on the right) to the lens called?

- ❑ Object distance (v)
- ❑ Focal distance (f)
- ❑ Image distance (b)

12. See Figure 5. The object distance is made equal to the image distance. Will the projected image be sharp?

- ❑ Yes, but the image will be virtual
- ❑ Yes
- ❑ No

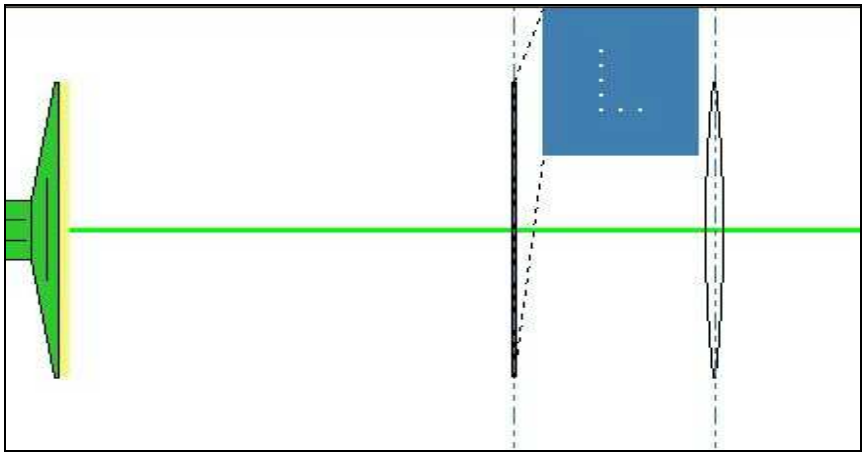
Figure 5



13. The magnification formula describes the relation between the size of a projected image and the size of an object. What does the formula look like?

- ❑ Magnification = focal disctance/size of object
- ❑ Magnification =size of object/size of image
- ❑ Magnification =size of image/size of object

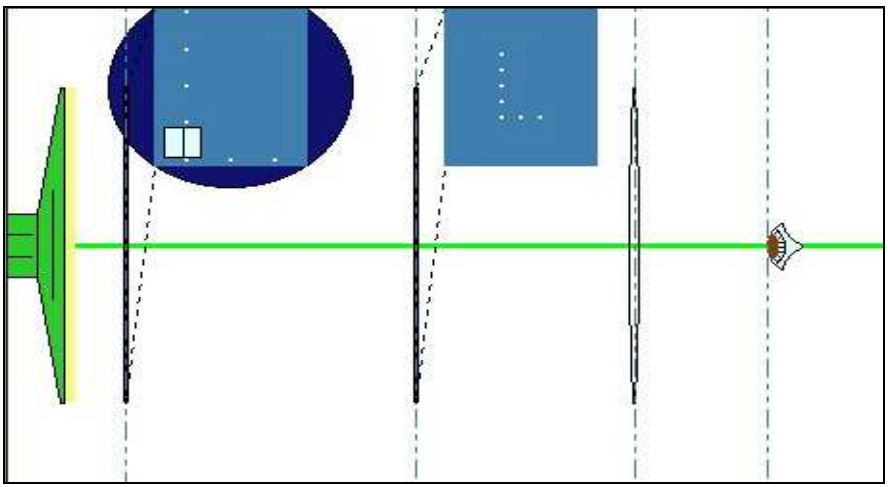
Figure 6



14. See Figure 6. An object is positioned within the focal distance of the lens. Is it possible to create an image on the screen?

- ❑ Yes, but the image will be virtual
- ❑ Yes
- ❑ No, because the image will be virtual

Figure 7



15. See Figure 7. The image is virtual. What does this mean for the distance of the object to the lens?

- ❑ Object distance is equal to the focal distance
- ❑ Object distance is smaller than the focal distance
- ❑ Object distance is larger than the focal distance

16. The relation between object distance, image distance, and focal distance can be expressed in a formula. What does this formula look like?

- ❑ $\frac{v}{f} + b = \frac{b}{f}$
- ❑ $\frac{v}{b} = \frac{1}{f}$
- ❑ $\frac{1}{v} + \frac{1}{b} = \frac{1}{f}$

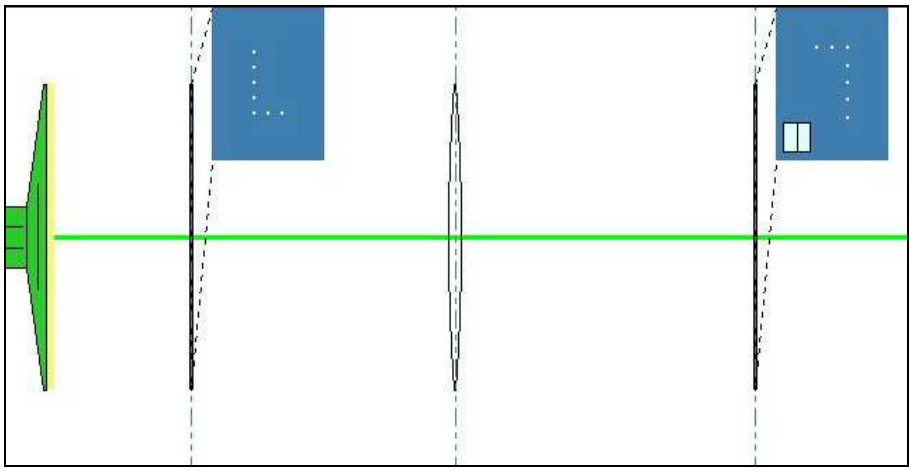
17. See Figure 8. The image distance is increased by moving the screen further away from the lens. What has to be done to the object distance to keep a sharp image?

- ❑ The object distance has to be decreased
- ❑ The object distance has to be increased
- ❑ The object distance has to stay the same

18. See Figure 8. The image distance is decreased by moving the screen closer to the lens. What has to be done to the object distance to keep a sharp image?

- ❑ The object distance has to be decreased
- ❑ The object distance has to be increased
- ❑ The object distance has to stay the same

Figure 8



19. See Figure 8. The object distance is increased by moving the plate with holes further away from the lens. What has to be done to the image distance to keep a sharp image?

- ❑ The image distance has to be decreased
- ❑ The image distance has to be increased
- ❑ The image distance has to stay the same

20. See Figure 8. The object distance is decreased by moving the plate with holes closer to the lens. What has to be done to the image distance to keep a sharp image?

- ❑ The image distance has to be decreased
- ❑ The image distance has to be increased
- ❑ The image distance has to stay the same

21. See Figure 8. The lens in the figure is replaced by a stronger lens. What has to be done to the image distance to keep a sharp image?

- ❑ The image distance has to be decreased
- ❑ The image distance has to be increased
- ❑ The image distance has to stay the same

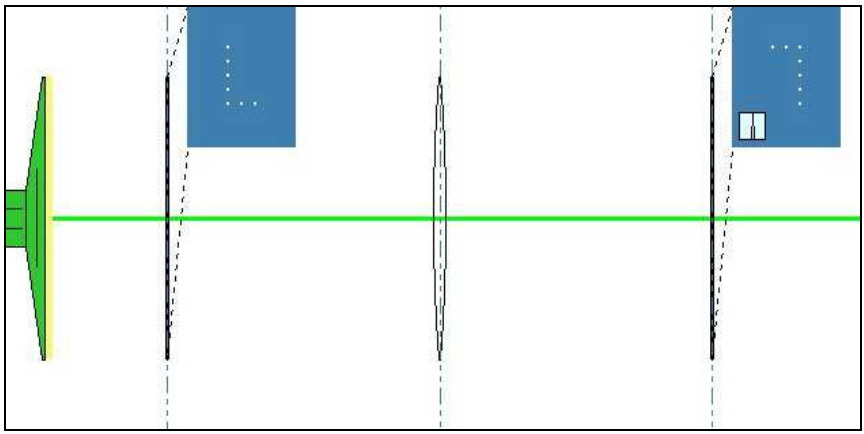
22. See Figure 8. The lens in the figure is replaced by a stronger lens. What has to be done to the object distance to keep a sharp image?

- ❑ The object distance has to be decreased
- ❑ The object distance has to be increased
- ❑ The object distance has to stay the same

23. See Figure 8. The lens in the figure is replaced by a weaker lens. What has to be done to the image distance to keep a sharp image?

- ❑ The image distance has to be decreased
- ❑ The image distance has to be increased
- ❑ The image distance has to stay the same

Figure 9



24. See Figure 8. The lens in the figure is replaced by a weaker lens. What has to be done to the object distance to keep a sharp image?

- ❑ The object distance has to be decreased
- ❑ The object distance has to be increased
- ❑ The object distance has to stay the same

25. See Figure 9. In this situation the magnification is 1. What can be done to get a smaller image than the object size?

- ❑ Move the object further away from the lens, and change the image distance until the image is sharp again
- ❑ Move the object closer to the lens, and change the image distance until the image is sharp again
- ❑ Just move the screen further away from the lens

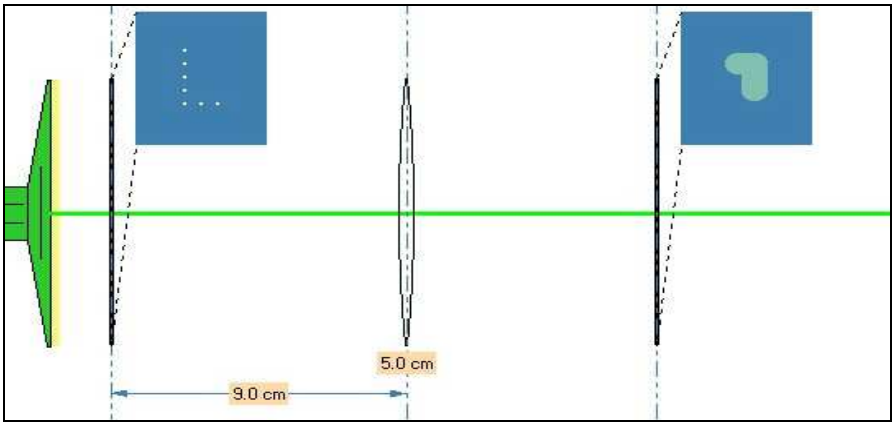
26. See Figure 9. In this situation the magnification is 1. What can be done to get a larger image than the object size?

- ❑ Move the object further away from the lens, and change the image distance until the image is sharp again
- ❑ Move the object closer to the lens, and change the image distance until the image is sharp again
- ❑ Just move the screen further away from the lens

27. See Figure 10. The image in the figure is not sharp. How does the screen on the right have to be moved to get a sharp image?

- ❑ The screen has to be moved further from the lens
- ❑ The screen has to be moved closer to the lens
- ❑ It is not possible to get a sharp image by moving the screen

Figure 10



28. An object is positioned 10 cm in front of a lens with a focal distance of 2 cm. What happens to the image?

- ❑ It is not possible to make an image
- ❑ The image is as large as the object, and not turned
- ❑ The image is as large as the object, turned and mirrored

29. Which of the following statements is true?

- ❑ A real image is magnified less when an object approaches a lens
- ❑ A real image is always shrunken in the same amount
- ❑ A real image is always magnified in the same amount

30.Which of the following statements is true?

- A virtual image is magnified less when an object approaches a lens
- A virtual image is always shrunken in the same amount
- A virtual image is always magnified in the same amount

Appendix C

Optics simulation: objects and tools

Here, an overview is given of the objects and tools that were used in the Optics 97, Optics 99, and Optics 2000 experiments. Students did not have all objects and tools available to them simulataneously. The configuration of the Optics simulation that was used in each study is described in the Method section of the empirical chapters.

Objects

Lens

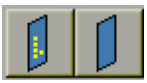
There are two types of lenses, positive and negative. In the environment these lenses have the same shape as their real-world counterparts, i.e. convex and concave. In addition, every lens has a specified focal point. In the experiment, the focal point of the lens could not be changed.

Lamp

In addition to the big lamp that is continually present in the first and third phase of the learning environment, in all phases a lamp with three divergent laser-like light beams is available. This lamp can be positioned anywhere in the working area.

Plate and screen

These two objects act in conjunction. The plate is a rectangle filled with holes that form a global L-shaped figure. The image of this figure can be projected on the screen, the source of light being the big lamp. Because the big lamp is positioned on the left side of the working area, the plate has to be put on the left side of the screen for the image to be projected.



Construction line

These vertical lines can be added to the working area to enable the student to measure any distance, not just the distance between objects. For example, in figure 1 a construction line is placed at the point where the outcoming light beams converge.

Distance measurement

The distance between any two objects, and between a construction line and an object, can be measured. The measured distance is displayed in the middle of the two objects. In Figure 2, two distances are displayed, including their units of measurements (centimeters). The distances that are shown are scaled down in order to make larger actual distances possible (a typical computer screen is much smaller than a real optical workbench).

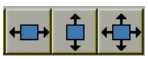
Eye

The eye is an advanced object that can be very useful to learn more about the virtual projection of images. When it's placed in the working area and to the right of a lens, it indicates the position on the principal axis where the L-plate *seems* to be when watched through the lens.

Tools

Object movement

The lamp can be moved in all directions, the other objects can only be moved horizontally. When an object is moved and there's a measurement line attached to it, the distance is automatically updated in real time.



Angular movement

After a lamp has been put in the working area, its light beams can be rotated. The rotation happens for all three beams at once, they cannot be rotated separately.

Angular measurement

The angle a light beam makes can be measured at any point in the working area. The angle shown is in reference to the principal axis.

Help line

A help line is a conceptual tool. When it's added to one of the lamp's light beams it extends that beam, showing the path that light beam would have followed if it had not been refracted by a lens (of course, this tool is only useful if the light beam actually goes through a lens).

Properties tool

This tool can be used to change various quantitative properties of every object, for example the distances of two objects can be changed quantitatively instead of by moving one of the objects. In our experiment, this tool had only a limited use, no advanced properties could be changed like the focal point distance of lenses. In addition to changing properties of an object, there was the possibility to give the object a name using this tool. Naming objects makes it easier to refer to an object in the hypothesis scratchpad. In addition to that, its essential for distances to have a name to be able to use a formula on them.

Formula

A formula that contains a computation on one or more variables (or just on two numbers) can be added to the working area. This is only possible if the relevant objects (or distances) have been labeled with a name (this is done with the properties tool). The outcome of the formula is displayed in the working area, and it is updated in real-time.

Notebook

At any time, subjects were able to note their ideas about the learning environment in a notebook (see figure 1). In this notebook a small replica of the current state of the environment was displayed, and subjects could write some lines of text to describe their ideas. The hypothesis scratchpad also enabled subjects to view old notes, but it was not possible to change these.

Appendix D

Bubbles simulation: tools

Here, an overview is given of the tools that were used in the Bubbles simulation. This computer simulation was used in the Optics 97 experiment.

Change amount in kettle

The amount of liquid that is used in the mix can be altered by moving a bolt (shown on that liquid's cask). The minimum is zero, maximum is 99. The unit of measurement is specified as deciliters.

Place Measurement line

A maximum of four measurement lines can be added to the graph. The position marks the time point at which the amount is measured.

Move measurement line

After a measurement line has been added to the chart, it can either be deleted or moved. Pressing this button (note the similarity between this button and the movement button in the Optics environment) allows movement of a measurement line.

Start experiment

After one or more casks is filled with liquid and at least one measure line is placed in the graph, it's possible to begin an experiment. The casks are emptied into the kettle, then the clock starts. A moving arrow under the graph indicates the time.

Stop/pause experiment

After an experiment is started, it ends when the moving arrow reaches the right point of the graph. Pressing this button also stops a running experiment. This works as a pause button. Nothing can be changed or reset when this button is pressed.

New experiment

When this button is pressed, a new experiment can be performed. The values with which the last experiment was run are restored. This means that should the start eperiment-button be pressed without changing anything, the last experiment will be repeated.

Appendix E

Bubbles simulation: underlying model

In the underlying model of the Bubbles computer simulation, a number of rules determines how much and in what way a liquid contributes to the overall reaction. This reaction is expressed as the amount of change in the amount of Magnum. The number of bubbles that originates every second is inversely related to this change. The rules can be expressed as a number of conditions that are evaluated during the running of an experiment.

The conditions are checked one at a time, and if one of them is satisfied the other conditions are skipped. The conditions are as follows:

1) if (t < (Kryton/10) → d(Magnum) = 0

The *t* variable denotes the time in seconds that has passed since the start of the current experiment. This condition implies that the amount of Magnum will not change if the amount of Kryton is not reached by the internal timer, or, the amount of Kryton determines how much the reaction will be delayed, because when the amount of Magnum doesn't change no bubbles originate. Only after this first condition is satisfied will the other conditions be considered.

2) if (Guernic < Sybar) → d(Magnum) = 0

This condition implies that the amount of Magnum will not change if the amount of Guernic is less than the amount of Sybar. Again, this means that no bubles will originate at all if this condition is satisfied.

3) if (Sybar >= (Guernic/2) → d(Magnum) = -Magnum * (Guernic - Sybar/3)/200

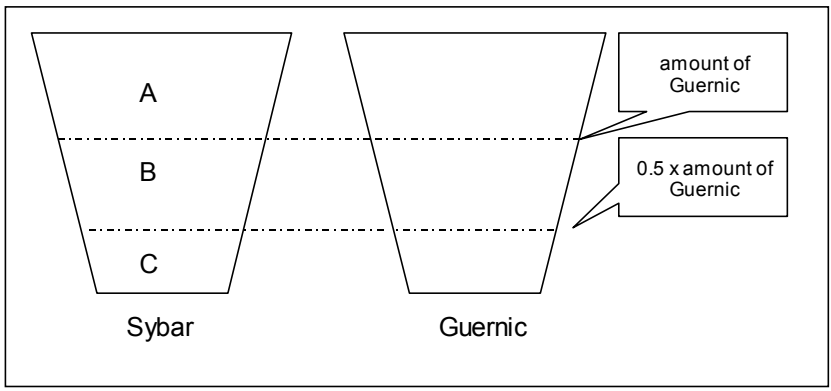
This condition indicates another relation between Sybar and Guernic. When there's a certain amount more Sybar than half the amount of Guernic, the reaction follows an exponential trajectory. The shape of this trajectory is partly dependent on the amount of Sybar, as well as the amount of Guernic.

4) d(Magnum) = -Magnum * Guernic/200).

This condition is only satisfied if the first three are not. When, for example, there is a only a large amount of Guernic, the reaction also follows an exponential trajectory. In this case the shape of the trajectory is only dependent on the amount of Guernic.

5) formula(true, bubbles = abs(d(Magnum)) * 10).

The second, third and fourth condition can be made more clear by showing the relation between Sybar and Guernic in a figure



Given a fixed amount of Guernic, we can specify three different reaction trajectories, dependent on the amount of Sybar. When the amount of Sybar is in region A, no bubbles originate from the reaction between Sybar and Guernic. When it is in region B, the amount of bubbles follows an exponential trajectory the properties of which are dependent on the amount of both Sybar and Guernic. Finally, when it is in region C, the reaction trajectory is also exponential, this time only dependent on the amount Guernic.