Signs for Logic Teaching

Tessa Eysink

2001

Ph.D. thesis
University of Twente

Twente University **Press**

Also available in print:
http://www.tup.utwente.nl/uk/catalogue/educational/logic-teaching/

Tessa Eysink

Signs for logic teaching

Doctoral committee

Chairman
Prof. dr. J.M. Pieters

Supervisors
Prof. dr. S. Dijkstra
Dr. ir. J. Kuper

Members

Prof. dr. W.J. van der Linden	(University of Twente)
Prof. dr. A.J.M. de Jong	(University of Twente)
Prof. dr. K. Stenning	(Edinburgh University, Scotland)
Prof. dr. J.J. Elshout	(University of Amsterdam)
Dr. J. van der Pal	(National Aerospace Laboratory, Amsterdam)

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: Dick van der Sar
Print: Grafisch Centrum Twente, Enschede

ISBN 9036516838

SIGNS FOR LOGIC TEACHING

THE EFFECT OF INSTRUCTIONAL VARIABLES ON THE DEVELOPMENT OF CONCEPTUAL KNOWLEDGE OF LOGIC

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 vrijdag 7 december 2001 te 16.45 uur

door

Tessa Henrike Susan Eysink

geboren op 28 juli 1974
te Oldenzaal

Dit proefschrift is goedgekeurd door de promotor:
prof. dr. S. Dijkstra

en assistent-promotor:
dr. ir. J. Kuper

Acknowledgments

Now that I have written my thesis and the end of my work as a PhD student is approaching, it is time to look back on the past four years and thank all the people who have helped me to reach this point.

I will start by thanking my supervisors. Sanne Dijkstra was always very confident in my capacities, sometimes even more confident than I was myself. His philosophy was one of giving a lot of freedom to stimulate creativity and giving support when necessary. Especially during the process of writing my articles, this support became very fruitful. During the important first year, I also received supervision from Hein Krammer and Jelke van der Pal, for which I was very grateful. My gratitude also goes to Jan Kuper who fulfilled this task in the following years. Although I have never seen Jan not being busy, he was always enthusiastic when he could learn Sanne and me something about logic. He made us realise that logicians approach the world in a completely different way than educational and cognitive scientists do.

During my PhD project I performed three experiments, which could not have been done without the help of a lot of people. I would like to thank: Jon Barwise and John Etchemendy for providing the source code of Tarski's World; Eugène Welling for putting time and subjects of his lectures at my disposal for my first experiment; Jakob for putting in a lot of effort helping me to convert my first experiment done on a Mac to my second and third experiment done on a pc, for all the other times he stood by to help me when I had questions about or trouble with my computer, and for his assistance during all the experiments; my colleagues for helping me piloting and supervising the experiments; my father and mother for doing the delicious catering during all experiments; and all the subjects for taking part in my experiments.

I am not the kind of person who likes to discuss her own research. But whenever I decided it was time again to discuss my work with my colleagues in the ProIST meetings, they were always very helpful, inspiring and motivating. Among them, I especially want to thank Mark, Casper and Koen who commented on almost every piece of work in my thesis. I would also like to thank the methodologists of the faculty, whom I asked for advice when I had trouble with the statistics. Moreover, I want to thank my father and Janet Heck for their linguistic comments, and Dick for designing the cover of my thesis. Furthermore, I would like to thank Bregje for the good time we had when we were roommates and the (sometimes scary, but always good) advices she gave, and Susanne, Renate and Bert for their friendship and good company during (and outside) lunch times.

Thanks, too, to Petra and all my friends for listening to all the ups and downs during my PhD project. Special thanks go to my parents who have helped me in every possible way and who have always been proud of me. Pap, mam, thank you for everything you have taught me and for having so much confidence in me. Finally, I would like to thank Fred. You have also helped me in all possible ways. You always have a refreshing look at things. I couldn't have done this without you. Thank you for being there!

Tessa

Enschede, October 2001.

Table of contents

1

The role of representations in the development of conceptual knowledge of logic

Abstract

The role of representations in learning logic is investigated. A theoretical review is given for representations, reasoning with representations and learning to reason with representations. The review focuses on the differences between everyday life reasoning and formal reasoning. Three main differences are found: (a) differences in abstraction; (b) differences in truth conditional aspects which involve differences in truth values between statements stated in everyday life and statements in logic, and (c) differences in user conditional aspects which concern pragmatic issues such as agreements for fluent conversations.

1.1 Context of research

Instruction is a communicative activity intended to initiate and support the development of knowledge and skills. It is part of a communication between an expert, for instance a teacher, a mentor or an instructional programme, and a novice, or learner. This instructional communication is about objects and conceptions about these objects and is conveyed through a medium (Dijkstra, 2001). As computer technology has advanced during the past decades, the use of this medium for instructional communication has increased resulting in a renewed interest in the design of instructions and instructional communications. The computer could provide all kinds of representations of objects and interactions with these objects. As a consequence, the effect of different types of instructional variables on the development of knowledge and skills has been studied and special interests for the effects of visualisation techniques, multiple representations and possibilities for the learners to interact with representations, have emerged. Researchers started to focus on theories that address issues such as authentic problem solving and visualising problem situations. Studies using these new ideas and theories often used materials from the empirical sciences such as biology, chemistry and physics. In the empirical sciences, representations are made on both a pictorial and a symbolic level, whereas the concepts and operations in the formal sciences, such as mathematics and logic, are often described in a purely symbolic way. As, for many learners, the instructions with only symbols as representations lead to a lack of understanding the formal rules and operations, the instructional communication for formal sciences are studied. More precisely, the main goal of the research described in this thesis is to study the effect of several instructional variables, such as different representations of reality and manipulation of these representations, on the acquisition of concepts and operations of logic, in particular *first-order predicate logic*. Logic is the science of (both human and machine) reasoning, which tries to discover conditions by which conclusions are justified and correct. Although logic is often presented by abstract, formal expressions, it can in fact be represented in several ways, ranging from abstract expressions in formal language on the one hand to concrete expressions in natural language on the other hand. The chosen form of a representation of formal objects will probably influence the mental images learners have of these objects and the (cognitive) operations they can perform on them, and this will thus influence the conceptions of the subject matter learners develop. The way of representing the objects of logic, and the visualisation and manipulation of these will be described and their effect on learning will be studied. This chapter will discuss several representations of formal objects and the advantages and disadvantages of these representations. Section 1.2 will focus on the topic of representations. First, some general ideas about representations will be elaborated followed by a description of language representations and visual representations. After that, section 1.3 will focus on the topic of reasoning with representations. This section will make an attempt to map out the similarities and differences in the meaning and practice of everyday life reasoning opposed to the meaning and practice of formal reasoning. The chapter continues with section 1.4 which discusses the topic of learning to reason with representations, followed by section 1.5 giving concluding remarks. The chapter ends with section 1.6 in which a short preview of the forthcoming chapters is given.

1.2 Representations[1]

The development of knowledge and skills takes place by experiencing reality. Experience results from the interaction with objects which on the one hand leads to the development of ideas about objects and on the other hand to the possibility to evaluate whether these ideas are correct or not. Experience can be divided into direct experience with reality or mediated experience. In the latter a *medium* is used to depict or describe reality. In education, mediated experience by using a representation of the reality is essential, as real objects are not always available or suitable to use. Dijkstra (2000) gave several reasons to use a representation instead of real objects: (a) objects are not always (easily) perceivable (e.g., bacteria, planets) or available (e.g., historic monuments in a foreign country); (b) experiments to determine the structure of reality are not always possible (e.g., there are no schools that have a particle accelerator at their disposal); (c) the duration of processes is sometimes too long or too short (e.g., evolutionary processes); and (d) using real objects can involve the risk of damage and personal safety (e.g., testing the safety of airbags by carrying out collision tests).

Mapping principles: the relation between reality and representation

The depiction or representation of objects is done by using *signs*. Every sign refers to an object or represents this object and every sign needs to be interpreted by an interpretant in order to be understood or to lead to a desired action (see Figure 1-1). Therefore, a sign can be described by a triadic relation (e.g., MacEachren, 1995; Markman, 1999; Seel & Winn, 1997): (a) a represented object, (b) a representing medium, and (c) an explanatory and contextual interpretation. The represented object, also called the referent or the signified, is the object which is represented by the representing medium, for instance the number of trees in your garden. The representing medium, also called sign-vehicle or signifier, is the 'carrier of meaning'.

[1]This literature review explicitly concerns external representations in contrast to internal representations, although most cognitive tasks are an interplay between external and internal representations (Zhang, 2000; Zhang & Norman 1994).

It is a representation in which (aspects of) the represented object are depicted or described, for instance the digit "3". The explanatory and contextual interpretation is the meaning to which the representing medium refers, for instance the meaning of the concept of cardinal numbers represented by digits. Other names used for this explanatory and contextual interpretation are meaning, sense, idea or content.

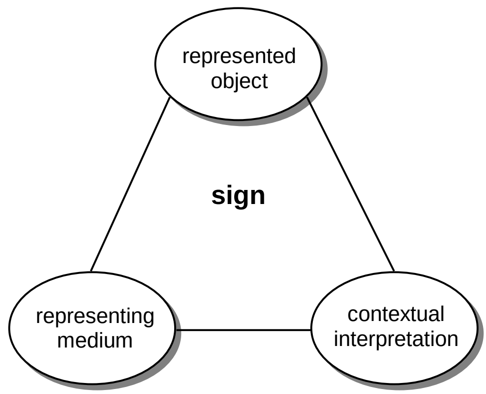


Figure 1-1. ***The triadic relation of a sign between represented object, representing medium and contextual interpretation.***

The representing medium (the representation) is related to the represented object (the reality) through a set of *mapping principles* that map elements of the reality to elements in the representation (Bernsen, 1994). Some representations are (almost) similar to the represented object, such as photographs or statues. In these cases, every element in the represented object is represented by a unique element in the representing medium, so that there is a one-to-one mapping or *isomorphism* between the two. If a representation is an abstraction of the represented object in which the characteristics relevant to the situation emerge and in which the characteristics irrelevant to the situation are left out, this is called *homomorphism* between the representing medium and the represented object. In this case, two or more elements in the represented object are represented by only one element in the representing medium. An example of this is the figure of a man or woman on a toilet door. In all cases in which a representation represents the represented object to some extent of similarity, the term *icon* is used. The relationship between an icon and the represented object depends on their 'mode of correspondence' (Blackwell & Engelhardt, in press) and is one of functioning (Eco, 1985). In different contexts, the same icon can have different functions, and can therefore have different meanings. For example, depending on the context, an icon of a wine glass may stand for 'a wine glass' (literal correspondence), for 'bar' (metonymic correspondence), or for 'fragile' (metaphoric correspondence).

There are also representations that have no similarity at all with their represented object. These are chosen arbitrarily by convention and are called *symbols*. Examples of these are the letters of the alphabet, words, punctuation marks and numerals. Because they are chosen arbitrarily, they are easily exchangeable. For instance, the number five can be represented as "five", "5" or "V". Other representations that have no similarities with their represented objects are *indices*. An index is physically connected with the represented object; it is an indication or cue that something exists or has occurred. For instance, a footprint means that someone has just walked by and smoke means that there is a fire.

Mapping principles: learning representations

Seel and Winn (1997, p. 298) argued that 'people's thinking consists of the use and manipulation of signs as media for the representation of ideas as well as objects'. This thinking, however, can only take place if the meaning of the signs used is clear to both sender and receiver. Especially, the meaning of symbols and abstract icons must be learnt, as additional knowledge of the mapping principles is needed for signs which do not have a (clear) intrinsic or intuitive relation with the represented object. For a person not familiar with chemistry, the formula H_2O has no meaning. However, after learning some conventions about chemistry, the learner will know that H_2O is the chemical formula for water and that this means that one molecule of water consists of two atoms H being hydrogen and one atom O being oxygen. The mapping principles play an important role here. The more natural the mapping principles are, the more they correspond to the user's knowledge, the better they are understood in advance, the smaller the chance of misinterpreting the signs or making mistakes (Bernsen, 1994). An example of this is the representation of temperature by using a line. A longer line is perceived as representing a higher temperature. However, when the mapping principle is used that a shorter line represents a higher temperature, this will be counter-intuitive and mistakes will easily be made. Norman (1988, p. 27) called this the principle of 'natural mapping'. He stated that taking advantage of physical analogies and cultural standards leads to immediate understanding.

As all information is conveyed through signs, and our thinking consists of using and manipulating these signs, the mapping principles used and thus the form of the representation will influence the perception of this information. Changing the kind of representation will have a direct effect on the ways people process, encode and recall this information (e.g., Seel & Winn, 1997). Examples of this are the numerous studies (for an overview of these studies, see Evans, Newstead & Byrne, 1993) on the Wason selection task (Wason, 1966), in which it became clear that subjects found it easier to solve a logical problem when the representations used in the problem were concrete (e.g., whether a person with a certain age was allowed to drink alcohol or not) than when they were abstract (e.g., whether a card with a certain letter on one side satisfied a certain rule). Other examples come from studies on reasoning under uncertainty in which subjects had less difficulty solving problems in which numbers of appearance were presented as frequencies than when the same problem was presented using representations of percentages (Cosmides & Tooby, 1996). These are just a few examples of studies showing that representation influences the perception of information. This suggests that representations also affect learning and, thus, choosing the appropriate representation or representations is assumed to be essential in reaching the learning goals and outcomes.

In the following sections, *visual (or pictorial) representations* and *language (or verbal) representations* will be discussed as two examples of types of representations often used for communicating scientific information.

1.2.1 Visual representations

Visual representations are often used to represent the reality. De Jong et al. (1998) discussed five dimensions along which visual representations can range: (a) perspective, that is from which point of view the representation is looked at; (b) precision, that is the level of accuracy or exactness of the information; (c) modality, that is the particular form of the representation that is used for displaying information; (d) specificity, that is the computational properties of the representation; and (e) complexity, that is the amount of information present. Dependent on the instructional purposes, visual representations can have several functions. For instance, Dijkstra, Jonassen and Sembill (2001) used the following categorisation: (a) the representation can depict the reality (e.g., a photograph, a drawing, a map); (b) the representation can depict a change of state of an object in the reality or an aspect of the reality (e.g., a graph showing the course of temperature in a city over a year); or (c) the representation can show a process occurring in reality (e.g., a demonstration, a simulation). Besides this categorisation, there are several other ways to group the functions that visual representations can have, dependent on the situation and the instructional purposes (e.g., Mayer & Gallini, 1990).

There are different types of visual representations with different functions and different names. In this thesis, the term 'visual representation' or 'visualisation' is used as a general concept. A specific subset of the visual representation is the 'graphical representation', which is used in the experiments described in Chapter 3 up to Chapter 5.

Benefits of visual representations

Several studies have tried to explain what gives visual representations so much power. For instance, Paivio (1986) and Clark and Paivio (1991) explained with their theory of dual coding that memory for pictures is better than memory for texts. According to this theory, there are two independent coding systems, the verbal coding system and the image coding system. Information coded in both systems is assumed to be remembered better than information coded in only one system. As texts are processed and encoded in the verbal system and pictures are processed both in the image and in the verbal system, pictures are more likely to be remembered better than text. Furthermore, theories of cognition state that advantages of graphical representations can be explained by the fact that these representations are more effective in communicating material, if they are understood. For instance, Larkin and Simon (1987) described the idea of 'locality': locating semantically related information together in a diagram can support extremely useful and efficient computational processes. Kulpa (1995) and Koedinger and Anderson (1994) described the idea of 'emergence': by coding information in diagrams, new information becomes visible. Stenning and Oberlander (1995) described the idea of 'inexpressiveness' or 'specificity': diagrams enforce representation of information, leading to weak expressiveness and limited abstraction and thereby facilitating processibility. An example that can clarify the idea of specificity, is the sentence "Colin is sitting next to Rene". In order to make a picture of this situation, the exact positions of Colin and Rene need to be expressed. When this information is not available, two or more pictures can be made of this single sentence. In other words, sentences can express abstraction, whereas pictures are specific.

Although the description of visual representations focuses on the characteristics that make these representations so powerful, it is important to realise that visual representations can also interfere with cognitive processes when not constructed in a proper way or when used in the wrong context. Cox (1999) stated that the effectiveness of representations depends upon a three-way interaction of (a) properties of the representation, (b) demands of the task, and (c) characteristics of the user, such as prior knowledge and cognitive style.

Visual representations in logic

In first-order predicate logic, the function of visual representations is the depiction of the reality. This reality can be an everyday life situation, a world of mathematical entities or any other situation. The reality consists, by consensus, of the objects the communication is about. The resources of this communication are, by consensus, the representations. By using visual representations, logical expressions can be related to a reality. Reasoning can be checked visually and it is supposed that steps made can be retained easily. Visual representations can vary in abstraction level. Many attempts to visualise logic concern abstract objects. Examples are Venn Diagrams, Euler Circles and Peirce's Existential Graphs. Figure 1-2 shows the visualisation of the expression $p \rightarrow q$ in these three representational systems (for a short description of logical operators, see Table 1-1). Venn Diagrams and Euler Circles both involve sets. Objects with the same characteristic constitute a set, which is depicted by a circle. As objects can have more than one characteristic, they can be members of more than one set, depicted by the intersection of two circles representing the characteristics. In Venn Diagrams, circles represent sets in which objects with a certain characteristic (represented by a propositional letter attached to the circle) may exist. Shaded circles represent (parts of) empty sets and an asterisk represents existence (i.e., there is at least one object present in the set of a certain characteristic). In Euler Circles, the universe of all objects is represented by a rectangle, whereas circles represent sets of objects within the universe with a certain characteristic (represented by a propositional letter included in the set). Peirce's Existential Graphs seem to visualise logic in the same way as Venn Diagrams and Euler Circles, but this is not the case. In Venn Diagrams and Euler Circles the propositional letters represent characteristics and circles represent sets of objects. In contrast, propositional letters (e.g., p) in Existential Graphs, represent true statements (i.e., p) and propositional letters enclosed by a circle negate these propositions (i.e., $\neg p$).

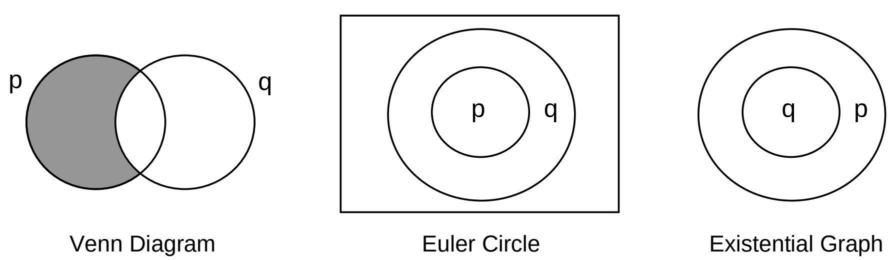


Figure 1-2. ***Three representations of $p \rightarrow q$.***

In Figure 1-2, all three pictures are representations of the proposition $p \rightarrow q$. In the Venn Diagram, all objects with the characteristic p lie within the p-circle, but since the shaded part of the p-circle is empty, all objects with the characteristic p are inside the q-circle as well. Hence, every object with the characteristic p has also the characteristic q, that is $p \rightarrow q$. In the Euler Circle, it is immediately clear that every object with the characteristic p has also the characteristic q. Hence, this too is a representation of the expression $p \rightarrow q$. The Existential Graph literally is a visual representation of the verbal description "it is not the case that p and not-q" (i.e., $\neg(p \wedge \neg q)$), since the inner circle represents the negation of q (i.e., $\neg q$) and the outer circle represents the negation of "p and $\neg q$" (i.e., $\neg(p \wedge \neg q)$). This then is equivalent to the expression "if p then q" (i.e., $p \rightarrow q$).

Although all three types of visual representations look quite similar, their semantics are clearly not. The mapping principles that are used differ, leading to differences in expressivity, that is the ability of the representational system to express information, and specificity, that is the extent to which the representational system requires specification of information (e.g., Dobson, 1998, 1999, Stenning, 1999).

1.2.2 Language representations

Language is a medium intended to convey information. When describing reality or when reasoning about reality, the use of language is inevitable. Lakoff and Johnson (1980, p. 10) spoke of the 'conduit metaphor'. In this metaphor, words and symbols are seen as carriers of meaning. It states that 'the speaker puts ideas (objects) into words (containers) and sends them (along a conduit) to a hearer who takes the idea/objects out of the word/containers'. There are several kinds of containers or languages that can be chosen to convey information. Each language has its own advantages and disadvantages, so that the suitability of the different languages depends on the situation the language is used in.

The structure of language

Three important aspects of language in general are *syntax*, *semantics* and *pragmatics*. Syntax analyses sentence structures and studies grammatical correctness. Semantics studies the expressions of a language in relation to what they mean. Aspects of meaning covered by semantics are called *truth conditional aspects of meaning* (Gamut, 1982), which means that changes of meaning can result in changes in truth. Pragmatics studies the use of expressions in different user contexts. All aspects which cannot be covered in terms of truth conditions fall under the *user conditional aspects of meaning* (Gamut).
Within the area of pragmatics, Grice (1975) developed a theory in which the concept of *conversational implicatures* takes a central role. A conversational implicature of a sentence is an implication of that sentence, but not in a strictly logical way. It is the suggestion which becomes clear from that sentence. People use these implicatures in a systematic manner with the *conversational cooperative principle* as an important one. The assumption underlying this principle is that all people in a conversation behave cooperatively. Everybody behaves in such a way that the common goal of communication is reached in the most optimal way. Within this general principle of cooperation, some specific principles, the so-called *conversational maxims*, can be distinguished.

Table 1-1. A short description of logical concepts

The (formalised) language of first-order logic contains *names* (*a*, *b*, …, *x*, *y*, …) to denote individual objects, and *predicates* to express properties of objects and relations between these objects. For example, *Large* is a predicate and *Large*(*a*) says that object *a* is large. Likewise, *Larger*(*a*, *b*) says that object *a* is larger than object *b*, and *Between*(*x*, *y*, *z*) expresses that object *x* is positioned between the objects *y* and *z*. Expressions of this form are called (*elementary*) *propositions*. They can be combined into more complex propositions by the *connectives*: *negation* (¬, not), *conjunction* (∧, and), *disjunction* (∨, or), and *conditional* (→, if… then..., sometimes called *implication*). For example, *Larger*(*a*, *b*) → *Smaller*(*b*, *a*) says that "*if a* is larger than *b*, *then b* is smaller than *a*".

Propositions express states of affairs about the world and they can either be *true* or *false*. The connectives are *truth functional*, that is, the truth or falsehood of a complex proposition is completely determined by the truth or falsehood of the propositions of which it is composed, as described by the *truth tables* given in the table below. Thus, for example, if p is true, then $\neg p$ is false. And also, if p is true and q is false, then $p \wedge q$ is false, but $p \vee q$ is true. Moreover, $p \rightarrow q$ is only false if p is true and q false. In all other cases $p \rightarrow q$ is true (see table below).

The formal language contains names to denote objects. There are two kinds of names: constants (a, b, …) denote specific objects (it is assumed that it is known which objects are denoted), whereas variables (x, y, …) denote arbitrary objects (it is not known which objects are denoted). In combination with variables two *quantifiers* can be used, the universal quantifier (∀, for all) and the *existential* quantifier (∃, there is at least one). For example, ∀*x Large*(*x*) says that *all* objects *x* are large, whereas ∃*x Large*(*x*) expresses that there is *at least one* object *x* which is large. A more complicated example is ∀*x* (*Cube*(*x*) → *Large*(*x*)). Literally, this formula says that "every object *x* is large, if it is a cube". In everyday language this is normally said as "every cube is large", or "all cubes are large". In this formula all occurrences of the variable *x* are said to be within the *scope* of the quantifier, and all occurrences of *x* are *bound* by the quantifier. If a variable is not bound by a quantifier, it is called *free*. If a formula contains a free variable, it cannot be known whether this formula is true or false, since it is not known which object is denoted by this variable. For example, it is unknown whether *Large*(*x*) is true or false, since it is not known which object is denoted by *x*. On the other hand, when *x* is bound by a quantifier, as in ∀*x Large*(*x*), we can know whether this formula is true or false, as *x* ranges over all objects now, and for each individual object it can be checked whether it is large or not. In general, it can be determined whether a formula is true or false whenever that formula does not contain free variables. Such formulas are called *sentences*.

p	q	$\neg p$	$p \wedge q$	$p \vee q$	$p \rightarrow q$
0	0	1	0	0	1
0	1	1	0	1	1
1	0	0	0	1	0
1	1	0	1	1	1

Note. p and q denote arbitrary propositions. 0 = false; 1 = true.

Examples of such conversational maxims are maxims of quantity (i.e., conversational contributions must be presented as informatively and as clearly as possible), maxims of relation (i.e., contributions must be relevant), maxims of quality (i.e., contributions must be -or thought to be- true), and maxims of manner (i.e., obscurity and ambiguity must be avoided). Hence, a conversational implicature is not a (logical) implication of a sentence, but it follows logically from the assumption that the speaker makes his contribution in accordance with the conversational maxims.

Language representations in logic

In logic, two types of language representations are used: natural languages and specifically designed formal languages. There are several differences between these two types of language representations. A formal language consists of precisely pre-defined constants, variables and operations, whereas a natural language develops during everyday life by producing new words or by forgetting existing words. In natural language it is acceptable (from a point of view of understandability) that sentences are syntactically incorrect, whereas in formal language a syntactically incorrect sentence is not a sentence and thus does not have any meaning. Moreover, meaning in natural language is often fuzzy and may give rise to ambiguities and misunderstandings, whereas the meaning of expressions in a formal language is determined uniquely. Finally, a formal language contains specially designed symbols so that formal expressions can be concise. Natural language, on the other hand, usually needs many words to express the same concept. In general, natural and formal languages differ in their level of abstraction.

Formal language plays a large role in modern logic. In Table 1-1, a short description is given of the concepts used in first-order predicate logic. Typically logical questions, such as validity, provability and rules of correct reasoning, can be analysed precisely by using a formal language representation. Still, natural language representations play a role in logic as well. They are used in everyday life reasoning and the correctness of everyday life reasoning is evidently important as well.

1.2.3 Summary

The topic of representations of reality in general was discussed followed by a discussion on visual representations and language representations as two representations suitable for instructional purposes in logic. Visual representations in logic are mostly used to *depict* the reality, whereas language representations are generally used to make *statements about* the reality and to reason about this reality. Both types of representations can vary in abstraction level. Furthermore, a distinction can be made between truth conditional aspects and user conditional aspects of representations.

1.3 Reasoning with representations

Human thought involves reasoning. This reasoning can either occur implicitly (e.g., knowing immediately that the woman in white is the bride) or explicitly (e.g., finding the cause why your computer does not work anymore). Experiences from the past can be generalised and this knowledge can be used in new situations in order to draw conclusions. Traditionally, reasoning is divided into *deduction* and *induction*[2]. A deductive inference is the reasoning process in which specific conclusions are drawn from general rules. The conclusion is implicitly present in the general rule and thus follows logically from that rule. An inductive inference is the reasoning process in which general conclusions can be drawn from specific instances. In this case, information is added and thus the conclusion does not necessarily follow logically. In everyday life, reasoning does not always follow strictly the rules of logic. Examples are pragmatic and statistical inferences and heuristic reasoning. Pragmatic inferences are inferences which, given the context, are plausible rather than that they follow logically from the premises (see Grice, 1975). Statistical inferences are inferences which can be drawn with a certain degree of probability, thus inferences which are very likely (e.g., Kahneman, Slovic & Tversky, 1982). The latter are highly influenced by the use of heuristics during the reasoning process. Heuristics are simple rules of thumb to reason with and to make decisions. Gigerenzer (1991, p. 254) describes them as 'strategies of discovery less general than a supposed unique logic of discovery but more general than lucky guesses'. Examples of heuristics are the availability heuristic (e.g., Tversky & Kahneman, 1973) and the representativeness heuristic (e.g., Kahneman & Tversky, 1973). Heuristics generally lead to correct answers, but in some cases they cause fallacies in the reasoning process.

Thus, heuristics and pragmatics interfere with the reasoning process and together with the abstractions which logic uses, they determine the differences between reasoning in everyday life and reasoning in formal logic. In order to develop a clear picture of these differences, an analysis of everyday life reasoning compared to formal reasoning will be performed. This is considered to be important, as only when these differences are mapped out, instruction can be adjusted in order to promote the acquisition of the meaning of the concepts of logic. To make an analysis of the differences between everyday life reasoning and formal reasoning, language representations are used as they are considered to be the most suitable representations for this purpose. The analysis starts with differences between natural language and logic on syntax, semantics and pragmatics (section 1.3.1). This is followed by a section on the differences in making knowledge explicit (section

[2] Aristotle was the first to make this distinction. He used the term 'apodeixis' for deduction and the term 'epagogè' for induction.

1.3.2) and a section on the differences in abstraction (section 1.3.3). Then the basic elements of language representations will be analysed (section 1.3.4) followed by a section on the formal definition of first-order predicate logic (section 1.3.5). Finally, a conclusion will be given summarising the differences found (section 1.3.6).

1.3.1 Differences in syntax, semantics and pragmatics

Sentences in natural language are constructed by applying rules of grammar which make up the syntax of natural languages. The semantics of natural language make it possible that syntactically incorrect or half-finished sentences are, most of the time, also understood. The latter is caused by the fact that human beings communicate about meaningful things and that they can interpret statements. Statements are considered to be true or false in the world or there may be some uncertainty whether they are true or false. In all these cases, the statements are given within a certain context, and the most probable meaning can be induced from this context. Moreover, pragmatics including aforementioned conversational maxims plays an important role in natural language.

Logic uses a syntax that is rigid and precise. For instance, leaving out one or more parentheses can make a logical expression syntactically incorrect or ambiguous. Putting one or more parentheses in a different place in the expression can completely change the meaning of the sentence. Moreover, expressions in (first-order predicate) logic are either true or false, there is no uncertainty. Furthermore, logic makes a distinction between *validity* of a conclusion and *meaning* of this conclusion. A conclusion can be logically valid to draw (i.e., can follow logically from other sentences), whereas it does not have to be true. It can even be a nonsense sentence, such as "If an elephant is a mammal, then Amsterdam is the capital of the Netherlands". This distinction originates from the condition that the rules of logic must be valid in all situations. As a consequence, these rules must be formulated in a general way, that is, in such a way that they are not restricted to a given context. In this way, every meaning can be given to these general elements, so that meaning and content do not play a role anymore. Furthermore, logic also abstracts from pragmatics.

1.3.2 Differences in making knowledge explicit

When humans communicate, they share knowledge that can be used in their conversations. This knowledge can be shared explicitly by stating it during the conversation, for instance to achieve common grounds, or it can be shared implicitly assuming that the knowledge is so obvious that there is no need to make it explicit. In logic, all knowledge that is used must be made explicit. This difference clearly emerges from the following example (taken from a study of Kahneman & Tversky, 1973, and revised for this purpose):

> A group of 100 people contains 30 psychology students and 70 computer science students. One person is chosen at random from the group. It is a 23-year old girl that likes to go out, has a busy social life and works in a nursery in her spare time. Do you think the girl is a psychology student or a computer science student?

Reasoning purely logically, the answer to the question is that the girl probably (with a probability of .7) is a computer science student. However, the majority of the respondents will probably say that it is most likely that the girl is a psychology student, as implicit knowledge about relations between characteristics of persons and the studies they are doing is used in the reasoning process.

1.3.3 Differences in abstraction

In the introduction, the formal, general and abstract format of logic was already mentioned. In this section, the topic of *abstraction* is elaborated. Abstraction means leaving out irrelevant aspects and focussing on relevant aspects. For instance, a sign indicating a (coin-) telephone booth shows a pictorial telephone with some coins as relevant characteristics whereas irrelevant characteristics like colour or size of the telephone are left out. Which features are relevant depends on the situation and the point of view. For instance, a mother wants to buy new shoes for her daughter. For the mother, the fit of the shoe is the most relevant characteristic when she chooses a shoe, whereas, for the girl, the colour of the shoe will be more important. The mother abstracts from the outward appearance of the shoe, whereas the child abstracts from the fit of the shoe. Besides being less than the original object by leaving out irrelevant aspects, an abstraction is also more than the original. By excluding specific characteristics of a single object, the abstraction refers to more objects than just that single one. In this way, the concept "chair" does not refer to one single chair. It refers to all chairs in general, that is all individual chairs that belong to the category of chairs. Furthermore, Kuper (in press) stated that abstraction 'creates new entities'. In natural language used in everyday life, the meaning of a word and the word itself are intrinsically related. When communicating, meaning is directly transferred through the use of words. There is an immediate connection between words and their meaning by the reference of the words to their meaning. For instance, the sentence "do not think about an ice-cream" cannot be read without thinking about an ice-cream. However, abstraction means breaking this intrinsic connection of form and content in two. Logic approaches objects in a language as meaningless instances to which meaning can explicitly be assigned. The object itself remains meaningless, but it refers to characteristics that give meaning to the instance (see Figure 1-3). Consequently, logic can discuss the word "saxophone" (e.g., a word of nine characters) and the object "saxophone" (e.g., a brass instrument). Only by explicitly mapping the word "saxophone" on the object "saxophone", meaning is assigned to the word. This mapping is arbitrary in the sense that the object could as well be assigned to the word "trumpet" instead of "saxophone".

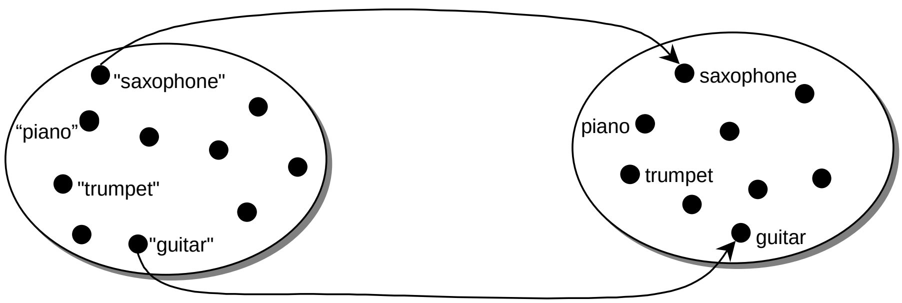


Figure 1-3. ***The mapping of objects within a set on objects within another set. The left set consists of words indicating musical instruments; the right set consists of objects indicating musical instruments.***

The separation of form and content becomes clear in the way logic makes (combined) statements about the world. The form is the appearance of the statement and the content is that what the statement is about. In determining the truth or falsity of a statement in the world, the relation between the two is important. However, in determining the validity of reasoning, only form is of importance, as logic then abstracts from content.
In the analysis of the basic elements of first-order predicate logic, these abstractions will be encountered again.

1.3.4 Differences emerging from the analysis of basic elements

In this section, the basic elements of first-order predicate logic will be discussed. This is done because these elements are often used in a different way in natural language than in formal language (Begg & Harris, 1982; Sanders & Antes, 1988), which can cause misconceptions when learning logic. The basic elements are the variables, the constants, the connectives - conjunction (∧), disjunction (∨), conditional (→), and negation (¬) - and the two quantifiers - universal quantifier (∀) and existential quantifier (∃). They will all be addressed individually concerning their truth conditional aspects (semantics) and their user conditional aspects (pragmatics).

Constants

Constants are symbols that are used to refer to *specific* individual objects, just as names are in natural language. The difference between names in natural language and constants in first-order predicate logic is that the latter are required to refer to exactly one object. In natural language, the name "Colin" can refer to all people named "Colin". In logic, there can only be one instance called "Colin". Furthermore, in natural language, there are names that do not refer to actually existing objects, for instance "Neptune". In first-order predicate logic, a name must always name an existing object. In both languages though, one object can have more than one name or objects can have no name at all.
In addition, in first-order predicate logic deictic words as "that" or "those" in, for instance, "those shoes" or "it" in the sentence "I am looking for my wallet, but it can be anywhere" cannot be used.

Variables

In logic, variables are introduced to denote *arbitrary* objects. These objects can have certain characteristics which are called *properties* when they are involved with only one object (e.g., pencil, red) and *relations* when they are involved with more than one object (e.g., married_to, older_than). The logical name for these characteristics is *predicates*[3]. In natural language, the variable and the object (with its properties) to which it is denoted are intrinsically related. Imagine, for instance, a red pencil. In natural language, this situation can be described by the sentence "there is a red pencil". The pencil is considered to be the object and this pencil has the property of being red. In first-order predicate logic, however, this situation is described by the sentence $\exists x\ (Pencil(x) \wedge Red(x))$, meaning "there is an object x with the properties of being a pencil and being red". In this case, the pencil is not the object itself, but the variable x is the object that has the properties of pencil and red. In general, in natural language substantive nouns are seen as objects, whereas adjectives are considered to be properties telling something about the object. In logic, both substantive nouns and adjectives are seen as properties of an arbitrary object x.

Conjunction

The logical symbol for the conjunction is $\wedge$, its natural language equivalent being "and". A conjunction $p \wedge q$ is true if both p and q are true. If one or both are false, the conjunctive expression is false as well. Other equivalents of the conjunction are, for example, "but", "although", "in spite of", "notwithstanding" and "as well as". Commas or periods can be equivalents of the conjunction, too. Equivalents of the conjunction

[3] This explains the name "predicate logic".

used in natural language can express subtle differences in meaning. For instance, in sentence (1) and sentence (2) the difference between the logical equivalents "and" and "but" becomes salient. The connective "but" is not just a conjunction; it also expresses a contradiction. First-order predicate logic abstracts from these subtle differences and only uses the five basic connectives to express ideas.

It is my birthday today and I am wearing a skirt (1)
It is my birthday today but I am wearing a skirt (2)

Logic also abstracts from internal relations, that is, statements can be made about two propositions p and q that do not have a relationship at all. For instance, sentence (3) lacks an internal contextual relation. In natural language, most people will experience this sentence as a complete nonsense sentence.

I brush my teeth and the ball is round (3)

As a consequence, logic does not recognise special relationships that express more than the two (or more) parts of the sentence. An example of this is given in sentence (4). In logic, this sentence can be expressed as "Barwise wrote a book" and "Etchemendy wrote a book". In natural language, however, it will be clear that they wrote a book *together*.

Barwise and Etchemendy wrote a book (4)

Another characteristic of the conjunction in logic is that $p \wedge q$ and $q \wedge p$ is logically equivalent. The expressions p and q can be exchanged without a change in meaning. In natural language, however, two expressions in a conjunction mostly have some kind of internal relation. Sentence (5) and sentence (6) differing in the order of the expressions within the sentences can illustrate this difference. In logic, these two sentences mean the same, as $p \wedge q$ and $q \wedge p$ are logically equivalent. In natural language, however, the two expressions in both sentences have an internal temporal relationship in which order matters. This is shown by the fact that Sally's clothes are still dry in sentence (5), whereas they are wet in sentence (6).

Sally took off her clothes and jumped into the water (5)
Sally jumped into the water and took off her clothes (6)

In natural language, people expect other people to speak in an orderly way. Therefore, when not mentioned explicitly, it is assumed that according to the speaker the two (or more) parts of a conjunction took place in the order mentioned leading to an internal causal relation.

A special instance of the conjunction concerns the class of conjunctive adverbials. Conjunctive adverbials are words that add information to sentences about the goal of a proof or argument and about relations between the information provided by the sentences within that proof or argument. They are often used to make the single steps within a proof more salient. They can express, for example, inferential relationships (e.g., therefore), temporal orderings (e.g., afterwards), contrasts (e.g., on the other hand), and paradoxes (e.g., yet) (Eulenberg, 1996). Logic abstracts from these extra meanings.

Disjunction

The logical symbol for the disjunction is $\vee$, its natural language equivalent being "or". A disjunction $p \vee q$ is false if both p and q are false and true if either p or q is true, but not both. When both p and q are true, the truth value depends on whether the disjunction used is inclusive (then $p \vee q$ is true) or exclusive (then $p \vee q$ is false). Other equivalents of the disjunction are, for example, "either…or…" or "unless". A comma or period, too, can be an equivalent of the disjunction (which makes the comma and period ambiguous, as they can also be used as conjunctions). Like the conjunction, equivalents of the disjunction used in natural language can also express subtle differences in meaning.

In natural language, both inclusive and exclusive disjunction can be used, and the meaning can again often be extracted from the context. For instance, everybody will understand that the disjunction in sentence (7) is meant to be inclusive (a 70-year old woman accompanying four children will also receive 25% discount), whereas the disjunction in sentence (8) is meant to be exclusive (when mother tells this to her daughter, the daughter will know that she is not allowed to buy a new sweater as well as a new pair of trousers).

To get 25% discount, one has to be older than 65
or one has to accompany more than two children (7)

You are allowed to buy a new sweater or a new
pair of trousers (8)

The examples (7) and (8) show sentences in which the two parts of the sentence do not exclude each other and meaning has to be derived from the context. There are also sentences in which the two parts of the disjunction exclude each other, as is the case in sentence (9).

Rick is in London or Rick is in Paris (9)

It will be clear that Rick cannot be in London and Paris at the same time. This leads to a pragmatic issue. Sentence (9) is true if Rick is in London or in Paris, but if someone knows that Rick is in London, then it will not be very cooperative to tell that Rick is in London or in Paris. Then the single utterance "Rick is in London" would be more informative. Thus, in natural language, a disjunction is only expected to be used when the speaker is not completely convinced of the truth of either part of the sentence.

Furthermore, the disjunction in natural language always assumes an internal relation, just as is the case with the conjunction. Logic again abstracts from this relation.

Conditional

The logical symbol for the conditional is $\rightarrow$, its natural language equivalent being "if…then…". A conditional $p \rightarrow q$ is false if p is true and q is false. In all other cases, the conditional $p \rightarrow q$ is true. Other equivalents of the conditional are, for example, "when", "provided that" and the *hidden implication* of the form "All P are Q", meaning "For all objects it counts that if it is P, then it is Q".

In everyday language, the meaning of the conditional differs in some aspects from the meaning in logic (a difference in the truth condition). If the first part of the sentence is true, the meaning of the complete sentence is the same in everyday life language as in logic. But if the first part of the sentence is false, then people in everyday life will state that you cannot know whether the complete sentence is true or false, such as, for instance, in sentence (10).

If my grandfather celebrates his hundredth birthday,
then he will give a big birthday party (10)

When my grandfather dies when he is 92, you will probably say that you will never know whether it was true what was said in expression (10). In logic, a conditional is always true when the first part of the sentence is false.

A second difference between the conditional in everyday life and in logic concerns the relations between the expressions p and q. In everyday life, these expressions have some kind of causal relationship. The sentence "if p, then q" expresses that p implies q which says that the truth of q is caused in some way by the truth of p. In logic, the conditional is defined in terms of truth and falsehood without the need for a relationship of causality or content between the expressions p and q. Therefore, it is possible to reason about nonsense sentences, like sentence (11).

If two plus two is four, then I like spinach (11)

Another difference is that in everyday life people do not normally reason with conditionals of which they know the first part is false. For instance, the antecedent of sentence (12) is false, as everybody knows that Monday does not come after Tuesday. In this case, this sentence does not tell anything in natural language, whereas in logic sentence (12) is true.

If Monday comes after Tuesday,
then my birthday will be on a Saturday (12)

As was the case with the disjunction, the conditional has also a pragmatic use in everyday life. Take as an example sentence (13).

If the kahikatea is a tree, then it has leaves (13)

Users knowing that the kahikatea is a tree would be more informative and clear, if they would say "Because the kahikatea is a tree, it has leaves" or "The kahikatea has leaves" instead of the conditional expressed in sentence (13). Users who are convinced that the kahikatea is not a tree would be more informative and clear, if they would express themselves with "The kahikatea is not a tree". Thus, in everyday life and according to the maxim of quantity of Grice (1975), the conditional is assumed only to be used when no stronger statement can be made, thus if the user does not know whether the antecedent or consequent is true or not.
In the experiments described in Chapter 3, Chapter 4 and Chapter 5, the goal is the acquisition of this connective, as the conditional is a difficult and counter-intuitive connective.

Negation

The negation is a special instance of the set of connectives, as it does not connect two statements p and q. The logical symbol for the negation is $\neg$, its natural language equivalent being "not". A negation $\neg p$ is true if p is false and is false if p is true. Other equivalents of the negation are, for example, "it is not true that…", "no…" or "neither…nor[4]". In addition to these types of negations, there are other ways to express a negation. Examples of these are prefixes such as un- (e.g., undesirable), im- (e.g., immoral), ir- (e.g., irrational), dis- (dismantle) and anti- (antisocial), suffixes such as -less (e.g., heartless), and words that mean the opposite of other words

[4] This negation in fact consists of a combination of two negations and one conjunction ($\neg p \wedge \neg q$).

(e.g., open - closed)[5]. Logic abstracts from these differences between different kinds of negations.

In natural language the meaning of a sentence depends on the position of the negation. For example, the sentences "not I am bad" and "I am not bad" express different ideas. Logic deals with this by a strict use of parentheses.

Differences between the negation used in natural language and the negation used in logic arise when a sentence contains two (or more) negations. In logic, $\neg\neg p$ means the same as p, whereas in natural language $\neg\neg p$ often expresses subtle differences in meaning (see sentence (14)) compared to its logical equivalent p.

Sophie thinks it is not impossible to be at home at six (14)

Although Sophie thinks it is possible to be at home at six, the sentence also implicitly states that it will not be an easy job for her to do.

Universal quantifier

The logical symbol for the universal quantifier is $\forall$, its natural language equivalent being "for all". Other equivalents of the universal quantifier are "each" and "every". Differences between the universal quantifier in natural language and in logic concern 'scope ambiguity' (e.g., Blackburn & Bos, 1997; Kamp & Reyle, 1993) in which the order of the quantifiers plays a critical role. A well-known example in this respect is given in sentence (15). This sentence can mean that all men love one and the same woman or it can mean that all men each love a woman.

All men love a woman (15)

Comparing sentence (15) with sentence (16) and sentence (17) shows that all three sentences have the same structure. However, context tells us that sentence (16) means that all Dutch people have one and the same queen, whereas sentence (17) means that each Dutch person has a bike of his own.

All Dutch people have a queen (16)
All Dutch people have a bike (17)

Thus, in natural language the scope ambiguity can be resolved by context, whereas in logic, this ambiguity is resolved by the order of the quantifiers and their connecting scope.

[5] Not all words that seem to be counterparts are each other's negation. For instance, "not ill" does not necessarily mean "healthy".

Existential quantifier

The logical symbol for the existential quantifier is $\exists$. The natural language equivalent is "there is". Another equivalent of the existential qualifier is "some". In logic, "some" refers to at least one individual, with the possibility of reference to all individuals. In natural language, "some" means that there is a smaller group with at least *more than one* individual *within a larger group* (Begg, 1987). This means that "some" is not used in natural language when one individual is referred to, as it would in that case be more informative to speak of "there is one". "Some" is neither used when more than half of the individuals in the group are referred to, as then "most" would be used. Furthermore, "some" is not used when all individuals in the group are referred to, as then the more informative "all" would be used. And "some are not" is not used when no individuals are referred to, as "no" would then be more informative (Begg & Harris, 1982).

1.3.5 A formal definition

In the previous sections, the differences between reasoning in natural language and reasoning with formal notations were discussed with the emphasis on those differences that can cause difficulties when learning logic. In this informal description of the logical notions and the differences with their counterparts in natural language, it became clear that most difficulties are caused by the abstractness of the logical language. In order to illustrate this difficulty and to give a complete picture, Table 1-2 gives the formal definition of the formal language of first-order predicate logic. Both the syntax and the semantics of the language are defined. The syntax describes the expressions of the language, and determines which expressions are correct and which are not. There are two types of expressions: (a) terms intended to denote individual objects in some – real or imaginative – world; and (b) propositions intended to make statements about the objects in the world. It is stipulated that the expressions, as defined by syntax, are *just* expressions. This means that they have no meaning by themselves. It is the semantics which determines the meaning of expressions explicitly. In other words, the semantics determines which objects in the world are denoted by terms, and whether propositions are true (indicated by 1) or false (indicated by 0) in the world. The semantics is defined as a function V, which assigns a value to terms or propositions. Before definitions can be given, it has to be defined what a world is and functions have to be given which define the meaning of the distinct symbols of the language, that is constants, variables and predicate symbols.

Table 1-2. First order predicate logic – formal definition

Preliminary definitions

- a *world* is a set of objects, which may have various properties, and between which several relations can exist. It is assumed that all properties and relations that are needed for the language, are present in the world;
- an *interpretation function I* gives values to elements of the language which do not change (i.e., constants and predicate symbols). So $I(a)$ is an object in the world, and $I(P)$ is a property or relation in the world;
- an *assignment function g* gives a value to a variable, so $g(x)$ is an object in the world.

Syntax

Terms can be of three kinds:

- *constants* (a, b, …) are intended to denote the same objects each time they are used;
- *variables* (x, y, …) may denote a different object each time it is used (though within the same expression it should denote the same object);
- *compound terms* are constructed from simpler terms by using so-called *function symbols*. They will not be considered in this thesis.

Propositions can be of two kinds:

- *elementary propositions* have the form $P(x)$, $Q(a,b,c)$ etc., where P, Q are predicate symbols. In general, when t_1, t_2, …, t_n are terms, and R is a predicate symbol, then $R(t_1,t_2,\ldots,t_n)$ is a proposition;
- *compound propositions* can be built up from simpler propositions by means of the logical connectives. More precisely, if p and q are propositions, then $\neg p$, $p \wedge q$, $p \vee q$, $p \rightarrow q$ are propositions. Furthermore, if x is a variable, then $\forall x\ p$ and $\exists x\ p$ are propositions too.

Semantics

The *meaning function V* determines the meaning of expressions of the language and is defined as follows:

- for constants: $V(a) = I(a)$;
- for variables: $V(x) = g(x)$;
- for elementary propositions: $V(R(t_1,t_2,\ldots,t_n)) = 1$ if and only if $V(t_1)$, $V(t_2)$, …, $V(t_n)$ together have the property or relation $I(R)$;
- for compound propositions:
    - $V(\neg p) = 1$ if and only if $V(p) = 0$;
    - $V(p \wedge q) = 1$ if and only if both $V(p) = 1$ and $V(q) = 1$;
    - $V(p \vee q) = 1$ if and only if $V(p) = 1$ or $V(q) = 1$ (or both);
    - $V(p \rightarrow q) = 1$ if and only if $V(p) = 0$ or $V(q) = 1$ (or both);
    - $V(\forall x\ p) = 1$ if and only if $V(p) = 1$ for all possible values of $g(x)$ in the world (for all other variables, g should remain the same);
    - $V(\exists x\ p) = 1$ if and only if $V(p) = 1$ for at least one value of $g(x)$ in the world (for all other variables, g should remain the same).

Although the definition of both syntax and semantics is in principle just a precise formulation of intuitions given before, this formal definition increases the abstractness of the logical language, and is therefore more difficult to read and to understand.

1.3.6 Conclusion

The analysis of everyday life reasoning versus formal reasoning shows three main differences:

A. The first concerns differences in the level of abstraction which can be subdivided into four types of abstractions which become salient in making statements about the world:
   1. Subtle differences: in natural language, the basic connectives - and, or, not, if…then - have equivalents which express subtle differences in meaning. Logic abstracts from these subtle differences and only uses the five basic connectives to express relations between predicates.
   2. Ambiguity: in natural language, sentences can be used which mean several things, that is, sentences which are ambiguous. By expressing these sentences in a context, the correct meaning can be extracted. In logic, this ambiguity has to be resolved by parentheses and by the order of the quantifiers and their connecting scope.
   3. Internal relations: in natural language, two (or more) parts of a sentence often have some kind of relation. This can be either a *contextual relation* which means that both parts express ideas about the same context; a *temporal relation* which means that the first part of the sentence expresses an event that took place before the event expressed by the second part of the sentence; or a *causal relation* which means that the event expressed in the second part was caused or triggered by the event expressed in the first part. Logic abstracts from these internal relations, as the rules of logic are valid for all possible situations and thus content does not play a role.
   4. Context: in natural language, sentences are uttered within a certain context. When humans communicate, they mostly share some kind of knowledge implicitly and use this knowledge in their conversations. Logic abstracts from this context. Every expression used during reasoning should be made explicit.

B. The second concerns differences in truth conditional aspects. Sentences stated in everyday life sometimes have different truth values compared to the same sentences approached purely logically. In everyday life, sentences can, besides being true or false, also be nonsense so that no truth value can be given or there can be uncertainty.

C. The third concerns differences in user conditional aspects. Natural language is permeated with user conditional aspects, whereas logic abstracts from these pragmatic issues.

1.4 Learning to reason with representations

Humans learn to reason during everyday life. Piaget (1976) stated, that from the age of twelve, children are able to consider possibilities, to think about abstract objects, to generate and test hypotheses; besides organising objects and events (e.g., classifying them) they can make propositions about them and explore in what way they are logically related; and they are able to systematically construct a list of all possibilities in a situation (i.e., the combinatorial system of 16 operations) and of all possible combinations of these possibilities in that situation (i.e., the INRC group of transformations). The child has shifted from the concrete operational stage to the formal operational stage. During this stage, the child also develops formal concepts which are constructed in a different way from empirical concepts (Piaget, 1973). Both types are developed by experiencing reality and Piaget distinguished two kinds of experiences in that respect: (1) the physical experience which resembles learning in the experimental sciences, and (2) the logic-mathematical experience which resembles learning in the formal sciences. The physical experience consists of abstracting information from the object itself. For instance, a child picking up balls of different sizes experiences different weights and can infer certain general rules from this. The logic-mathematical experience, however, consists of abstracting knowledge by *operating* on the objects and not of abstracting knowledge from the objects themselves. In addition to characteristics already present, new characteristics are attributed to objects. Experience, then, refers to the relation between the characteristics attributed to the objects by manipulating them or operating on them, and not to the characteristics the objects already possess. In this sense, knowledge is seen to be abstracted from the operations as such and not from the physical features of the object. For instance, a child learns the concept of order by ordering different balls to size. In this case, size is a feature all balls possess, order is added by operating on the balls. The child understands that operating on the balls does not change the characteristics of the balls themselves.
At a certain moment, the applications of operations on physical objects become superfluous and the logic-mathematical operations are being integrated in symbolic operators, which can be applied in different contexts. Therefore, from a certain moment, pure logic and mathematics are left, for which no (concrete) experience is needed. Formal concepts and operations can be abstracted from reality and these representations can be operated on mentally. However, various studies have shown that especially learning these formal, abstracted concepts without concrete experience is difficult (e.g., Pintrich, 1990; White, 1993). The analysis which was

done in the light of the present research between everyday life reasoning and formal reasoning has exposed the differences that can account for these difficulties. The conclusion is that, in order to make logic easier to learn, instruction must be adjusted in such a way that the difficulties have been resolved or taken into account. Furthermore, learning should gradually shift from this simple, concrete and context-bound everyday life knowledge towards the more abstract, complex and sophisticated models of scientific knowledge (White, 1993). Misconceptions should be addressed, so that the incorrect existing knowledge can be changed and correct existing knowledge should be addressed so that this can be extended.
In the next sections, language representations and visual representations are discussed in order to determine how learning logic can be facilitated.

1.4.1 Learning logic with language representations

As logic expresses itself in language representations, learning logic should also involve language representations. Yet, learning the rules of logic can be done in different language settings ranging from natural language as used in everyday life on the one hand to formal logic as used in abstract situations on the other hand. Both languages have advantages and disadvantages. As natural language in everyday life is permeated with pragmatics, this can easily lead to false inferences about the rules of logic. Irrelevant aspects of the problem can interfere with the reasoning process, as is also the case with earlier acquired (incorrect) experiences in everyday life situations. However, the natural language sentences can be understood very easily and no abstract symbols need to be learnt in advance. In contrast, formal logic not only abstracts from pragmatic issues, it also abstracts from meaning. Abstract rules must be applied to expressions consisting of abstract symbols. Instead of acquiring understanding of these rules, this can easily lead to shuffling symbols without understanding[6]. An advantage of abstract sentences is that pragmatic issues do not interfere with the reasoning process. Another advantage is that the logical structure of sentences becomes clear within formal representations in which abstract symbols are used (O'Brien et al., 1989). For instance, it is immediately clear that the structures of the three logical expressions (18) up to (20) are the same. In contrast, the structures of sentences (21) up to (23) appear to differ from each other. However, expressions (18), (19) and (20) are the formal representations of sentences (21), (22) and (23) respectively. So, in fact all six sentences have the same logical structure. By making the structure of sentences more visible, learners can more easily understand the rules of logic.

[6] In fact, shuffling symbols without understanding is precisely one of the intentions of formal logic, so that the rules and concepts of logic can be used by computers.

$\forall x\ (Elephant(x) \rightarrow Large(x))$ (18)
$\forall x\ (Flower(x) \rightarrow Smells(x))$ (19)
$\forall x\ (Tree(x) \rightarrow Green(x))$ (20)

All elephants are large (21)
Every flower smells (22)
For all objects it counts that, if it is a tree then it is green (23)

1.4.2 Learning logic with visual representations

The rules of empirical sciences have been inferred from experiences in everyday life situations and experiments in laboratories. By having several experiences (physical as well as logic-mathematical), regularities became salient and from these regularities rules or laws could be constructed, mostly in the form of mathematical formulae. In logic, though, agreements have been made and rules could be deduced logically, as they concerned validity of reasoning. Therefore, the rules of logic cannot be tested empirically. This can only be done when the notion of truth is introduced. For learning logic, this notion of truth and the semantics of the rules are important, as learning logic should not only involve learning the rules of logic and being able to apply them. By relating statements to reality, the truth or falsity of the statements can be determined and the meaning of operators, rules and concepts can be learnt. By relating the statements to a manipulable visual representation (which then makes up the reality), learners can learn logic from logic-mathematical experiences in this 'reality' by operating on objects and abstracting knowledge about logical concepts and operators from their operations.

Learning logic with visual representations representing reality can be done in various ways. The way in which the reality can be represented can range from a direct, everyday reality via an entirely pre-structured and well-defined reality (e.g., geometrical figures) to complete abstraction (e.g., abstract mathematical objects, elements, sets). As was the case with the language representations, both levels of abstraction of the visual representations have advantages and disadvantages, too. The drawback of learning in an everyday life context is that students are tempted to pay attention to irrelevant aspects of the problem, which can easily lead to misunderstandings. When solving problems in an everyday life context, learners will use ideas developed by experiences in everyday life in which pragmatic or user conditional aspects as preferences and hidden assumptions can play a role. Students will use their everyday life ideas about what is correct reasoning, whereas they should learn to abstract from the given context and learn to reason according to the rules of logic. Truth conditional aspects are assumed to be learnt best without any interference of user conditional aspects. Giving reality as a complete abstraction also shows some drawbacks. Learners then receive abstract, conceptual knowledge which is isolated from the situations in which this knowledge is normally used.

Learners will not always understand what the concepts and rules are about and the knowledge will not be imbedded into prior knowledge. When learners try to overcome this abstraction by imagining concrete objects for abstract expressions, they will use their own situations and in doing so mistakes can be made, which can lead to the development of incorrect knowledge.

1.4.3 Summary

Learning logic can be done with different types of representations and within these types of representations in different levels of abstraction. Abstract, formal representations have the disadvantage of being isolated from concrete, meaningful situations, whereas concrete, everyday life representations have the disadvantage of interfering pragmatic issues.

1.5 Concluding remarks

This thesis started with a short description of influences of new techniques and theories of instructional communication emerging from advancements in technology. These new ideas involved techniques and theories such as embedding rules and concepts in a (concrete) context, visualising problem situations and giving learners possibilities to interact with the representations by problem solving. As a reaction to the fact that mainly materials from the empirical sciences were used, it was assumed that these new ideas and theories could also facilitate learning formal sciences. The claim made in that respect, was that the chosen form of a representation influences the way learners perceive, and thus learn the subject matter. In order to find evidence for this claim, the topics of representations, reasoning with representations and learning to reason with representations were elaborated. An analysis was made in which different types of representations were examined and from which the difficulties learners perceive when learning logic can be inferred. It is assumed that resolving these difficulties can facilitate learning logic. In the following chapter, decisions will be made as to which representations will be used in the experiments, given the theoretical points made in this chapter.

1.6 Preview of the following chapters

The goal of the research that is described in this thesis is to study the effect of different representations of reality on the acquisition of concepts and operations of logic, in particular first-order predicate logic. In the following chapters, the way of representing the objects of logic, and the visualisation and manipulation of these will be described and their effect on learning will be studied.

Chapter 2 will focus on the representations chosen for the present project. The chapter starts with a description of the decisions that were made based on the theoretical issues described in Chapter 1. The analysis that was done to map out the similarities and differences between natural language and logic is supposed to provide grounds for decisions made in the instruction of logic used in the experiments described in this thesis. The chapter continues with a section on problem solving, as many current instructional designers agreed that the most effective way to learn is by solving problems (e.g., Dijkstra, 2000, 2001; Jonassen, 1997, 2000; Merrill, 2001). During the problem solving process, learners are stimulated to interact with the environment. These interactions result in experiences and observations. Regularities can be found and hypotheses about rules and concepts in the domain can be formed, tested and evaluated leading to the development of new knowledge or the change of incorrect, existing knowledge. In the following section, a computer-based learning environment for learning first-order predicate logic, Tarski's World (Barwise & Etchemendy, 1992), will be introduced in this respect. This environment is suitable for representing different types of representations and combinations of representations, and problems can be presented to the learners to be solved. This section is followed by a section in which the conceptualisations of the decisions made are described. This section starts with the functions of and the relations between the chosen representations, followed by a description of the conceptualisation of the problem solving process and the conditional, the connective the learners have to master.

In *Chapter 3*, a report of the first experiment is given. This experiment is partly replicated from a study done by Van der Pal and Eysink (1999). In this experiment, the added value of the first-order predicate logic representation will be studied. The aim is to determine whether a visual representation in addition to a natural language representation is sufficient to develop knowledge about logical concepts and rules or that an extra symbolic representation (the first-order predicate logic representation) is necessary in order to develop this knowledge. This leads to a comparison between a condition in which only-graphical instruction is given, a condition in which a combined graphical and formal instruction is given and a control condition. Knowledge development was measured by a transfer test consisting of a pre-, post- and retention test. The results showed that the scores of the students in the combined condition did not change significantly from pre- to post- to retention test, whereas the scores of the students in the only-graphical condition and the control condition increased from pre- to post test. The process data showed that the students in the combined condition had difficulties in translating the Dutch natural language sentences into formal expressions. Therefore, the process of translation will have interfered with the process of learning the truth conditional aspects of first-order predicate logic.

In *Chapter 4*, the study of Eysink, Dijkstra and Kuper (2001a) is described in which the combined graphical and formal instruction of the first experiment was studied more closely. It appeared that in this condition two variables were confounded: (a) the use of a graphical reality; and (b) the manipulations of the objects in this reality. In the second experiment, the influence of both these variables, visualisation and manipulation, was studied. A graphical depiction of the situation was compared to a verbal description of the situation, and a world of objects which could be manipulated was compared to a world of objects which could not be manipulated. The data showed that students needed visual depictions of the situation in order to solve the problems. Students presented with verbal descriptions produced visualisations themselves. Furthermore, the results showed that no differences were found between the students who had the possibility to manipulate and students who did not have this possibility. The process data showed that this can be explained by the fact that students did not use the possibilities of manipulation at all and that they stayed in situations they already knew the answer to.

Chapter 5 describes the study of Eysink, Dijkstra and Kuper (2001b) in which the manipulation variable was studied more closely. In order to interpret the effect of the manipulation variable on learning logic, the students had to be stimulated to use the possibilities of manipulation. Therefore, a new variable, guidance, was introduced. This variable consisted of guiding students to all possible, and especially difficult, problem situations by stimulating them to manipulate. Results showed that students not guided in their problem solving process did not profit from instruction as they did not make use of the possibilities to manipulate. Students guided in their problem solving process did profit from instruction, independently of the possibility to manipulate. The most interesting finding, however, became clear after two weeks: students who were guided and did not have the possibility to manipulate decreased their scores again after two weeks, whereas students who were guided and who had the possibility to manipulate increased their scores even more after two weeks.

Chapter 6 summarises the results of the three experiments done and discusses these. Some general conclusions concerning the instructional design and the instructional variables are given. The overall conclusion, which can be drawn from the results gained from the three experiments, is that the instructional variables often used in learning empirical sciences can also be used to develop knowledge and skills in first-order predicate logic. Furthermore, suggestions for further research and implications for logic teaching are put forward. Suggestions for further research concern the transfer tests and the manipulation variable. Implications for logic teaching concern solving the difficulties learners encounter when learning logic, the importance of learning the semantics of logic, embedding a learning environment in a social context, choosing an instructional method, choosing appropriate problem situations, and being aware of the consequences of open learning environments.

2

From theory to experiments

Abstract

On the basis of a theoretical review, decisions are made for using representations in order to facilitate the development of conceptual knowledge of logic. It is decided to choose for an instruction in which a formal system is used which is perceptible and concrete. This results in a world consisting of geometrical objects and relations between these objects. Abstract principles are related to concrete objects, so that meaningfulness and understanding can be reached and as the context is controlled, unwanted characteristics of the context do not have any influence and learners can focus on the truth conditional aspects without any interference of user conditional aspects.

2.1 Introduction

In Chapter 1, the claim that representation matters, was put forward and different types of representations and levels of abstractness were examined in that respect. This chapter tries to make a connection between the theoretical issues discussed in Chapter 1 and the experiments presented in the following chapters. The selection process of the representations and the issues playing a role in this selection process will be described in section 2.2. The instructional method that will be used for learning logic is problem solving which will be discussed in section 2.3. The chapter continues with section 2.4 on microworlds in general, followed by a description of Tarski's World (Barwise & Etchemendy, 1992), a microworld which was used in the experiments. In section 2.5, the conceptualisations of the experiments are described, that is, the function of and relations between the representations chosen are described, the way in which the problem solving process is conceptualised is described, and the conditional, the connective the learners have to master, is discussed. The chapter ends with some concluding remarks in section 2.6.

2.2 The selection of representations

In everyday life, human beings develop ideas about rules of logic. When these everyday life rules resemble the rules that logic uses, learners will not experience any difficulties. The difficulties arise when there are differences between the naïve notions learners developed in everyday life and the rules of logic. In Chapter 1, it became clear that there are truth conditional aspects and user conditional aspects in natural language and everyday life reasoning, but that logic and formal reasoning only acknowledge the truth conditional aspects and ignore the pragmatic user conditional aspects. Moreover, differences between both types of reasoning were found within the truth conditional aspects. Conclusions obvious in everyday life appear to be incorrect in logic, the so-called *counter-intuitive situations*. Furthermore, logical abstractions cause differences between everyday life reasoning and formal reasoning, too. These differences were grouped into four categories: (a) subtle differences, (b) ambiguity, (c) internal relations, and (d) context. Everyday life reasoning uses these four categories, whereas formal reasoning abstracts from them.

As one of the main goals of an introductory course in first-order predicate logic is learning the conditions under which sentences are true and the conditions under which they are false, misconceptions and counter-intuitive situations form an interesting starting point. Therefore, the focus of the present research will primarily be on these counter-intuitive truth conditional aspects of logic. As a consequence, the decision is made to use representations in which user conditional aspects are avoided, so that the learners can completely focus on the truth conditional aspects without any interference of other, unclear pragmatic aspects. This means that representations concerning everyday life objects or situations cannot be used. However, from the analysis of representations it also became clear that representations consisting of abstract objects have drawbacks as well if used to learn logic. There is no embedding into a context and real understanding is difficult to grasp. Therefore, an instruction is chosen in which a formal system is perceptible and concrete. This results in a world consisting of geometrical objects and relations between these objects. Such a world makes all operations possible and at the same time shows what happens when certain operations are applied. The world is well defined, in such a way that errors can be precluded, since irrelevant characteristics of the problem situation are left out of the context. Abstract principles are related to concrete objects, so that meaningfulness and understanding can be reached and because the context is controlled, unwanted characteristics of the context will have no influence and learners can focus on the truth conditional aspects without any interference of user conditional aspects.

Yet, by using the domain of geometrical objects to learn to reason logically, the general, abstract format of the rules of logic, which are valid in *all* situations, disappears, despite the fact that a common goal of logic teaching is to be able to apply logic in all situations. It is supposed, however, that abstract concepts can be learnt best by generalising them from concrete objects or situations, as learning of and thinking in abstract, formal terms without any reference to real world objects or situations has been shown to be difficult (Pintrich, 1990). In fact, this is precisely how Piaget (1973) stated that children learn to generalise. What is first developed by (physical and logic-mathematical) experience in situations with concrete objects, can (at a later stage) be logically deduced, too. The same can apply for students learning logic: first, knowledge can be developed by experiences in situations with concrete objects, then the underlying abstract concepts can be deduced. A representation of geometrical objects is thus assumed to support learning.

2.3 Problem solving

Many current instructional designers agree that the most effective way to learn is by problem solving (e.g., Dijkstra, 2000, 2001; Jonassen, 1997, 2000; Merrill, 2001). A problem is a question without an immediate answer. It is an unknown entity or situation that is perceived as having social, cultural, or intellectual value when solved. To give an answer, the problem has to be elaborated or a procedure to solve the problem has to be found. It requires manipulation of the problem space, either internally or externally. In instruction, a problem should either support the development of the student's knowledge or help the student to develop a procedure and practise it. Merrill stated that learning is facilitated when (a) the learner is engaged in solving real-world problems; (b) new knowledge and skills are built on the learner's existing knowledge; (c) new knowledge and skills are demonstrated to the learner; (d) new knowledge and skills are applied by the learner; and (e) new knowledge is integrated into the learner's world. During the problem solving process, learners are stimulated to interact with real objects or with representations of these objects in a learning environment. By giving them problems to solve, the learners have a goal and they can be directed to certain features of objects and to relations between features of one or more objects. They can especially be directed to situations which are often experienced as problem situations. The learners can observe the objects, interact with them by manipulation and they can ask the learning environment for feedback on the changes. These interactions result in observations and experiences. Similarities and differences between features of objects can be found, regularities in the change of objects can be observed and studied, and laws can be described. Hypotheses can be formulated about the functions of features and about the change of objects, and based on this predictions can be made. These activities lead to the development of concepts and theories in the domain and thus to the development of new knowledge or the change of incorrect, existing knowledge.

When learners form, test and evaluate hypotheses, their existing knowledge plays an important role. In situations in which the existing everyday life knowledge is correct, learners will receive confirmatory feedback and the existing knowledge remains the same or is strengthened. In situations in which the existing everyday life knowledge is incorrect and leads to incorrect predictions, the problems cannot be solved with the learners' existing knowledge. Learners will receive infirmative feedback and they can react to this feedback in different ways (e.g., Chinn & Brewer, 1993; Dunbar, 1993; Mynatt, Doherty, Tweney, 1978). They can discard or reject the feedback following the error, they can search for the error made in order to explain it, or they can try out the same, similar or other situations to experience the same or new feedback. In the latter case, the learners learn from the feedback they receive after making errors and, sooner or later, they can make inferences that lead to new knowledge or to a change and extension of existing knowledge. Naturally, these situations, in which existing everyday life knowledge is in conflict with the rules of logic, the so-called counter-intuitive situations, support and facilitate the development of new knowledge and skills.

2.4 Microworlds

Computers are extremely suitable for providing representations of all sorts. For instance, computer technology offers the opportunity to visualise representations of objects and to manipulate these objects. Furthermore, the computer can be used as a medium for instructional communication. The learner can, for example, ask for feedback, the computer can compute what effects certain actions have and it can present problems to the learner. This has led to numerous types of computer-based learning environments, one of which is the microworld. Microworlds are supposed to be a good way to learn mathematical or scientific concepts (e.g., White, 1993). The term 'microworld' was first introduced by Papert (1980, p. 122), who referred to a self-contained world in which the learners become 'the active, constructing architects of their own learning'. Edwards (1993, p. 129) called them open-ended exploratory computer environments 'which "embody" or "instantiate" some subdomain of mathematics or science'. Thompson (1987) added the qualifier 'mathematical' to distinguish environments concerning mathematical domains from environments concerning, for example, natural phenomena, microworlds which Papert (1987) called simulations. Thompson described a mathematical microworld as:

> …a system composed of objects, relationships among objects, and operations that transform objects and relationships… [It] incorporates a graphical display that depicts a visualization of the microworld's initial objects. The display in conjunction with operations upon the microworld's objects constitutes a model of the concept or concepts being proposed to the students. In a very real sense, the microworld embodies the structure of the concept. The students' task is to internalize that structure and make it their own. (1987, p. 85).

White (1993) emphasised that, in microworlds, abstract concepts are linked to concrete representations, so that abstractions become concrete, manipulable and meaningful. The underlying regularities or laws are implicitly present in the microworld waiting to be observed by the learner during problem solving. Teodoro (1992, p. 451) emphasised that the possibility of direct manipulation in such environments enables the learners 'to establish a much closer relationship with actions, which is fundamental for experiental learning'.

In the present research the definition of Thompson (1987) will be used, as it gives a precise description containing all elements considered to be important for a microworld.

2.4.1 Tarski's World

A computer-based learning environment in the domain of first-order predicate logic which meets the above-mentioned characteristics of a microworld is *Tarski's World* designed by Barwise & Etchemendy (1992). The programme was named after the logician Alfred Tarski (1902-1983) as a tribute to his pioneering work in the development of the semantic conception of logic. Table 2-1 shortly addresses the history of logic and makes some remarks about the place Tarski had in this development.

Table 2-1. The history of logic

The Greek philosopher Aristotle (384-322 B.C.) was one of the first pioneers in the field of logic. Aristotle analysed correct thinking by looking for general patterns in reasoning. He found that, by using simple inference rules, a conclusion could be deduced from two other assertions, a type of reasoning called syllogistic reasoning. A well-known example of syllogistic reasoning is "All men are mortal; Socrates is a man; therefore Socrates is mortal". Aristotle tried to unravel all general patterns which were possible in this type of reasoning. Examples of some general inference rules are the 'modus ponens' (i.e., if a true conditional $p \rightarrow q$ is given, and the antecedent p is true, then it follows that its consequence q is true) and the 'modus tollens' (i.e., if a true conditional $p \rightarrow q$ is given, and the consequence q is false, then it follows that its antecedent p is false). By analysing correct thinking, Aristotle made an important contribution to logic, later called traditional logic.

In 1854, George Boole (1815-1864) formalised this traditional logic into a mathematical theory which was the beginning of propositional logic. Later, in 1879, Gottlob Frege (1848-1925) extended this theory with quantifiers resulting in predicate logic. These formal languages were so powerful that for many domains rules could be written down in these languages by adding only a few symbols concerning the specified domain. Around 1900, Peano (1858-1932) thus formulated a theory for arithmetic in the language of first-order logic. Cantor (1845-1918) and later Zermelo (187-1953) did this for set theory and Hilbert (1862-1943) for geometry. All these developments marked the switch from traditional logic to modern mathematical logic.

Modern mathematical logic gave a framework for precise analysis of inconsistencies and paradoxes. For instance, Russell's paradox showed that, originally, set theory was inconsistent. Tarski (1902-1983) analysed the so-called Liar Paradox: "This sentence is false". If this sentence is assumed to be true, then it turns out to be false, and vice versa. The paradox brought him to a theory of truth in which he studied the relation between language and the world, often illustrated by the following example: "The sentence 'snow is white' is true if and only if snow is white". Tarski's theory of truth provides the basis for the semantics of modern logic.

During the past decades, modern logic has been expanded by types of logic which are extensions to this logical system, each describing other aspects. For example, in modal logic expressions as *possible* and *necessary* can be used; in temporal logic expressions can be used which concern time, such as *until now*, *from now on*, *ever*, *never*; in epistemic logic *believing* and *knowing* can be used; and in deontic logic obligations, such as *must* and *can* can be used.

Since logic is about reasoning in general, it is not only relevant for human reasoning, but also for computer reasoning. Modern mathematical logic, nowadays, is a direct foundation for several programming languages and it is extensively used in, for instance, artificial intelligence.

By defining a world of visible, geometrical objects with certain characteristics and relations, users learn the semantics of first-order predicate logic and they learn to determine the truth values of formulae. Within the programme, all situations are well defined and the programme is able to provide feedback on syntactic aspects and the truth of logical formulae. A typical example of a situation in Tarski's World is shown in Figure 2-1. Tarski's World is used in the experiments described in Chapter 3, Chapter 4 and Chapter 5.

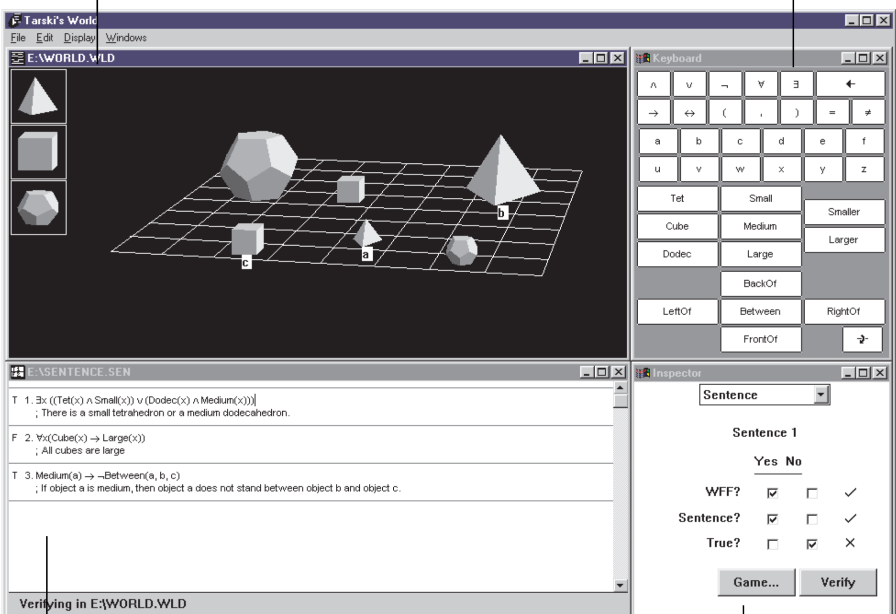
The microworld consists of four main components: (1) the world module, in which the students can place the objects of a certain size and shape in the proper position; (2) the sentence module, in which the formal sentences appeared; (3) the keyboard module for constructing sentences in the sentence module; and (4) the inspector module, in which sentences from the sentence module can be checked to verify whether they are well-formed, syntactically correct, and true/false in relation to the world in the world module.

In a typical problem, students are given natural language sentences. As the programme is not able to recognise or parse natural language sentences, these sentences appear after the semicolon in the sentence module, which is an indication for the computer that this information should be ignored. Students have to translate these natural language sentences into logical expressions. They can do this by mouse clicking on the keyboard module in which all symbols are given. When clicking on a predicate symbol, the keyboard module automatically adds parentheses, commas and the number of arguments that go with the predicate. For instance, clicking the button "Between" in the keyboard module leads to "Between(, ,)" in the sentence module, with the cursor blinking on the first instance to be filled in. This feature of the programme allows students to quickly concentrate on the semantics of the conceptions and operations of the language. When all sentences are translated, a world can be constructed. In the world module, the students can click on the objects on the left side of the screen. The objects then appear on the grid and can be given another position, size, or shape, and they can be given names. Changing position can be done by simply dragging the objects to the right position. Changing size, shape or giving names can be done by selecting the object and choosing the right size, shape or name in the part of the inspector module that automatically pops up when selecting the objects.

After translating the natural language sentences into logical expressions and after constructing a world, the students can start to evaluate their actions in the inspector module. In the inspector module, three questions can be asked. Two concern the translation of the natural language sentences into the logical expressions and one concerns the relation between the sentences and the world. When a natural language sentence is translated into a logical expression, the student can check in the inspector module whether the sentence is syntactically correct, that is whether it is a well-formed formula ("WFF?"), and whether it is a sentence, that is whether all variables are bound to a quantifier ("Sentence?").

World module: the student can construct a world of geometric objects by mouse clicking on the objects on the left of the screen. The objects can vary in shape (Tet, Cube, Dodec), size (Small, Medium, Large) and position (LeftOf, RightOf, BackOf, FrontOf, Between).

Keyboard module: the student can construct logical formulae by mouse clicking on the keys.



Sentence module: the sentences the student is constructing by using the keyboard module appear in this screen. The computer ignores the semicolons followed by sentences in English. By giving T(rue)s and F(alse)s, the computer gives feedback about the truth value of the sentences in the world given in the world module.

Inspector module: the student can check whether a logical sentence is syntactically correct (WFF?), whether all variables are bound to a quantifier (Sentence?) and whether the formula is true (True?) in the world given in the world module, by selecting a box. The computer gives immediate feedback (✔, ✖).

Figure 2-1. ***Tarski's World: world module, sentence module, inspector module and keyboard module***.

The computer cannot check whether the logical expressions are precise translations of the Dutch sentences, as the computer cannot recognise or parse natural language sentences. When a world is constructed, the students can check whether the sentences are true or false in the world ("True?").
In order to get feedback to these questions, the students have to select one of the two checkboxes ("Yes" or "No") for each question they want an answer to. When the button "Verify" is clicked, the computer gives two types of feedback. The first type of feedback is given within the inspector keyboard, which is an answer to the questions of the students. When the student answers the questions correctly, a "✓" is given behind the checkbox of the corresponding question; when the answer of the student is incorrect, a "✗" is given. The second type of feedback is given within the sentence module. When a logical expression is not a well-formed formula or a sentence, a "★" is given in front of the logical expression. In addition, when a logical expression is not a sentence, a pop-up window appears with the text "Formula is not a sentence. Free variable: x". When a logical expression is true in the world, a "T" is given in front of the logical expression, when it is false, an "F" appears. After evaluating, the logical expressions and/or the world can be changed and new feedback can be asked.
Tarski's World also contains a game-button in the inspector module. The features of this game will not be outlined here, as it is not used in the experiments.

2.5 The conceptualisations

In this section, the conceptualisations of the decisions made with reference to the theory are described. This includes a description of the representations used in the experiments, a description of the type of problems used and a description of the conditional, the connective which the students are supposed to learn.

2.5.1 The representations

This section describes the representations used in the experiments. In all experiments Tarski's World is used. Dependent on the goals of the individual experiments, the programme is adapted to fit the different conditions. In all cases, the reality represented concerns *geometrical objects*, such as cubes, tetrahedrons and dodecahedrons. The advantage of geometrical objects is their unequivocalness and the absence of difficulties concerning user conditional aspects. Furthermore, aspects as subtle differences, ambiguity, internal relations and context can be avoided easily. By presenting a world of geometrical objects, the students can focus on the semantics of logic by profiting from the advantage of concrete objects without the additional features and use of everyday life objects.

The representations used to represent these geometrical objects involve *language representations* as well as *visual representations*. The language representations can be divided into *natural language representations* and *first-order predicate logic representations*. An example of a natural language representation is given in sentence (1) and its logical counterpart is given in sentence (2). The visual representation is a *world of pre-defined 3-D geometrical objects*.

A large cube stands in front of a small tetrahedron (1)

$$\exists x \exists y\, ((Large(x) \wedge Cube(x)) \wedge (Small(y) \wedge Tet(y)) \wedge FrontOf(x, y)) \quad (2)$$

In the following sections, the function of the visual representations and the language representations will be described, followed by a description of the relations between the different representations.

Visual representations

The function of the visual representation is representing or depicting the reality. The geometrical objects on the computer screen make up the 'reality'. The objects can be considered as representations of real objects (the representing medium) or as the reality itself (the represented object). As the interface of the programme is transparent, the students will not make this distinction. There is an isomorphism between the represented objects and the representing objects, that is the geometrical objects in Tarski's World resemble prototypical geometrical objects in everyday life. The mapping principles use literal correspondence, that is the geometrical objects in the world must be interpreted literally as being geometrical objects. In the experiments, there are two kinds of visual representations, one in which the objects can be manipulated (objects can be added to the world, they can be removed and they can be given another size, shape and position) and one in which this is not possible. The manipulation of objects on the screen can be done by direct manipulation with the mouse.

Language representations

The function of the language representations is one of making (true or false) statements about the visual representation. The formal language representations consist of concrete predicates and symbols of logic. The mapping principles of the predicates are rather natural, as everybody knows what most predicates used in Tarski's World mean. Exceptions might be the predicates 'Tet' (tetrahedron) and 'Dodec' (dodecahedron) which will, therefore, be explained to the students. The symbols of logic are less straightforward and students have to learn these.

The natural language representations use the same concrete predicates as used in the formal language representations, but use natural language instead of formal symbols. Only the five basic connectives are used and ambiguities are avoided. In all experiments, the natural language representations make statements about the world. Yet, in the second experiment, the function of the natural language representations is extended. There are still natural language representations which make (true or false) statements about the world, but there are also natural language representations which *describe* the situation (instead of the visual representation). Apart from their function, these two types of natural language representations are the same.

Relations between visual and language representations

When using multiple representations, the relations between these representations are important. Figure 2-2 gives an overview of the representations used in the experiments and their relations to each other.

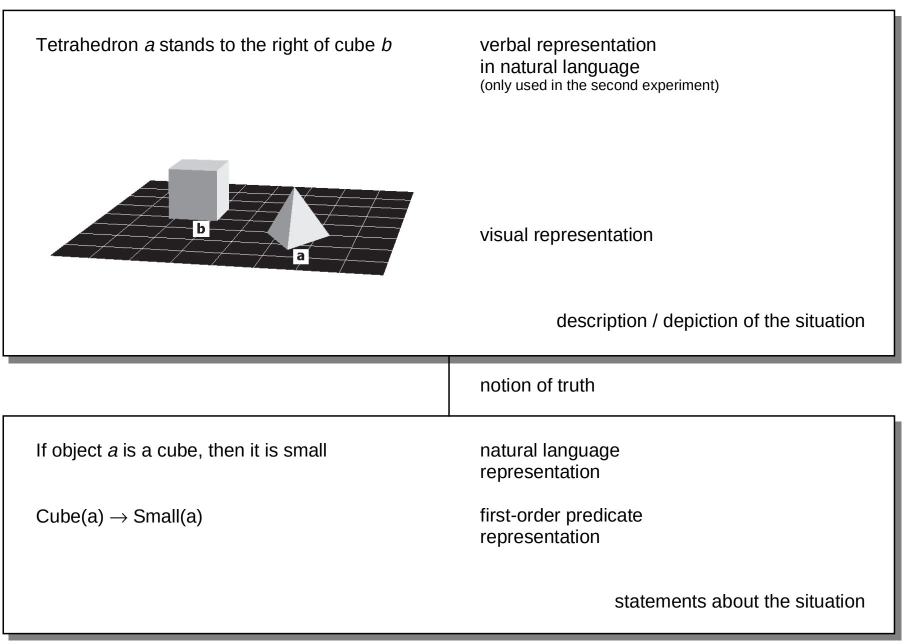


Figure 2-2. ***The functions of and relations between the representations as used in the experiments.***

In the experiments, students have to translate the natural language sentences into formal language representations (see lower part of Figure 2-2). If translated

correctly, these two language representations represent the same and are informationally equivalent (Simon, 1978). Larkin and Simon (1987) stated that two representations are informationally equivalent when all information that can be extracted from one representation can also be extracted from the other representation. Then, it is possible to construct the one from the other and the other way around.
The visual representation and the language representation describing a situation (see upper part of Figure 2-2) are not informationally equivalent. Although they both represent a situation (either by depicting it in a visual representation or by describing it in a language representation), the visual representation is more specific. A graphical representation enforces a more precise representation of certain information, whereas non-graphical representations such as language allow for abstraction or indefiniteness. This is what Stenning and Oberlander (1995) called specificity of graphic representations. A visual representation can be described by a language representation in several ways, and a language representation can result in more than one visual representation.
The relation between the representations describing or depicting the situation (the upper part of Figure 2-2) and the representations making statements about this situation (lower part of Figure 2-2) concerns a relation of truth. They are not supposed to represent the same.

The experiments

Within and between the experiments described in Chapter 3 up to Chapter 5, different representations and instructional variables are used and compared. The differences evolved as a result of the experiments and were dependent on the aims of the experiments. In the previous section, the functions of and the relations between the different representations were discussed. This section describes which representations are used in which conditions of the three experiments, so that the similarities and differences between the experiments become clear. Table 2-2 gives an overview of the representations and instructional variables in the three experiments, and Appendices A, B and C give examples of problems for each experiment and for each condition.
In the first experiment, the added value of the *first-order predicate logic representation* will be investigated. To be more specific, the aim is to determine whether a visual representation in addition to a natural language representation is sufficient to develop knowledge about logical concepts and rules or that an extra symbolic representation (the first-order predicate logic representation) is necessary in order to develop this knowledge. For this purpose, three conditions will be

compared (see Appendix A): (a) the only-graphical condition S, (b) the combined graphical and formal condition SF, and (c) the control condition C[1].

Table 2-2. ***The instructional variables used in the various conditions of the experiments***

	Condition						
Verbal description	●						
Visual depiction		●	●	●	●	●	
Manipulable repr.			●	●		●	
Natural language repr.	●	●	●	●	●	●	
First-order logic repr.	●	●		●	●	●	
Guidance					●	●	
Experiment 1			S	SF			C
Experiment 2	SN	GN		GM			
Experiment 3		Man- Guid-		Man+ Guid-	Man- Guid+	Man+ Guid+	

In the second experiment, the effect of *visualisation* and *manipulation* of objects will be investigated. To determine the effect of visualisation, a condition in which a verbal description (i.e., a natural language description) is presented (in addition to first-order predicate representations and natural language representations) will be compared with a condition in which a visual representation is presented. To determine the effect of manipulation, a condition in which a visual representation is given (in addition to first-order predicate representations and natural languages) which cannot be manipulated will be compared with a condition in which this visual representation can be manipulated. As the combined formal and visual representation and the visual representation which can be manipulated overlap, this leads to three conditions which will be compared (see Appendix B): (a) the sentential, non-manipulation condition SN; (b) the graphical, non-manipulation condition GN; and (c) the graphical, manipulation condition GM.

In the third experiment, the effect of *manipulation* of objects and *guidance* will be investigated. To determine the effect of manipulation, a condition in which a visual representation is given (in addition to first-order predicate representations and natural languages) which cannot be manipulated will be compared to a condition in which this visual representation can be manipulated. To determine the effect of guidance, a condition in which problems are given which guide the learners to all different types of basic problem situations and which stimulate them to (physically or

[1] The names of the conditions are derived from preceding experiments of Van der Pal (1995) and Van der Pal and Eysink (1999), in which the "S" stood for "Situated" and the "F" for "Formal".

mentally) manipulate the geometrical objects of the visual representation is compared to a condition in which this guidance is not given. As a result, four conditions will be compared (see Appendix C): (a) the manipulation / guidance condition Man+Guid+, (b) the manipulation / non-guidance condition Man+Guid-, (c) the non-manipulation / guidance condition Man-Guid+, and (d) the non-manipulation / non-guidance condition Man-Guid-.

2.5.2 Problem solving

In the experiments of Chapter 3 up to Chapter 5, the students will be given problems that have to be solved. The problems will be formulated in such a way that students can focus on the truth conditional aspects of logic. This means that no other connectives will be used than the standard connectives, ambiguities in sentences will be avoided as much as possible, no internal relations within sentences will be used and there will be no interfering context.

The problems will be presented according to the idea of 'model progression' (White & Frederiksen, 1990). One of the general principles of model progression is to structure the learning environment in such a way that the model of the subject matter increases from simple to complex, so that not too many ideas are introduced at the same time. This results in the concepts to be introduced in the following order: (a) predicates and constants, (b) connectives and parentheses, (c) quantifiers and variables, and (d) conditionals.

Given are two sentences:

1. If object *a* is a tetrahedron, then object *b* is a cube
2. If object *b* is a cube, then object *a* is not a tetrahedron

Assignment:

- Translate the two natural language sentences in correct first-order predicate logic.
- Construct a world in which both sentences are true.

Figure 2-3. ***A typical problem solving task as used in the experiments.***

The types of problems vary within and between the experiments. An example of a problem solving task is given in Figure 2-3. The task is formulated in such a way, that the students will encounter the counter-intuitive aspects of the conditional. The students have to translate the natural language expressions into correct first-order logic expressions. Then they have to construct a world in which all sentences are true. This can only be done by giving object a the shape of something else than a tetrahedron. By constructing the world and evaluating the truth of the sentences, the students will experience that their (everyday life) conceptions (e.g., when the first part of a conditional sentence is false, then the whole conditional sentence is false) are not always in line with the rules of logic (when the first part of the conditional sentence is false, then the whole conditional sentence is true). The experiences by interactions with the learning environment will result in new hypotheses about these rules. The goal is that the students eventually induce these general rules from the set of specific experiences.

2.5.3 The conditional[2]

All the experiments aimed at learning students the conditional, as this connective is difficult to understand for learners because of its counter-intuitive aspects. The general form of the conditional is $p \rightarrow q$. The truth table is given in Table 2-3. From this table the notions of *necessary* and *sufficient* conditions can be distilled. If p implies q, then this means that p is a sufficient condition for q, that is q *must* occur if p does. For instance, given is the sentence "If I turn on the light, then it is light in the room". In this case, the expression "I turn on the light" is a sufficient condition for "It is light in the room". But turning on the light is not a necessary condition for having light in the room, as it can be daytime as well. If p implies q, it also means that q is a necessary condition for p, that is p *cannot* occur unless q does. For instance, given is the sentence "If I turn the car key, the engine starts running". Turning on the car key is a necessary condition for a running engine, a running engine cannot occur unless the car key has been turned.

Table 2-3. ***Truth table of the conditional p → q***

p	q	$p \rightarrow q$
0	0	1
0	1	1
1	0	0
1	1	1

Note. p and q denote arbitrary propositions. 0 = false; 1 = true.

The conditional has several difficulties, one of which is the counter-intuitive case of a false antecedent. In logic, a conditional with a false antecedent is true, whereas in

[2] The term 'conditional' is often used by psychologists and programmers. The equivalent term '(material) implication' is mostly used by logicians.

everyday life reasoning such a conditional appears to be meaningless, nonsense, irrelevant or at least it cannot be told whether the statement is true or not[3].

Another difficulty is the fact that, in everyday life, conditionals are often used as if they are equivalents, that is "if and only if p, then q" (e.g., Rumain, Connell, Braine, 1983). This results in a truth table in which $p \leftrightarrow q$ is true when both p and q are either true or false. An example of this confusion is a father giving his son the conditional "If you finish your homework within an hour, then we will go to the movies". Although the child knows that father means that if his homework is not finished within an hour, then they will not go to the movies (the equivalent interpretation), the statement logically does not tell what happens when homework is not finished within an hour (the conditional interpretation). This confusion between the conditional and the equivalence becomes salient in Table 2-4 in which frequently occurring inferences associated with the conditional are presented. Two of these inferences, the 'modus ponens' and the 'modus tollens', are correct, whereas the other two, 'denial of the antecedent' and 'affirmation of the consequent', are fallacious inferences for the conditional. For the equivalence, however, all four inferences are valid.

Table 2-4. ***Four frequently occurring inferences of the conditional*** $p \rightarrow q$

First premise	Second premise	Conclusion	Inference
$p \rightarrow q$	p	q	Modus Ponens
$p \rightarrow q$	$\neg p$	$\neg q$	Denial of the Antecedent
$p \rightarrow q$	q	p	Affirmation of the Consequent
$p \rightarrow q$	$\neg q$	$\neg p$	Modus Tollens

Note. p and q denote arbitrary propositions.

The last difficulty of the conditional, already discussed in Chapter 1, is that two premises within a conditional often have some kind of internal relation. This relation depends on the function of the statement. Examples are given in Table 2-5. Logic abstracts from these internal relations.

[3] Wason (1966) suggested that people use a three-value logic for conditionals. Statements are considered to be true, false or they do not apply under the current circumstances.

Table 2-5. ***Examples of functions of conditional statements (from Evans, Newstead and Byrne, 1993)***

Contingent universal	If the animal is a fish then it is cold-blooded
Temporal/causal	If the glass is dropped then it will break
Advice	If you work hard then you will do well in life
Promise	If you clear up your toys then I will give you an ice cream
Threat	If you do that again I'll hit you
Warning	If you break the speed limit then the police will catch you
Counterfactual	If I had made some putts I would have won easily
Non-truth functional	If you want a good book to read there is one on the table

2.6 Concluding remarks

In this chapter, decisions were made concerning the types of representations to be used in the experiments of Chapter 3 up to Chapter 5. The theory discussed in Chapter 1 provides grounds for the decisions made. The result is a formal and perceptible system of geometrical objects. The representations used to represent these geometrical objects involve language representations, which can be divided into natural language representations and first-order predicate logic representations, in combination with visual representations, which can be manipulated or not. In the following chapters, experiments will be done with these representations in order to see which instructional variables can play a role in facilitating learning logic.

3

The role of formalisation of language and visualisation of objects

Abstract

The effect of a formal logical expression in addition to a visual representation in learning to use the logical connective, conditional, was investigated. Instructions for 49 first-year computer science students were varied in the computer-based learning environment of Tarski's World, designed for teaching first-order logic (Barwise & Etchemendy, 1992) resulting in three conditions (an only-graphical condition S, a combined graphical and formal condition SF and a control condition C). For the combined SF condition, the scores on the transfer tests showed no significant increase in understanding the conditional, whereas for the only-graphical condition as well as for the control condition, the scores on the transfer tests showed a significant increase. It is concluded that, in order to learn concepts and rules of logic, only-graphical instruction is not sufficient, as this condition yielded the same results as the control condition in which no further practice was given. Furthermore, instruction in the combined SF condition did possibly not focus enough on the process of translating natural language sentences into formal expressions. This led to interference between the translation process and the truth conditional aspects to be learnt, resulting in worse instead of better learning results.

3.1 Introduction

Various studies have shown that learning and using logic often causes difficulties for a majority of humans (e.g., Pintrich, 1990). They develop ideas about reasoning and rules of logic in everyday life, but these often ill-defined ideas do not always meet the rules of logic as used in formal contexts (Begg & Harris, 1982; Sanders & Antes, 1988). It is assumed that the difficulties can be explained by the fact that rules and concepts of everyday life are used in learning logic and that these everyday life ideas differ from the rules used by logic. Differences between everyday life reasoning and formal reasoning concern (a) differences in abstraction, (b) differences in truth conditional aspects, and (c) differences in user conditional aspects. Differences in abstraction are caused by the fact that logic abstracts from the subtle meaning that words can have in natural language sentences. Furthermore, logic does not recognise -and thus abstracts from- ambiguities, logic abstracts from internal relations (i.e., contextual, temporal or causal relations within sentences), and logic abstracts from the context. Differences in truth conditional aspects involve the fact that sentences stated in everyday life sometimes have different truth values than the same sentences approached purely logically. This difference can lead to conclusions which are obvious in everyday life but which appear to be incorrect in logic. Differences in user conditional aspects concern agreements which are made to take care of fluent conversations (Grice, 1975). In everyday life, it is possible to draw conclusions from sentences which, in fact, cannot be drawn from a strictly logical viewpoint. In these cases, a conclusion can be drawn because the sentence and its context suggested this conclusion.

As the main goal of an introductory course in first-order predicate logic is learning the conditions under which sentences are true and under which they are false, the differences in truth conditional aspects between everyday life reasoning and formal reasoning are the subject matter to be learnt. The remaining two differences can account for the difficulties learners encounter. The first difficulty concerns the abstractness (and the allied generality and formality) of logical expressions. Learners themselves often attribute the difficulty of learning logic to this characteristic of logic. Students indicate that they have difficulties translating sentences of natural language into sentences of formal language and that they have problems manipulating formal notations and understanding the meaning of this formal language. They find it difficult to think in abstract terms and to handle abstract concepts as used in the general, formal and mathematical techniques of logic (Fung, O'Shea, Goldson, Reeves & Bornat, 1994; Goldson, Reeves & Bornat, 1993). The second difficulty concerns the interference of irrelevant, everyday life aspects in the reasoning process of the students. These aspects can interfere with or even inhibit learning the rules of logic as used in formal logic.

It is supposed that a careful selection between different types of representations can overcome these difficulties. The chosen form of one or more representations can influence the way learners perceive, process, encode and thus learn the subject matter.

The developments in computer technology increased the choice of different representations, which led to a renewed interest in the design of instructional communication. The effect which different types of instructional variables have on the development of knowledge and skills was studied. In these studies, theories and techniques such as the effects of visualisation, the effects of multiple representations and the effects of the possibilities for the learner to interact with representations, were studied, mainly by making use of materials from the empirical sciences such as biology, chemistry and physics. These studies showed that linking abstract concepts to concrete representations led to meaningful learning and better and faster understanding (e.g., Clark & Paivio, 1991; White, 1993). Visualisation of the subject matter led to better memorisation and to more knowledge development (e.g., Clark & Paivio, 1991; Paivio, 1986; Rieber, 1996). Interaction with the subject matter led to better understanding (e.g., White, 1993).
In formal sciences such as logic, the effect of instructional variables on the development of formal concepts has not been explored in such a systematic and elaborated way. However, over the years, different representations which were supposed to facilitate learning logic have been used in education. For instance, in order to facilitate syllogistic reasoning various kinds of visual representations have been used. Examples are the Euler Circles and the Venn Diagrams in which syllogistic reasoning is visualised by using sets. By drawing and combining diagrams of the given premises, conclusions can easily be drawn from the resulting diagram. Rothbart's (1998) lessons in syllogistic reasoning concerned learning students the skill to draw valid conclusions from a given set of premises about everyday life situations stated in natural language by using abstract, symbolic expressions as cognitive tools. The students found it hard to come to a correct answer when asked to draw a logical conclusion from several given natural language statements. After translating the natural language sentences into abstract symbolic expressions, though, the students managed to solve the questions easily. Using a combination of different types of language representations helped the students to develop a greater sensitivity to language and reasoning.
Besides syllogistic reasoning, learning the meaning of the operators or connectives as used in logic has also been influenced by visual representations. An example is the educational game, Rocky's Boots (Robinett & Grimm, 1987), in which learners must construct machines which are able to select objects with or without certain characteristics. For instance, to build a machine that selects "all crosses or blue triangles", the learner must combine an OR-gate and an AND-gate. When the machine has been constructed, it can be tested. Various kinds of objects float by and when the machine has been constructed correctly those objects confirming the description will be selected (Burbules & Reese, 1984). In the game, a language

representation (i.e., the description of the types of objects to be selected) is related to a visual representation (i.e., the machine that selects floating objects meeting the description). Cox, Stenning and Oberlander (1994, 1995) studied this relation between language and reality (i.e., a visual representation in which a reality is represented) more explicitly. The computer-programme Hyperproof (Barwise & Etchemendy, 1994), which was constructed for teaching logic, was used by them to compare a combination of sentential and graphical representations with only sentential representations. They distinguished different types of cognitive styles or modality preferences in order to see what effect these differences had on learning to reason with both types of instruction. The results showed that there was an interaction between instructional design and modality preference. 'Visual learners' increased their scores when learning with both sentential and graphical representations, but decreased on the only sentential representations, whereas 'verbal learners' decreased their scores in the combined instruction and increased their scores in the only sentential instruction.

These were just a few examples of studies using different types of representations for learning logic. All the studies used multiple representations, either two types of language representations or a combination of a language representation and a visual representation. The methods differed, among other things, in the objects that were used in the statements (e.g., concrete everyday life objects, concrete but well-defined mathematical objects or abstract objects) and in the level of abstraction of the statements (e.g., natural language or symbolic expressions). In the present research, the effect of combining a visual representation with a formal language representation on learning the meaning of the connectives of first-order predicate logic will be studied. For this purpose, natural language representations, formal language representations and a visual representation are selected. The objects of communication encompass geometrical objects which can be used as concrete objects to reason with. By choosing geometrical objects, both difficulties in learning logic are resolved: (a) abstract principles are related to concrete meanings, so that meaningfulness and understanding can be reached; and (b) because the context is controlled, unwanted characteristics of the context will have no influence and learners can focus on the semantics or truth conditional aspects without any interference of pragmatic or user conditional aspects. The study described in this chapter partly replicates the study of Van der Pal and Eysink (1999) who made a comparison between three conditions: (a) F-condition: a condition in which Dutch natural language sentences had to be translated into first-order predicate logic; (b) S-condition: a condition in which Dutch natural language sentences had to be checked to be true or false in a given or self-constructed graphical representation; and (c) SF-condition: a condition in which Dutch natural language sentences had to be translated into first-order predicate logic and in which these sentences had to be checked to be true or false in a given or self-constructed graphical representation. It was hypothesised that learners in the F-condition would not profit from instruction, as they would lack feedback based on the given or self-constructed graphical representation. The formal instruction was said to lack an appropriate real-life context through which knowledge could be constructed resulting in a bad transfer from an abstract context to a real-life context. However, embedding instruction in a context (i.e., the S-condition) was neither supposed to be sufficient to yield good learning results. In this case, everyday concepts might be changed or developed during instruction, but these concepts are said to lack the necessary abstractness for transfer to other tasks or to generalise rules from specific cases. Students in this condition will have to abstract from natural language in which often the underlying structure is not clear. Therefore, the combination of formal instruction and instruction embedded in a context was supposed to lead to better results than either instructions on their own. In such a combined instruction (i.e., the SF-condition), knowledge can be constructed through activity in a context, that is the graphical representation, and the formal language can be used as a tool to abstract the formal structure underlying the natural language sentences. Results of this study showed that the combined instruction SF indeed yielded the best performances.

In the study described in this chapter, some changes have been made in comparison to the experiment of Van der Pal and Eysink (1999). One of the major adjustments is the addition of a control condition C and the omission of the formal condition F. In the control condition C, students do not receive any instruction apart from a general instruction, in which an introduction into first-order predicate logic is given. The control condition can show learning effects not caused by instruction, for instance test effects. The condition with the only-formal instruction has been left out as it is seen as a condition that cannot be compared to the other two conditions. In the F-condition, the students are supposed to translate Dutch sentences into first-order predicate expressions. After having done this, they have to check their translations in a - for them - invisible graphical representation, which is then only an extra check on their translations. In the other two conditions, the conditions in which instruction is embedded in a context by using graphical representations (condition S and condition SF), the sentences have to be checked to be true or false in these (visible) graphical representations. Thus in these conditions, the notions of truth and falsity become important, whereas this notion is neglected in the F-condition. In the graphical conditions S and SF, the learners have to check Dutch sentences to be true or false in a graphical representation. The graphical representation can be used to check the steps in the reasoning process. Yet, students in the SF-condition have an extra representation in addition to the natural language representation and the graphical representation. This extra representation, the formal logical expression, can be used to abstract from the context and to induce general rules. The formal format of the logical expression lends itself adequately to this purpose, as the structure of the expression and thus the truth conditional aspects to be learnt become salient.

In order to measure the transfer of knowledge of the students concerning the conditional, four types of the Wason selection task (Wason, 1966) will be administered. Since its introduction, the task has led to numerous studies varying it in all sorts of forms. This interest arose from the fact that despite its apparent simplicity, a great majority of subjects failed to solve it. Wason and Johnson-Laird (1972) even reported a percentage of less than 10% correct answers. The early studies (e.g., Wason, 1968, 1969, Wason & Johnson-Laird, 1972) concerned the abstract Wason task, as shown in Figure 3-1. Later studies used more thematic contents (e.g., Cosmides, 1989; Griggs & Cox, 1982; Johnson-Laird, Legrenzi & Legrenzi, 1972; Wason & Shapiro, 1971), which gave rise to even more discussions about the possible causes which made this task so difficult (for an overview, see Evans, Newstead & Byrne, 1993; for an overview and discussion, see Poletiek, 2001). Johnson-Laird and Wason (1970) developed a model in which subjects could be grouped into one of three states of insight: (a) no insight: subjects in this state attempt to verify the rule and choose either p alone or both p and q, depending upon whether they hold a conditional or biconditional reading of the conditional; (b) partial insight: subjects in this state look for choices which both verify and falsify the rule and choose p, q and *not*-q; and (c) complete insight: subjects in this state correctly try to falsify the rule and choose p and *not*-q. In the present research, the development of new knowledge will be measured by calculating the differences in correct answers between pre-, post- and retention tests and by calculating the differences between pre-, post- and retention tests in the number of students in each state of insight. These differences will be compared between conditions in order to be able to make a conclusion concerning the effect of the instructional variables.

In sum, three conditions will be compared to study whether the instructional variables, formalisation and visualisation, influence learning logic. The three conditions are a combined graphical and formal condition (SF-condition), an only-graphical condition (S-condition) and a control condition (C-condition). It is expected that the combined SF-condition will yield better results than the S-condition, as the extra, formal expression of the SF-condition makes the structure of the sentence and thus the truth conditional aspects salient. Furthermore, both graphical conditions (SF condition and S condition) in which instruction is given are supposed to lead to more knowledge development than the control condition in which no instruction is given.

3.2 Method

3.2.1 Participants

The participants were 49 first-year computer science students (46 male, three female; mean age 18.4 years, *SD* 0.9). They all took part in an obligatory course at the University of Twente in which they were given an introduction into logic. As part of the course, the students could volunteer to participate in the experiment. All students reached high math-levels in secondary school and they attended several mathematics and statistics courses after secondary school. Most students had some experience in computer programming. None of them had any experience in logic.

3.2.2 Learning environment

The computer-based learning environment Tarski's World 3.1 for Apple Macintosh (Barwise & Etchemendy, 1992) was used. Tarski's World provides an introduction in first-order logic. In the problems to be solved, a well-defined, simple world of three kinds of geometrical objects (cubes, tetrahedrons and dodecahedrons) was used. Participants could change the size of the objects (small, middle or large) and the position of the objects (to the left of, to the right of, at the back of, in front of, and between). The learning environment consisted of four main components (see Figure 2-1): (a) the world module, in which the students could place the objects of a certain size and shape in the proper position; (b) the sentence module, in which the formal sentences appeared; (c) the keyboard module for constructing sentences in the sentence module; and (d) the inspector module, in which sentences from the sentence module could be checked to verify whether they were well formed, syntactically correct, and true/false in relation to the world in the world module.

The programme was adapted to fit the experimental design. Two versions of Tarski's World were made which corresponded with the S- and the SF-condition (see Appendix A). The students in the control condition did not receive instructions in Tarski's World. In the SF-condition, the students had to translate Dutch sentences into first-order predicate expressions. Then, they had to check these sentences in the corresponding world. The students in the S-condition did not have to translate the Dutch sentences. They only had to check these sentences in the world. For this purpose, the logical expressions were made invisible for the students in the S-condition.

Moreover, the following changes were made: (a) the menu bar was made invisible, so that students were not able to give commands themselves; (b) the programme was translated from English into Dutch, so language could not interfere with the results; and (c) worlds and sentences were automatically loaded and saved when starting and finishing a task.

The instruction accompanying Tarski's World was provided by Hypercard 2.1.

3.2.3 Learning materials

The learning material comprised the *conditional*. Two sentences p and q can be combined into a new sentence with the symbol of the conditional ($\rightarrow$). The new sentences will look like $p \rightarrow q$; its English counterpart being "If p, then q".

3.2.4 Tests and questionnaires

To measure the students' knowledge of the meaning of the conditional ($p \rightarrow q$), a transfer test was administered. The test consisted of four Wason selection tasks (1966). Figure 3-1 shows a typical example of an item of this task, as used in the experiment.

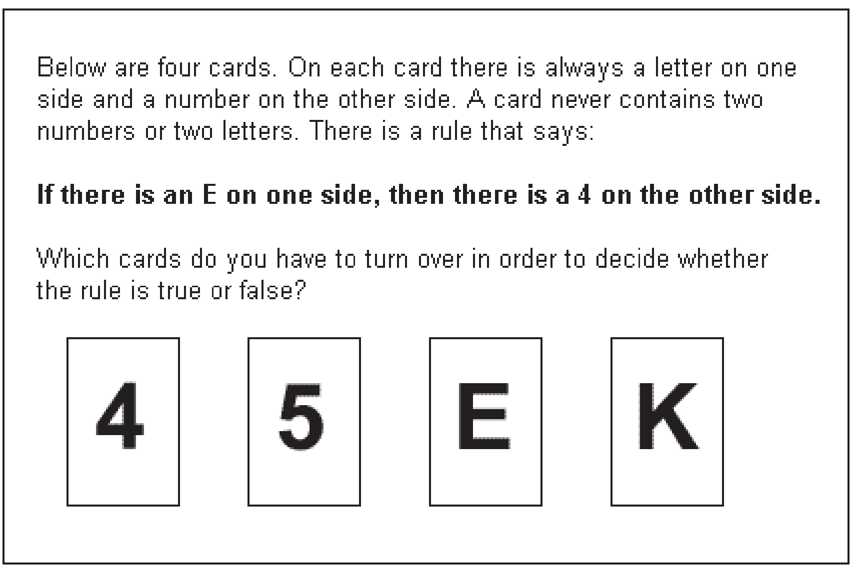


Figure 3-1. ***A typical example of the Wason selection task (Wason, 1966). The correct answer is to turn over the cards with the E and the 5 written on them.***

Logically, the cards with the E (the p-choice) and the 5 (the *not*-q-choice) have to be turned over. The E-card might have a 4 on the back which confirms the rule, or might have another number than 4 on the back which falsifies the rule. Hence, this card must be turned over. The 5-card might have an E on the back which then falsifies the rule "If there is an E on one side, then there is a 4 on the other side" or it might have another letter on the back which then does not falsify the rule. Thus, this card must be turned over as well. Furthermore, the 4-card may or may not have an E on the back, but this is of no concern as there is no claim that there must be an E on the back of a 4. So, this card cannot falsify the rules and, therefore, should not be turned over. The same applies to the K-card, because the rule has no implications for cards with letters other than the E.

In the present research, one abstract Wason (card)task and three concrete, non-arbitrary Wason tasks were used. The students had to complete three comparable versions of the transfer test, namely a pre-, post- and retention test. These tests were designed to measure the knowledge developed after the various instructions.

One questionnaire was administered concerning the former education of the students on mathematics and logic.

All students received the same tests and questionnaires, administered on the computer.

3.2.5 Log files

Two log files were generated during the experiment. The first log file logged the status of the sentences and the matching world every time the students checked this combination on well-formed formula (WFF), sentence and/or truth. The second log file logged the answers of the students on the questionnaire and on the transfer tests. Both log files contained time registration.

3.2.6 Design and procedure

Experimental conditions

Three conditions were administered (see Appendix A): (a) S-condition: a condition in which Dutch sentences had to be checked in a given or self-constructed world; (b) SF-condition: a condition in which Dutch sentences had to be translated into first-order predicate logic and in which these sentences had to be checked in a given or self-constructed world; and (c) C-condition: a condition without any instruction besides the general instruction. The subjects were randomly assigned to one of the three conditions.

Procedure

The experiment was held in three consecutive sessions; the first (pre test, general instruction and training) and the second (exercises and post test) on two successive days and the third (retention test) one week later.

The first session took about three hours and started with an introduction after which the students had to complete a questionnaire about their previous education in mathematics and logic. This questionnaire was followed by a pre test, which consisted of four Wason selection tasks. Successively, the subjects were given a general instruction in which they received an introduction into first-order logic and in which they learnt to work with Tarski's World by demonstrating the basic Tarski's World operations on the computer. After this instruction, the students in the S and SF-conditions received a general training depending on the condition to which they were assigned. During the training, the students had to complete twelve assignments in which they learnt to work with Tarski's World and with the logic operators. Students in the C-condition also received the general instruction, but did not receive the general training.

The second session started with six problems with a total of 18 sentences addressing the conditional. These exercises had to be solved by the students in the S- and SF-conditions. The students in the C-condition did not receive these exercises. After the students in the S- and SF-conditions had completed the exercises, a post test was administered to all students. The post test consisted of the same type of items as used in the pre test.

One week after the experiment, the students had to return for the retention test. This test consisted of comparable items as were used in the pre- and post test.

3.3 Results

3.3.1 Reliability

The reliability of pre-, post- and retention test, as measured with Cronbach's α, was $\alpha = .93$; $\alpha = .96$; and $\alpha = .97$ respectively.

3.3.2 Pre-, post- and retention tests

Table 3-1 shows the means and standard deviations on the pre-, post- and retention test for the three conditions C, S and SF. The maximum score was 4. A significant overall mean effect was found from pre- to post- to retention test ($F(2, 45) = 4.73$, $p < .001$). This effect was caused by a significant increase from scores on the pre test to scores on the post test in condition C ($F(1, 13) = 4.92$, $p < .05$) and in condition S ($F(1, 17) = 6.23$, $p < .05$). In condition SF, scores on the pre-, post- and retention test did not differ significantly ($F(2, 32) = .77$, $p > .05$). Furthermore, no significant differences on scores were found between conditions ($F(2, 46) = .47$, $p > .05$).

Table 3-1. ***Means and standard deviations for each condition on pre-, post- and retention tests***

	Condition		
Test	C	S	SF
Pre test			
M	.05	.94	1.18
SD	1.29	1.43	1.78
Post test			
M	1.21	1.67	1.47
SD	1.67	1.94	1.94
Retention test			
M	1.21	1.89	1.65
SD	1.67	1.97	2.03

Note. Maximum score = 4. C = control condition; S = graphical condition; SF = graphical, formal condition.

Table 3-2 shows the number of answers in several answer categories on the Wason selection tasks for the pre-, post- and retention test. Analysis showed that no differences between conditions were found in the distribution of answers in the answer categories. More than half of the answers were given in the "p-category", which did not change significantly from pre- to post- to retention test ($F(2, 96) = 1.01$, $p > .05$). The number of answers in the correct category ($p \wedge \neg q$), however, increased significantly from pre- to post test ($F(1, 48) = 6.36$, $p < .05$).

Table 3-2. ***Number of answers in a certain answer category on the Wason selection task ($p \rightarrow q$) for the pre-, post- and retention tests***

	Test		
Answer	Pre test	Post test	Retention test
$p \wedge \neg q$	45	72	79
p	106	99	93
$p \wedge q$	21	6	6
q	9	6	5
rest-category	15	13	13

Note. The answer ($p \wedge \neg q$) is the correct answer. The rest-category consists of answer categories with less than 6 responses.

Table 3-3 shows the number of students in one of the three states of insights according to the model of Johnson-Laird and Wason (1970) for each transfer test. A subject was assigned to be in a certain state when three or four of the four answers of the subject were in that state. The data show that most students started in the no-insight state and some began in the complete-insight state. After instruction, however, this distribution changed significantly ($\chi^2 = 6.26$, $p < .05$). The number of students in the no-insight state decreased, whereas the number of students in the complete-insight state increased. During the pre-, post- and retention test, hardly any students were in the partial-insight state.

Table 3-3. ***Number of Students in the Three States of Insight on the Wason selection task for the Pre-, Post- and Retention Tests***

	Test		
State of Insight	Pre test	Post test	Retention test
Complete insight	10	17	19
Partial insight	3	3	3
No insight	36	29	27

3.3.3 Process data

Students in the S-condition needed an average of 1 hr 18 min to complete the exercises and the students in the SF-condition took 1 hr 24 min to complete the exercises, which is a non-significant difference ($F(1, 31) = .72$, $p > .05$).

Furthermore, the process data show that students in the SF-condition made many errors when translating the natural language sentences into the formal expressions. The errors especially concerned using the wrong quantifier ($\forall$ instead of $\exists$, and vice versa), using the wrong connectives (e.g., $\wedge$ instead of $\rightarrow$), and the incorrect use of parentheses, such as parentheses that do not match (e.g., in the expression $\exists x\ (Cube(x) \wedge Large(x)$ a parenthesis is missing at the end) and not using parentheses resulting in free variables (e.g., in the expression $\exists x\ Cube(x) \wedge Large(x)$ the last variable x is free) and ambiguity within expressions (e.g., the expression $(Cube(a) \wedge Large(a) \vee Small(a))$ is ambiguous). Moreover, when given the feedback that they made an error, the students often (incorrectly) tried to rectify this error in the construction of their world instead of in their translation.

3.4 Discussion

3.4.1 Reliability

The reliability of the pre-, post- and retention test is high. This can be explained by the fact that the four items in the pre-, post- and retention test were all Wason tasks. Consequently, students completely understanding this task will have answered all items correctly, whereas students not understanding the task will consistently have given the same, but incorrect answers or will have been guessing.

3.4.2 Overall learning results

The overall scores on the pre-, post- and retention tests clearly show that students in the S and C- conditions profited from instructions. The students were able to solve significantly more logic problems correctly on the post- and retention tests in comparison to the pre test. Students in the SF-condition did not profit from instruction, as their scores did not increase from pre- to post- to retention test. These results contradict the hypothesis that combined formal and graphical instruction would help students to get a better understanding of the conditional.
However, especially the finding that no practice led to improvement is questioned. Students in the control condition without any further practice increased their scores from pre- to post test. This can be explained by a test-effect of the transfer tests or it is possible that all students learnt from the general training. Furthermore, the only-graphical instruction S did not lead to better learning results, as scores in this condition did not increase more from pre- to post test than scores in the control condition without any further practice. Thus, in order to learn concepts and rules of logic, only-graphical instruction is not sufficient. Finally, students in the combined formal and graphical instruction had to translate the natural language sentences into formal expressions. The data show that students had major difficulties with this translation process. It is possible that the translation became the most important hurdle to take, so that this interfered with giving attention to the truth conditional aspects to be learnt. Moreover, when given feedback that their combination of sentence and world was incorrect, the students often (incorrectly) thought that the error was caused by an incorrect world instead of an incorrect translation of the sentence. The hypothesis concerning the added value of the formal expression was that the formal expression makes the underlying structure of the natural language sentence more salient, so that truth conditional aspects could be inferred more easily. However, if the sentences were not translated correctly, then the underlying structure of the natural language sentence was not understood and the students received feedback on another structure than they thought. Therefore, in further

research, students should receive an instruction in which more attention is given to the translation of natural language sentences into formal expressions.
Thus, the results which Van der Pal and Eysink (1999) found in their study were not replicated. However, the combined instruction in the present study yielded different results to only-graphical instruction. Possibly, the translation process interfered with the learning process of the truth conditional aspects. Therefore, the translation of natural language sentences into formal, symbolic expressions needs more practice.

3.4.3 The graphical condition

The effect of a combined formal and graphical instruction was investigated in comparison to an only-graphical instruction. However, the description of the graphical instruction only included the advantage of visualisation in which the steps in the reasoning process could be checked and could be retained better. Yet, the graphical instruction included (a) a visualisation in which (b) geometrical objects could be manipulated. Thus, the question arises whether a possible effect has been caused by combining the formal sentences with the visual representation of objects, or whether the effect has been caused by the combination or interaction of both the visual representation of objects and the possibility to manipulate these objects in this visualisation.

3.4.4 The Wason selection tasks

In order to measure the transfer of knowledge of the students concerning the conditional, four types of the Wason selection task were administered. Table 3-3 shows that, in the present study, most students were in the 'no-insight' state, some were in the 'complete-insight' state and hardly any students were in the 'partial-insight' state. After instruction, this distribution changed in such a way that several students from the 'no-insight' state shifted towards the 'complete-insight' state. This indicates that, after instruction, more students knew how to solve the Wason selection task and thus more students understood the conditional. It is not an indication, however, that the students who did not answer the Wason selection tasks correctly, did not understand the conditional. As aforementioned, the Wason selection task is a difficult task in which other complicated issues such as precise formulation of the instructions and content can also influence the performance. Consequently, it is possible that the students understood the conditional within the domain of Tarski's World, but were not able to transfer their newly developed knowledge to a complicated task like the Wason selection task. Therefore, in further research it is a sensible idea to extend the transfer test with items to which the knowledge of the conditional can more easily be applied. Then, the students' knowledge of the conditional is also measured in less difficult situations leading to a clearer idea of the ability of the students to apply the knowledge of the conditional to different types of situations.

4

The role of visualisation and manipulation of objects[1]

Abstract

The effect of two instructional variables, visualisation and manipulation of objects, in learning to use the logical connective, conditional, was investigated. Instructions for 66 first-year social science students were varied in the computer-based learning environment Tarski's World, designed for teaching first-order logic (Barwise & Etchemendy, 1992). For all instructional conditions, the scores on the transfer tests showed a significant increase in understanding the conditional. Visualisation, operationalised as presenting a graphical reality or a sentential description of this reality, showed no differences on the transfer test. If presented sentential descriptions, about half of the participants needed to make drawings of the objects, especially when the problems increased in complexity. The manipulation condition, in which the participants could either construct a graphical world or were presented a fixed, graphical world, significantly influenced the participants' cognitive processes in solving the logic problems. The students worked affirmatively and were tempted to stay in familiar situations. The results support the authors' view that visualisation facilitates cognitive processing. Moreover, the results are congruent with Piaget's theory of the development of knowledge of formal science concepts from the action with objects.

---

[1] Eysink, T. H. S., Dijkstra, S., & Kuper, J. (2001). Cognitive processes in solving variants of computer-based problems used in logic teaching. *Computers in Human Behavior, 17*(1), 1-19. (with minor adjustments).

4.1 Introduction

The central concern of logic is the correctness of human reasoning. Reasoning occurs in all sciences and in all possible contexts. The rules of logic are valid in all these situations. To use the same rules in every possible situation, they must be formulated in a general way, that is, in such a way that they are not restricted to a given context. This makes logic abstract and general. Furthermore, the language of logic is formal. Agreements are made about symbols to be used and about the way these symbols can be connected to each other to form formulas in a formal and precise way.

Various studies have shown that a substantial part of all students have difficulties with learning and using these abstract and formal characteristics of logic. Students experience logic education as being difficult, too abstract and boring (e.g., Bender, 1987; Fung, O'Shea, Goldson, Reeves & Bornat, 1994; Goldson, Reeves & Bornat, 1993). Besides this, Barwise and Etchemendy (1998) stated that students often see logic as the manipulation of logical expressions by applying formal, meaningless rules. They do not get sufficient practice in finding the relation between abstract representations and real-life meanings and therefore have difficulties in applying abstract principles to everyday phenomena (White, 1993). This results in students not grasping any real understanding of the concepts and rules of logic.

In addition to this, studies have shown that abstract reasoning is difficult to improve. Only near-transfer effects (Cheng, Hollyoak, Nisbett & Oliver, 1986) or effects of years of formal training (Lehman & Nisbett, 1990) have been found. Freudenthal (1991) supposed that abstract reasoning is difficult to improve, because common-sense ideas often obstruct scientific ideas. In everyday life, people develop naïve notions about logical reasoning. Sometimes, these often ill-defined concepts and rules do not meet the rules of logic. If the learners develop certain ideas, it is difficult to change their minds and to convince them they should replace the old (incorrect) knowledge with the new (correct) knowledge. Teachers in logic are often confronted with the problem of how to teach students to solve logic problems and to translate the real world statements into formal statements. Moreover, they have to make clear to the students how to use the logical connectives, conjunction ($\wedge$), disjunction ($\vee$), conditional ($\rightarrow$) and negation ($\neg$). To clarify the meaning of logical expressions, Barwise and Etchemendy (1992) constructed a reality of computer-generated geometrical objects, which students could use to construct logical statements and to check whether these statements were true. This world of geometrical objects, which was labelled Tarski's World, was presented on a screen and could be constructed and manipulated by using a mouse. Tarski's World has been reviewed (Fung, O'Shea, Goldson, Reeves & Bornat, 1996; Goldson, Reeves & Bornat, 1993) as being an easy and fun to use programme that is 'capable of teaching a great deal about a formal language, its interpretation, models, counterexamples and consequence'. Van der Pal and Eysink (1999) designed an instruction for learning formal logic in which Tarski's World was used. In addition to formal expressions, students could manipulate objects in this graphical world. This instruction was compared to an instruction in which only formal logical expressions were given. Results showed that students who were given the experimental instruction, performed better on transfer tests than students who were given the formal instruction. In the experimental instruction, however, two instructional variables were confounded: (a) the use of a graphical reality, and (b) the manipulation of the objects. Thus, it was possible that the effect could be caused either by one of the two variables, by both or by an interaction of the two. The purpose of the present study was to investigate which variable was critical. In this respect, two issues received attention. The first concerned the extent of *visualisation*, that is, the effect on students' performances of presenting a graphical reality in addition to formal expressions. The second concerned the extent of *manipulation*, that is, the effect on students' performances of manipulating the objects presented. Both instructional variables were supposed to facilitate the students' problem solving process and thus the development of the students' knowledge and skills.

4.1.1 Abstraction and reality

In order to learn formal logic, cognitive development has to be in a stage in which formal concepts can develop and in which the logical operators can be used in a meaningful way. Piaget (1976) called this stage the formal operational period and he claimed that it is reached at about the age of twelve. During this stage, children start to think abstractly. They can formulate hypotheses without actually manipulating concrete objects, and when more adept in this, they can test hypotheses mentally. They can generalise from a real object to another and from a real object to an abstract notion.

Many adults, however, still have problems learning abstract, formal concepts without any reference to real world objects. They are unable to solve formal problems, in which only symbols are used. They can only do mental operations with real (concrete) objects, events or situations. It was estimated that this is the case for 40% to 70% of all adults (Pintrich, 1990). Freudenthal (1991) also recognised the difficulties learners have when studying mathematics. He proposed to connect the formal, abstract mathematics to reality, so that the learners could infer these formal concepts from this reality. This led to the suggestion, that learners need to be offered concrete problem situations, which can be imagined and can be used to develop mathematical knowledge and skills, so that the learners will understand the concepts and be able to work with them.

To study the relationship between abstract concepts and reality, it has to be clear how reality is described in formal sciences, and what will be the best possible way to represent this relationship. The essential characteristic of logic is that it can be applied to all situations and to all worlds. The world or set of worlds is the reality and the formal language describes this reality. The way in which this reality can be represented can range from a direct, everyday reality via an entirely pre-structured reality (e.g., geometrical figures) to complete abstraction (e.g., abstract mathematical objects, elements, sets). The drawback of learning in an everyday life context is that students are tempted to pay attention to irrelevant aspects of the problem. Students' prior knowledge consists of ideas about often ill-defined concepts and rules. When solving problems in an everyday life context, the students will use these naïve ideas in which pragmatic aspects as preferences, intuition, and hidden assumptions can play a role. Language is also permeated with *conversational implicatures* (Grice, 1975), that is, sentences often express suggestions without explicitly stating them. When learning logic by solving logic problems in an everyday life context, students will use their everyday life ideas and expressions about what is correct reasoning, whereas they should learn to abstract from the given context and learn to reason according to the rules of logic.

Giving reality as a complete abstraction also shows some drawbacks. Students then receive abstract, conceptual knowledge that is isolated from the situations in which this knowledge is normally used. Students will not always understand what the concepts and rules are about and the knowledge will not be imbedded into prior knowledge. They may only learn to shuffle the abstract symbols without comprehending what they are doing and why. When trying to overcome this abstraction by imagining concrete objects for abstract expressions, they will use their own situations and in doing so mistakes can be made.

Although a common goal of logic education is to be able to apply logic and to reason logically in all situations, the use of a graphical world to learn to reason logically will probably support learning: it makes all operations possible and at the same time shows what happens when certain operations are applied. The world is well defined, in such a way that errors can be precluded, since irrelevant characteristics of the problem situation are left out of the context. Abstract principles are related to concrete meanings, so that meaningfulness and understanding can be reached. Because the context is controlled, unwanted characteristics of the context will have no influence. Stenning, Cox and Oberlander (1995) added to this that a graphical world shares the property of specificity with the internal representations used by humans in their reasoning. It is the specificity of the world that makes it cognitively manageable and more concrete.

4.1.2 Tarski's World

The computer-based learning environment Tarski's World was designed by Barwise and Etchemendy (1992). By defining a world of visible, geometrical objects with certain characteristics and relations, users learn semantic structures of first-order predicate logic and they learn to determine the truth values of formulae. Within the programme, all situations are well defined and the programme is able to provide feedback on syntactic aspects and the truth of logical formulae. A typical example is shown in Figure 4-1.

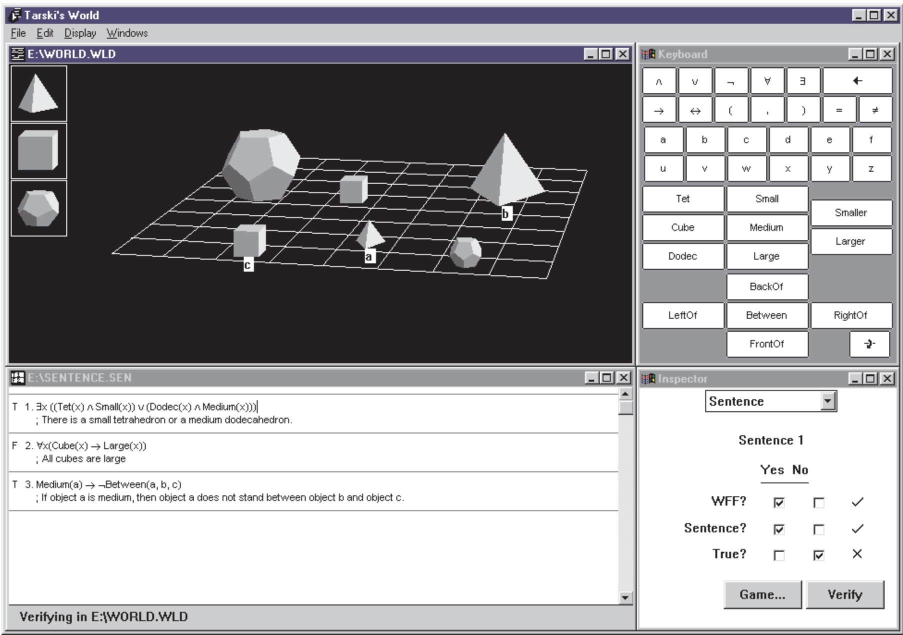


Figure 4-1. ***A typical example of a problem in Tarski's World.***

The problems that the students have to solve lead to the construction of logical expressions. By examining types of errors and sequences of errors, the students' reasoning process could be mapped out. For example, in Figure 4-1 sentence (1), the problem "There is a small tetrahedron or a medium dodecahedron" leads to the logical expression

$$\exists x \, ((Tet(x) \wedge Small(x)) \vee (Dodec(x) \wedge Medium(x))) \qquad (1)$$

If programmed accordingly, the programme is able to register all errors the students made when constructing a formula and when reasoning. As can be seen from the T's and F's at the left of the expressions in the sentence module, two of the displayed expressions are true, whereas one is false. Expression (1) claims that there must be a small tetrahedron or a medium dodecahedron. As object *a* is a small tetrahedron, the sentence is true in the world given. But as can be seen in the inspector module, the student thought Expression (1) was false. The programme provided feedback by telling the student that the expression was indeed syntactically correct, that it was indeed a sentence, but that the answer that the sentence was

false, was incorrect. These three errors, (a) whether the logical expression was syntactically correct, (b) whether the logical expression was a sentence, and (c) whether the truth value given by the student to this logical expression was correct, were being recognised by the feedback mode of Tarski's World.
Another example in Figure 4-1 is sentence (2), "All cubes are large" which has a hidden conditional. Another way of saying "All cubes are large" is "For all objects it holds that, if it is a cube, then it is large". The correct translation is given in the logical expression

$$\forall x\ (Cube(x) \rightarrow Large(x)) \tag{2}$$

However, many subjects do not recognise the hidden conditional and render "All cubes are large" as

$$\forall x\ (Cube(x) \wedge Large(x)) \tag{3}$$

which means "All objects are large cubes". In this case, the student made a faulty translation and Tarski's World is not able to recognise this. This brings us to two errors not being recognised by the feedback mode of Tarski's World. These errors are (d) errors in the translation of the Dutch sentences into logical expressions, and (e) errors during the problem solving process of finding an answer. The last two types of errors could only be traced by analysing the log files in which all the students' actions were logged.

4.1.3 The visualisation variable

In order to study whether a simple world of geometrical objects facilitates logic learning, two instructional conditions were designed and constructed. In one condition, the students were presented a graphical world in addition to Dutch sentences that had to be translated in first-order predicate logic. In the other condition, the students were given a sentential world, that is, a textual description of a world, in addition to Dutch sentences that had to be translated in first-order predicate logic. It is assumed that the students who were given a graphical world would outperform the students who were presented a sentential world. The students in the first group would be able to visually check their reasoning in the available world and easily retain the steps made. When only presented a sentential world, imaginations arise automatically on the basis of the verbal descriptions. Johnson-Laird (1989) called these imaginations the 'mental models of discourse', making explicit the structure of the situation as it is imagined instead of the exact sentence. This results in a higher load on working memory and thus a greater chance of making errors and less concentration on the conceptions and operations of logic.

Learning results were measured by administering a transfer test. It was supposed that students understood the subject matter, if they were able to apply their newly acquired knowledge and skills to new situations in which the subject matter was not learnt. The transfer test consisted of items that measured the students' ability to apply the rules of logic to everyday life problems. Psychologists have long been sceptical about the extent in which logical skills generalise to domain independent reasoning skills as people use in everyday life (see Nisbett, Fong, Lehman & Cheng, 1987 for an overview). It is widely held that for most people teaching logic only influences the algebraic symbol shuffling skills. Logic is seen as a syntactic mechanism of reasoning and because humans do not reason syntactically, teaching logic will neither help them to reason, nor to understand what their reasoning means. However, it is assumed that if instruction is given in which the reasoning can be applied to a graphical world, understanding is reached, so that this knowledge and these skills can also be used adequately in everyday life problems.
In addition to the transfer results, the authors' interest concerned the acquisition process together with the errors that were made within this process. Therefore, it was decided to study the errors that the students made during the course of the problem solving process, although there was no specific prediction about the errors that would be made. It was assumed that presenting a simple world of geometrical objects would not influence the number or type of errors concerning syntax. However, it was hypothesised that the problem solving process of the students given a graphical world was different from that of their colleagues in the other condition, because the former could use the graphical world to check the steps in the reasoning process and to use the objects to retain the steps made, whereas the latter had to imagine the world themselves and had to cognitively operate on the imagined objects. This makes the problem solving process more difficult which will manifest in students needing more time to solve the problems, needing more checks of logical expressions in worlds and making more errors.

4.1.4 The manipulation variable

To solve logic problems, knowledge of abstract objects as well as the skill to perform logical operations has to be developed. Piaget (1973) stated that learners need to act in the environment if knowledge development is to ensue. Knowledge is constructed through actions on objects in the environment. He added to this that the development of knowledge of formal concepts is realised in a different way from the development of empirical knowledge. Piaget distinguished two kinds of experiences: (1) the physical experience that resembles learning in the experimental sciences and (2) the logic-mathematical experience that resembles learning in the formal sciences. The physical experience consists of abstracting information from the object itself. For instance, a child picking up balls of different sizes experiences different weights and can infer certain general rules from this. The logic-mathematical experience, however, consists of abstracting knowledge by operating on the objects and not from the objects themselves. In addition to characteristics already present, new characteristics are attributed to objects. Experience, then, refers to the relation between the characteristics attributed to the objects by manipulating them or operating on them, and not to the characteristics the objects already possessed. In this sense, knowledge is seen to be abstracted from the operations as such and not from the physical features of the object. For instance, a child learns the concept of order by ordering different balls to size. In this case, size is a feature all balls possess, order is added by operating on the balls. The child understands that operating on the balls does not change the characteristics of the balls themselves.

At a certain moment, the applications of operations on physical objects become superfluous and the logic-mathematical operations are being integrated in symbolic operators, which can be applied in different contexts. Therefore, from a certain moment, pure logic and mathematics are left, for which no (concrete) experience is needed. Formal concepts and operations can be abstracted from reality and these representations can be operated on mentally. This theory was extended to the acquisition of concepts and rules of logic.

Therefore, it was supposed that for solving logic problems the manipulation of objects will support the development of formal logic concepts and the use of logical operators. By adding and removing objects, by changing size or position, students can see what happens with the truth value of the logical expression they construct.

To investigate whether operating on objects facilitated the development of formal concepts, two conditions were compared. In one condition, students were given a graphical world in which they could manipulate concrete objects. In the other condition, students were given a graphical world in which the objects could not be manipulated. Learning results were again measured by a transfer test. It was assumed that students who were given the opportunity to manipulate objects would profit more from the environment and better understand the meaning of the logical expression than students who lacked this opportunity. Furthermore, differences in problem solving processes were expected, although the authors did not have specific expectations of the errors made.

4.1.5 Summary

In this study, the effect of two instructional variables was studied. The first variable concerned the extent of visualisation of the subject matter: a graphically given world versus a sententially given world. The second variable concerned the extent in which the students could manipulate objects. As manipulating objects could only occur in a computer-based graphical world, the two dimensions, visualisation and manipulation, partly overlapped. As a result, three conditions were administered (see Appendix B): (a) the sentential, non-manipulation condition SN; (b) the graphical, non-manipulation condition GN; and (c) the graphical, manipulation condition GM. It was supposed that the students in the third condition would profit most from the instructions.

4.2 Method

4.2.1 Participants

The participants were 66 first-year social science students (38 male, 28 female; mean age 19.4 years, *SD* 1.0). They volunteered for the experiment for which they were paid a fee of 50 Dutch guilders (approximately € 23). None of the students had any experience in computer programming or logic.

4.2.2 Learning environment

The computer-based learning environment Tarski's World 4.1 for Windows (Barwise & Etchemendy, 1992) was used. Tarski's World provided an introduction in first-order logic. In the problems to be solved a well-defined, simple world of three kinds of geometrical objects (cubes, tetrahedrons and dodecahedrons) was used. Participants could change the size of the objects (small, middle or large) and the position of the objects (to the left of, to the right of, at the back of, in front of, and between). The learning environment consisted of four main components (see Figure 4-1): (a) the world module, in which the students could place the objects of a certain size and shape in the proper position; (b) the sentence module, in which the formal sentences appeared; (c) the keyboard module for constructing sentences in the sentence module; and (d) the inspector module, in which sentences from the sentence module could be checked to verify whether they were well-formed, syntactically correct, and true/false in relation to the world in the world module.
The programme was adapted to fit the experimental design by making three versions that corresponded with the three conditions (see Appendix B). In all three conditions, the students had to translate Dutch sentences into first-order predicate language. These sentences had to be checked in a world. In the sentential, non-manipulation condition SN, this world was given by a textual description. In the graphical, non-manipulation condition GN, the world consisted of geometrical objects, which could not be manipulated. In the third, graphical, manipulation condition GM, students had to construct and manipulate the world themselves.
The following changes in Tarski's World were made: (a) the menu bar was made invisible, so that students were not able to give commands themselves; (b) the programme was translated from English into Dutch, so language could not interfere with the results; (c) worlds and sentences were automatically loaded and saved when starting and finishing a task; (d) in the graphical, non-manipulation condition a

certain world was given which could not be manipulated by the students; and (e) in the sentential condition the graphical world was made invisible.

The instruction accompanying Tarski's World was provided in the browser of Netscape Communicator 4.06. The changes in the browser were as follows: (a) the menu options were disabled, so that students could not navigate completely freely in the browser nor surf on the internet; (b) the browser was linked to Tarski's World, so commands in one programme resulted in actions in the other programme.

4.2.3 Learning materials

The learning material comprised the *conditional*. Two sentences p and q can be combined into a new sentence with the symbol of the conditional ($\rightarrow$). The new sentence will look like $p \rightarrow q$; its English counterpart being "If p, then q".

4.2.4 Tests and questionnaires

To measure the students' knowledge of the meaning of the conditional, a transfer test of eleven items was administered. Figure 4-2 shows a typical example of an item of the Wason selection task (1966), as used in the experiment. In the transfer test, one abstract Wason (card)task, two concrete, non-arbitrary Wason tasks and one near-transfer Wason task in a Tarski's World setting was used.

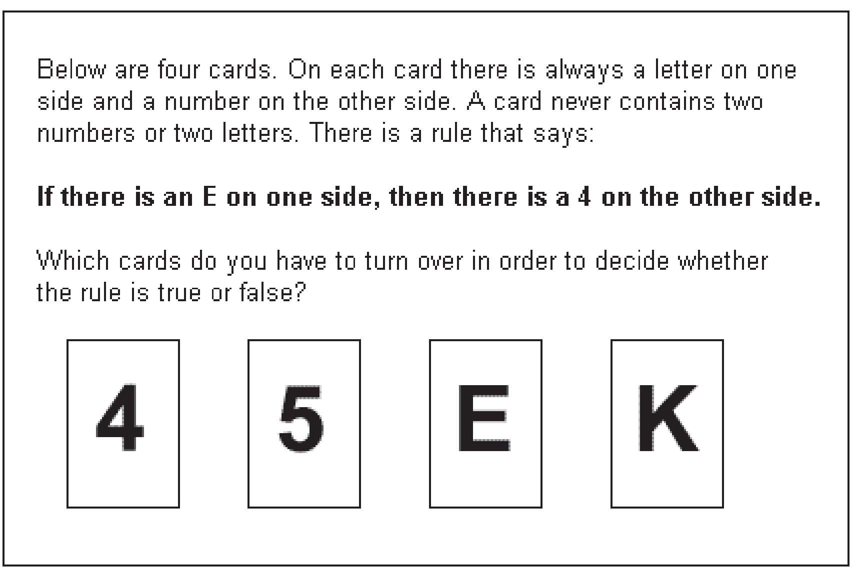


Figure 4-2. ***A typical example of the Wason selection task (Wason, 1966). The correct answer is to turn over the cards with the E and the 5 written on them.***

The remaining seven items included one Reduced Array Selection Task (RAST, Johnson-Laird & Wason, 1970), two items to be solved best by using sets and four logical deduction problems in which a statement was given and the students had to decide whether this statement was true or false or whether you could not tell from the information given. Figure 4-3 shows a typical example of the latter.

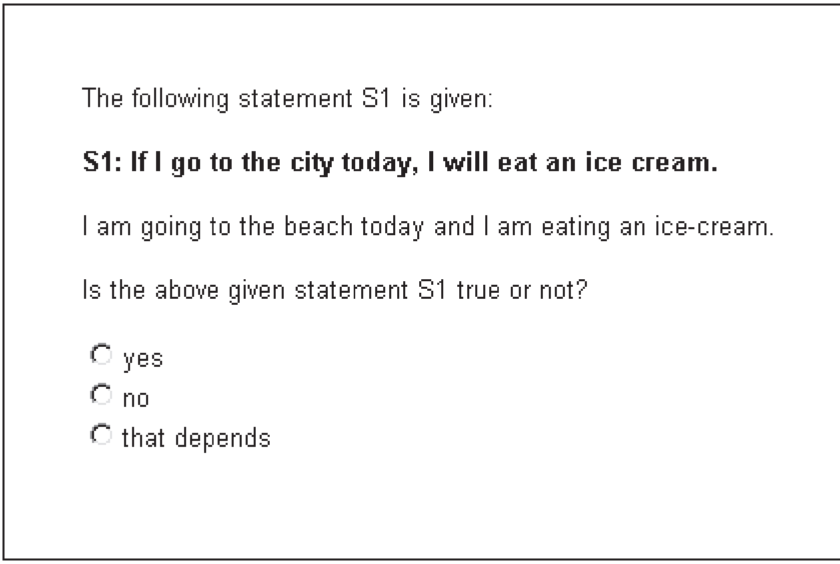


Figure 4-3. ***A typical example of a logical deduction problem. The correct answer is "yes".***

The students had to complete three comparable versions of the transfer test, namely a pre-, post- and retention test. These tests were designed to measure the knowledge gained after the various instructions. All students received the same tests.

Two questionnaires were administered. The first concerned the former education of the students on mathematics and logic. The second questionnaire was an evaluation of the instruction in combination with Tarski's World.

All tests and questionnaires were administered on the computer.

4.2.5 Log files

Two log files were generated during the experiment. The first log file logged all the actions of the students while working in Tarski's World. It logged the status of the sentences and the matching world every time the students checked this combination on well-formed formula (WFF), sentence and/or truth. The second log file logged all the actions of the students while working in the browser. This, among others, concerned answers of students on two questionnaires, answers of students on the transfer tests, and time registration.

From these log files, different results could be taken. First, the mean time students were working on the problems during general training and experimental problems could be calculated. Second, the mean number of checks per student could be computed. As the subject matter concerned the conditional with the general format $p \rightarrow q$, these checks could be divided into the four possible truth-falsity combinations ($1 \rightarrow 1$, $1 \rightarrow 0$, $0 \rightarrow 1$, $0 \rightarrow 0$, in which 1 = true, and 0 = false). Third, the mean number of errors in the final answer per student could be determined. These 'final' errors could be divided into two kinds: (a) 'indicated' errors, that is, errors indicated by the feedback module of Tarski's World, but deliberately ignored by the students; and (b) 'non-indicated' errors, that is errors not indicated by the feedback module of Tarski's World, so that the students did not detect them. When the latter type of errors occurred, it was mostly because Tarski's World was not able to check whether the logical expressions were correctly translated from the Dutch sentences.

4.2.6 Design and procedure

Experimental conditions

The subjects were randomly assigned to one of three conditions (see Appendix B). In all conditions, the students were given Dutch sentences, which they had to translate in first order predicate language. In the sentential, non-manipulation condition SN, the students had to check the sentences to be true or false in a given, sentential world. In the graphical, non-manipulation condition GN, the students had to check the sentences to be true or false in a given, static graphical world. In the graphical, manipulation condition GM, the students had to make graphical worlds in which they had to check the sentences to be true or false. In all conditions, students were allowed to use scrap paper if they wished.

Procedure

The experiment was held in three consecutive sessions; the first (pre test, instruction and general training) and the second (exercises and post test) on two successive days and the third (retention test) three weeks later.

The first session started with an introduction after which the students had to complete a questionnaire about their previous education in mathematics and logic. This questionnaire was followed by a pre test, which consisted of eleven problems measuring the knowledge of the subjects of several aspects of the conditional. Successively, the subjects received a verbal instruction in which they got an introductory course into first-order logic, as used in Tarski's World. This instruction gave the subjects an idea of what logic can be used for, what Tarski's World can do, what logic operators and quantifiers are available, how these operators and quantifiers can be used and what truth and falsity meant. This, together with some examples was the knowledge the students were equipped with. After the instruction, the students received a general training of about two hours depending on the condition to which they were assigned. During the training the students learnt to work with Tarski's World and with the logic operators. For this, model progression was used, an idea introduced by White and Frederiksen (1990). One of the general principles of model progression is to structure the rich information source and to keep the environment manageable by not introducing too many ideas at one time. Model progression entails starting with a simplified version of a model and gradually offering more complex versions of the model. In this case, the model was the field of predicate logic. The concepts were introduced in the following order: (a) predicates and constants, (b) connectives and parentheses, (c) quantifiers and variables, and (d) conditional. If the students had any questions, assistance was given by one of the experimenters present.

The second session started with six problems which had to be solved by the students in the three conditions. The first two problems were presented to refresh the knowledge acquired the day before. Consequently, four exercises addressed the conditional. In all the exercises students were asked to translate Dutch sentences into first-order predicate logic and to check the truth of the sentences in the graphical or sentential world. After the students had completed the exercises, a post test was administered. In this post test, the students were again tested on their knowledge of the conditional. The post test consisted of the same type of items as used in the pre test. Also the second questionnaire in which the instruction and Tarski's World was evaluated, was administered. Three weeks after the experiment, the students had to return for the retention test. This test consisted of comparable items as were used in the pre- and post test.

4.3 Results

4.3.1 Reliability

The reliability of pre-, post- and retention test, as measured with Cronbach's α, was $\alpha = .49$; $\alpha = .68$; and $\alpha = .75$ respectively. Deleting items from the test did not lead to significant higher reliabilities.

4.3.2 Pre-, post- and retention tests

Table 4-1 shows the means and standard deviations on the tests for the three conditions GM, GN, SN. The maximum score was 11. Scores on the pre-, post- and retention tests increased significantly for condition GM ($F(2, 42) = 3.85$, $p < .05$), condition GN ($F(2, 42) = 7.91$, $p < .01$), and condition SN ($F(2, 42) = 11.39$, $p < .001$).

To study the effect of the visualisation variable, the SN- and the GN-condition were compared. To study the effect of the manipulation variable, the GN- and GM-condition were compared. In both cases there was no difference between the conditions on the pre test ($F(1, 42) = .27$, $p > .05$ and $F(1, 42) = .03$, $p > .05$ respectively). The use of visualisation did not yield significant differences between the conditions SN and GN on the post test ($F(1, 42) = .02$, $p > .05$) and on the retention test ($F(1, 42) = .00$, $p > .05$). Neither did the manipulation of the objects in the world show a significant difference between the conditions GN and GM on the post test ($F(1, 42) = .40$, $p > .05$) and on the retention test ($F(1, 42) = 1.29$, $p > .05$).

Table 4-1. ***Means and standard deviations for each condition on pre-, post- and retention tests***

	Condition		
Test	SN	GN	GM
Pre test			
M	4.45	4.18	4.09
SD	1.97	1.50	1.74
Post test			
M	4.91	5.00	4.59
SD	2.31	2.27	1.99
Retention test			
M	6.05	6.00	5.14
SD	2.48	2.60	2.44

Note. Maximum score = 11. SN = sentential, non-manipulation condition; GN = graphical, non-manipulation condition; GM = graphical, manipulation condition.

4.3.3 Process data

The students' actions, which were stored in the log files, are summarised in Table 4-2. As can be seen, condition GM distinguished from condition GN and SN on several aspects: (a) students in condition GM needed more time to complete the experimental problems than students in condition GN ($F(1, 42) = 14.31$, $p < .001$) and students in condition SN ($F(1, 42) = 5.42$, $p < .05$); (b) students in condition GM used more checks on all sentences than students in condition GN ($F(1, 42) = 10.72$,

$p < .05$) and students in condition SN ($F(1, 42) = 13.24$, $p < .001$), especially on sentences in which both antecedent and consequent were true ($1 \rightarrow 1$); and (c) students in condition GM deliberately ignored less 'indicated' errors, although this was not significant compared to students in the condition GN ($F(1, 42) = 3.08$, $p > .05$) and compared to students in the condition SN ($F(1, 42) = 3.14$, $p > .05$), and they made far more 'not-indicated' errors, that is errors of which Tarski's World did not indicate they were made, compared to students in the GN-condition ($F(1, 42) = 7.78$, $p < .01$) and compared to students in the GM-condition ($F(1, 42) = 7.68$, $p < .01$).

Table 4-2. ***Summary of process data of students working in Tarski's World***

	Condition		
	SN	GN	GM
time[a]			
General training	1:41:20	1:28:51	1:41:46
Exp. Problems	0:24:39	0:21:32	0:30:27
Total	2:05:59	1:50:23	2:12:13
# checks			
all sentences[b]	31.5	29.7	47.4
$1 \rightarrow 1$[c]	4.3	3.9	12.6
$1 \rightarrow 0$[c]	3.7	3.5	4.2
$0 \rightarrow 1$[c]	1.8	2.1	1.3
$0 \rightarrow 0$[c]	3.8	4.0	2.1
# final errors[d]			
Indicated[e]	29	27	15
Non-indicated[f]	44	44	80

Note. SN = sentential, non-manipulation condition; GN = graphical, non-manipulation condition; GM = graphical, manipulation condition.

[a] *mean time students were working during resp. the general training, the experimental exercises, and the sum of these two.*

[b] *mean number of checks per student over all sentences during the experimental exercises.*

[c] *mean number of checks per student made on the four possible checks of the conditional ($p \rightarrow q$) during the experimental exercises (1= true; 0 = false).*

[d] *mean number of final errors students made during the experimental exercises.*

[e] *errors indicated by the feedback module of Tarski's World.*

[f] *errors not indicated by the feedback module of Tarski's World.*

All students were allowed to use scrap paper. It turned out that 59% of the students in the SN-condition made use of this possibility, especially when the problems increased in complexity. In all cases, the paper was used to draw the given, sentential world. Students in the other two conditions did not use the scrap paper.

4.3.4 Questionnaire

At the end of the experiment, 82% of the students stated that they enjoyed working with the programme. Furthermore, 72% of the students in the GM- and GN-condition stated that the concrete, visual representation made it easier to work with the logical formulae.

4.4 Discussion

4.4.1 Reliability

The fact that the reliability of the pre test was lower than the reliabilities of the post- and retention test is explained by the small number of correct answers on the pre test (M = 4.24, SD = 1.73). Apparently, the pre test was difficult, so that the students may have been guessing when answering the items, which has a negative influence on the reliability of the test.

4.4.2 Learning results

The scores on the pre-, post- and retention tests clearly show that all students profited from the instructional conditions. The students were able to solve significantly more logic problems correctly on the post- and retention tests in comparison to the pre test. Because the items of the post- and retention test also comprised the Wason selection task, the knowledge and skills acquired in Tarski's World were transferred to very different problem situations. This is evidence that far transfer is possible. The results show that, although scores on the retention test are still rather low, even non-technical students are able to do better on the difficult items of the transfer tasks after instruction. It is supposed that the findings are fostered by the advantages of Tarski's World. In this learning environment, it is easy to construct logical expressions, as the programme automatically shows parentheses, commas and the number of arguments that go with a predicate. This allows the student to quickly concentrate on the conceptions and operations of logic. Also, the sentences are easily checked on syntactical correctness by using the inspector module. The worlds allow the student to visualise the objects and their relationships to one

another and to test the truth of logical expressions in a given world. Learning occurs by successively testing logical formulae in worlds and by the immediate feedback that is given to the student.
The information on the pre-, post- and retention tests suggests, that in the interval between administering the post- and retention test, the students continued reflecting on the logic problems, which had a positive effect on the learning results. The scores on the retention tests were significantly higher than scores on the post tests. The students needed time for integrating their newly developed knowledge with their existing knowledge.
Though for all conditions the scores on the post- and retention tests increased, no differences were found between the three instructional conditions. In the experiment, the instructional variables, visualisation and manipulation, did not influence the test results. It is assumed that this result is probably due to the time spent on the instructional conditions. The students had to solve only six introductory problems. The number of steps in solving the problems was small and the students may have imagined a world and remembered the cognitive steps they made without the need for help from visualisation and/or manipulation. Also the use of scrap paper may have provided support that interfered with the visualisation variable. In a next experiment, the complexity of the problems will be increased to study the relevance of the variables for instruction.

4.4.3 Cognitive processes

The process data of the students in the graphical condition GN did not differ much from the process data of the students in the sentential condition SN. This might be due to the design of the SN-condition. By introducing scrap paper, the students in this condition were still able to use visualisations, if they needed it. In this condition, 59 % of the students used this opportunity, especially when the problems became more complex. This shows that most students need a graphical representation in which their problem solving process can be made concrete and in which steps can be retained instead of keeping these into working memory. The other 41% of the participants did not use the scrap paper to solve the problems. They were apparently able to use mental objects for this situation instead of perceivable objects.
The process data of the students in the manipulation condition GM clearly differed in number of checks and amount of time used from the data of the students in the non-manipulation condition GN. Because of these differences, the working method of the students in condition GM was given a closer look. Three findings will be discussed here.
First, students in this condition had more freedom to explore the different combinations of sentences in worlds. The students were expected to manipulate the objects in the world and to infer knowledge about the conditional by induction. However, this was not what happened. In the exercises, students first translated the

Dutch sentences into first-order logic. Then, they constructed a world that matched the first sentence. However, if it was the case that the world had to be changed to have the second sentence match the world, the students only added objects to the world. They never removed objects or started all over, not even when the problem could then not be solved correctly.
Second, it appeared that students played around in such a way that they were confronted with the subject matter they already understood, but that they did not confront themselves with the subject matter put in a new situation. They were tempted to stay in familiar situations, even when given freedom in exploring. This can be concluded from the high amount of checks in situations in which both the antecedent and consequent were true ($1 \rightarrow 1$). They kept on the safe side, repeatedly checking things they already knew, instead of trying out something new. They headed straight for the solution without straying from their path, even if this could have resulted in a better solution. Apparently, more guidance is needed to lead them to less familiar situations.
The last finding was that students in the GM-condition made a world on the basis of information from the sentences. If a sentence was about something, the students put these objects in the world. This is a way of working human beings use in everyday life. They work affirmatively, they do not start sentences using negations, and they do normally not reason about things that are not present. For instance, the sentence "all cubes are large" is a complete nonsense sentence in everyday life when no cubes are present. In logic, however, this sentence is true. Differences in language between everyday life and logic are explained by the theory of Grice (1975; for an overview, see Gamut, 1982).
In addition to differences in working method, the students in the GM-condition had an extra difficulty in their instruction. As Tarski's World was not able to recognise Dutch sentences in natural language, the programme only checked whether the sentence in first-order language was correct. As a consequence, the programme did not check whether the logical sentence was the correct translation of the Dutch sentence. Therefore, it could happen that the students thought they correctly translated the sentences, whereas this was not true. If the students then checked the Dutch sentences in the world instead of the logical sentences, it was possible that they deduced the wrong principles. So, students in the GM-condition can have learnt wrong conceptions. The data in Table 4-2 support this assertion. Students in the GM-condition were more often not aware of making wrong worlds or sentences compared to students in the other two conditions.

4.4.4 Affective reception

The students in the GM- and GN-condition were positive about the use of a world. Of these students, 72% stated that the concrete, visual representation made it easier to work with the logical formulae. Of all students, 82% stated that they enjoyed working with the programme.

5

The role of manipulation of objects and guidance of learners[1]

Abstract

The effect of two instructional variables, manipulation of objects and guidance, in learning to use the logical connective, conditional, was investigated. Instructions for 72 first- and second year social science students were varied in the computer-based learning environment Tarski's World, designed for teaching first-order logic (Barwise & Etchemendy, 1992). Guidance, which was operationalised by giving the learners problems that guided them to all different types of basic problem situations that could be derived from the conditional and stimulated them to (physically or mentally) manipulate the geometrical objects, significantly influenced the scores from pre- to post test. Manipulation, which was operationalised by giving the learners a visual representation in which (concrete) geometrical objects could be manipulated, yielded significant results from post- to retention test. The results support the authors' view that guidance in combination with the possibility to manipulate objects in a domain, support the acquisition of knowledge and skills in that domain. Guidance leads to the students seeing all situations that are relevant for the development of new knowledge, which has a positive effect directly after instruction. The possibility to manipulate leads to the students playing around in such a way that they (systematically) experience the results of their actions by the feedback given, having an effect after two weeks of non-instruction.

---

[1] Eysink, T. H. S., Dijkstra, S., & Kuper, J. (2001). The role of guidance in computer-based problem solving for the development of concepts of logic. Manuscript submitted for publication. (with minor adjustments).

5.1 Introduction

Various studies have shown that a substantial part of all students in higher education have difficulties learning and using logic. Logic is often seen as boring and difficult (e.g., Bender, 1987; Fung, O'Shea, Goldson, Reeves & Bornat, 1994; Goldson, Reeves & Bornat, 1993) as well as meaningless (e.g., Barwise & Etchemendy, 1998). Students indicate that they especially have difficulties translating sentences of natural language into sentences of formal language, that they have problems to break down a problem into a series of smaller problems to be solved, and that they have trouble manipulating formal notations and understanding the meaning of this formal language. Fung et al. concluded that these difficulties are often caused by a lack of mathematical background and an unfamiliarity with formal notations and language.

In sum, students generally find it difficult to think in abstract terms and to handle abstract concepts as used in the general, formal and mathematical techniques of logic. This can either be caused by the fact that the cognitive development of the students has not yet reached the stage of development in which abstract concepts can evolve, or by inappropriate instruction. In his theory of cognitive development, Piaget (1976) stated that children reach the formal operational stage in which mathematical abstract concepts can develop at the age of twelve. He said, that at this age, children start to think in abstract terms. They start to generalise from (classes of) real objects to abstract notions. The child learns to think abstractly by deducing abstract concepts from *concrete* objects or situations. Learning of and thinking in abstract, formal terms without this reference to real world objects or situations is difficult, even for a large percentage of adults (Pintrich, 1990). People often develop and use abstract concepts in close connection to concrete objects in their everyday world. This then raises the questions what makes mathematical and logical abstraction so difficult for most people, and what instructional variables play a role in order to improve mathematical and logical skills. In this article, several instructional variables in logic teaching are discussed, including the connection between abstract concepts and concrete objects. A computer-based learning environment in which this connection has been realised is Tarski's World (Barwise & Etchemendy, 1992). In this environment, information from formal logical sentences is combined with information from visual representations of (concrete) geometrical objects. As a result, students can focus on the semantics of logic rather than on the syntax.

In order to study the idea that connecting formal sentences to concrete objects would indeed lead to a better understanding of the concepts of logic, Van der Pal and Eysink (1999) designed an instruction in which Tarski's World was used. They compared three conditions: (a) a condition in which situations were described by formal sentences (sentential condition); (b) a condition in which situations were described by visual representations of geometrical objects (graphical condition); and (c) a condition in which situations were described by formal sentences as well as visual representations of geometrical objects (combined sentential and graphical condition). The results showed that the last condition yielded the best performances, although an unpublished replication experiment (Eysink, 1998) did not confirm these results. Thus, a clear answer to the question whether the combination of formal sentences and concrete objects yielded better understanding of the concepts of logic could not yet be given. Further analyses of both studies revealed that, in the combined instruction, two instructional variables were confounded. The subject matter was represented as (a) a visualisation in which (b) geometrical objects could be manipulated. Thus, the question arose whether a possible effect was caused by combining the formal sentences with the visual representation of objects, or whether the effect was caused by the combination or interaction of both the visual representation of objects and the possibility to manipulate these objects in this visualisation. Moreover, two claims were made concerning visualisation and manipulation of objects. The first claim concerning visualisation was based on the phenomenon which Johnson-Laird (1989) called the 'mental models of discourse'. When presented sentential descriptions, imaginations or mental models arise automatically on the basis of these descriptions. The mental models make explicit the structure of the situation as it automatically arises from the sentential descriptions instead of making explicit the structure of the situation as described exactly by the sentence. As logic is precise, these slightly different imagined visualisations can cause a greater chance of making errors. Furthermore, a visual representation gives the possibility to visually check the reasoning process in the available visualisation and to easily retain the steps made in this reasoning process. Thus, it was expected that giving a visual representation of geometrical objects would lead to better results than giving a sentential description of the situation.

The second claim on manipulation of objects was derived from Piaget's theory (1973) that learners need to act in the environment if knowledge development is to ensue. Knowledge is constructed through actions on objects in the environment. The logic-mathematical experience (in contrast to the physical experience) consists of abstracting knowledge by operating on several objects. In addition to characteristics already present, new characteristics are attributed to objects by these operations. Experience, then, refers to the relation between the characteristics attributed to the objects by manipulating them or operating on them. In this sense, knowledge is seen to be extracted from the operations as such and not from the physical features of the object, as is the case with physical experience. For instance, a child can learn the concept of order by the activity of ordering different balls to size. In this case, size is a physical feature the balls already possess, order is a logic-mathematical concept being added by actually operating on the balls. A parallel can be drawn to the logical connectives, conjunction ($\wedge$), disjunction ($\vee$), conditional ($\rightarrow$) and negation ($\neg$). Learners can learn the concepts of these connectives by actually manipulating or operating on the concrete sentences connected to each other by one (or more) of the connectives and they can see whether the combined sentence is true or false. This can be elaborated on further by creating a possibility to manipulate objects in a visual representation as described by sentences. Thus, by introducing real objects that can be manipulated, the development of understanding logical combinations of sentences can be supported. Therefore, it was expected that students having the possibility to manipulate, would profit more from the instructional material than students who did not have this possibility.

In order to study the correctness of these claims, Eysink, Kuper and Dijkstra (2001a) designed an experiment in which two instructional variables, *visualisation* and *manipulation* of geometrical objects were investigated. For this purpose, three conditions were administered: (a) a sentential, non-manipulation condition SN; (b) a graphical, non-manipulation condition GN; and (c) a graphical, manipulation condition GM. The results showed that not all the claims made were confirmed. First of all, students who were given a visual representation of geometrical objects did not profit more from the instructional material than students given a sentential description. However, this was explained by the fact that students given a sentential description had the possibility to draw their own visual representation, something the students frequently used, especially when problems became more complex. Clearly, the students needed a visualisation to solve the problems. Secondly, students who were given the possibility to manipulate did not outperform students not given this possibility. A closer look and further analysis of the data showed that this result could be explained by the fact that students did not fully use the possibilities of manipulation. They played around in such a way that they were confronted with the subject matter they already understood, but that they did not confront themselves with the subject matter put in a new situation. The students were tempted to stay in familiar situations, even when given freedom to explore. They repeatedly checked things they already knew, instead of trying out something new and they headed straight for the solution without straying from their path, even if this could have resulted in a better solution.

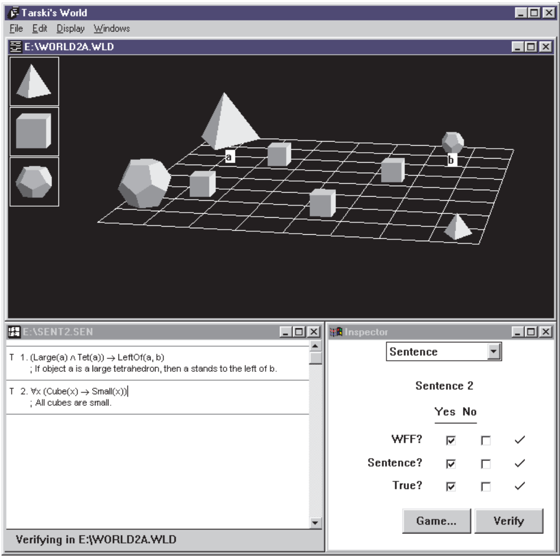
The finding that students did not fully use the possibilities of manipulation fits in with results from other studies. Several researchers (e.g., Dunbar, 1993; Klayman & Ha, 1987; Kuhn, Schauble & Garcia-Mila, 1992) have found that learners tend to stay in familiar situations when testing a certain hypothesis, the so-called *confirmation bias*. They mainly conduct experiments that confirm their current hypothesis. They keep repeating the same situations to which they (think they) know the answers. The learners do not think of conducting experiments that will disconfirm their current beliefs. When, nevertheless, confronted with evidence that is inconsistent with their current hypothesis, they just ignore this evidence (e.g., Chinn & Brewer, 1993; Dunbar, 1993; Mynatt, Doherty & Tweney, 1978). Instead of changing their current theories, learners reject or distort these findings. One way of doing this, is stating that the anomaly in the data must have been caused by error. In order to test this, learners can conduct a series of the same experiments, and only when all these experiments give disconfirming evidence, they will start thinking of a new explanation for this finding.

Apparently, most learners are not able to regulate their own problem solving process. In complex domains, the learners can have difficulties getting an overall notion of all aspects that cover the domain. When no feedback is given on the suitability of the working method used by the learner, this can lead to unsystematic behaviour resulting in missing parts of the subject matter, not seeing what is important and what are side-issues, and working without a goal or mission, even if the problem is well defined. In addition, learners find it difficult to doubt and change existing knowledge and to increase their skills.

A solution to this problem is to regulate the problem solving process by giving guidance to learners. Guidance is a form of instruction that supports and facilitates the development of knowledge and skills. In the present study, guiding the learners in their problem solving process will be realised by giving problems in which the learners are directed to all possible problem situations, especially to those in which the development of new knowledge is most likely to occur, that is in situations in which problems are presented that cannot be solved with the learners' existing knowledge. In these situations critical for the development of new knowledge, the learners experience that the feedback they receive is in conflict with their existing ideas about logic. Learners can react on this infirmative feedback in different ways. They can discard or reject the feedback on the error, they can search for the error made in order to explain it, or they can try out the same, similar or other situations to experience the same or new feedback. In the latter case, the learners learn from the feedback they receive after making errors and, sooner or later, they can make inferences that lead to new knowledge or to change and extension of existing knowledge. However, the firmer the knowledge base of existing beliefs, the harder it will be to abandon or modify these beliefs and the clearer the evidence for an alternative theory should be, so the more often the learners should be confronted with the subject matter involved. Without guidance, no supervision or feedback is given on the working method of the learner. This can increase the possibility of working unsystematically without a goal. Learners can get lost, they are not forced to study every possible situation and thus they will not as often be confronted with situations that are critical for the development of new knowledge, as they can decide not to confront themselves with situations they do not know the answer to, and after having been confronted with unexpected feedback the learners can decide to avoid this problem situation. In such situations, guidance of the learner may help to develop new knowledge.

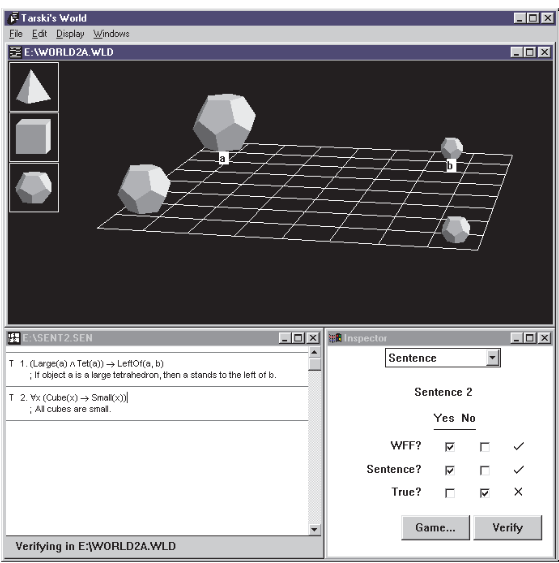
In order to investigate the influence of guidance on learning logic with or without the possibility to manipulate objects, an experiment needs to be designed in which two instructional variables, *manipulation* and *guidance*, will be varied systematically. Manipulation will be operationalised by giving the students the possibility to manipulate geometrical objects in a visual representation. By adding and removing objects, and by changing their features like size or position, students will be able to observe and (physically) experience whether the logical sentences they constructed become true or false in the visualisation and what effect certain changes in the visualisation has on the truth or falsity of the logical sentences. The group of students in the non-manipulation conditions will be given a fixed visual representation of geometrical objects in which they can evaluate whether the logical sentences are true or false.

The guidance variable will be operationalised by giving problems in which the students are directed to all possible problem situations, especially to those problem situations that are in conflict with the students' existing knowledge, so that new knowledge can be constructed. In these problem situations the students will be given tasks to (mentally or physically) manipulate the geometrical objects in the visual representation, so that the students experience what influence every change in the visualisation has on the truth or falsity of the elements of a sentence and what impact this has on the truth or falsity of the complete sentences composed of the elements combined by one or more connectives. By doing this in such a way that the different types of basic problem situations that can be derived from the conditional are all presented, the students can experience what impact every change in the visualisation has on the formal logical sentences concerning their truth or falsity. A typical example of a problem in which the students are guided, is given in Figure 5-1. In the upper part of Figure 5-1, the student has already solved the first two tasks of the problem presented. The natural language sentences have been translated into first-order predicate language and the truth or falsity of the sentences has been checked by the student. The answer of the student appeared to be correct: both sentences are indeed true in the world given, as in both sentences the first part of the sentence as well as the second part is true. The next task of the problem concerns the guidance to a problem situation of which students are assumed to have difficulties with, namely the situation in which the first part of the sentence is false and the situation in which the objects mentioned in the first part do not exist at all. The students are stimulated to change the world in such a way that they are confronted with both problem situations (see for the changed world the lower part of Figure 5-1). Both sentences are checked again in this new problem situation. The existing everyday life knowledge of the student can easily lead to the conclusion that both sentences have become false by the change made. The feedback given by the computer programme shows the student that this answer is not correct. If the first part of the sentence before the conditional is false or if the objects mentioned in this part do not exist, the complete sentence is true. The guidance to these problem situations and the feedback of the programme on the answers given by the students can cause the students to change existing knowledge or develop new knowledge.



<u>Assignment:</u>

1. Translate the two given sentences into correct logical expressions.
2. Predict whether the sentences are true or false in the given world and verify this
3. Change the world in such a way that:
   - no cubes are left
   - all tetrahedrons are changed into dodecahedrons
4. Predict again whether the sentences are true or false in the changed world.



<u>Assignment:</u>

1. Translate the two given sentences into correct logical expressions.
2. Predict whether the sentences are true or false in the given world and verify this
3. Change the world in such a way that:
   - no cubes are left
   - all tetrahedrons are changed into dodecahedrons
4. Predict again whether the sentences are true or false in the changed world.

Figure 5-1. ***A typical example of a problem situation in which the students are guided in their problem solving process.***

When systematically confronting the students with the different kinds of situations that are possible and showing which changes in the truth or falsity of the elements of a sentence are responsible for changes in the truth or falsity of the combined sentence, the students will realise that their initial, everyday life beliefs about logic are not always consistent with the rules logic holds. It is supposed that this newly learnt knowledge will (eventually) change the students' concepts of logic.

The students who are not guided in their problem solving process, will be given problems in which they have the possibilities to create the same problem situations as the students who receive guidance, with the distinction that the students who are not guided, will not especially be directed to critical problem situations in which new knowledge is most likely to develop. Furthermore, students who are not guided, will not be stimulated to (mentally or physically) manipulate the geometrical objects in the visual representation.

Students are only supposed to profit from the possibility to manipulate if this possibility is combined with problems that guide the students through the subject matter to all possible, and especially critical situations. When guided through all different situations, learners receive feedback on every possible situation. Manipulation without guidance is supposed to be less helpful, as students in this case probably will only stay in situations they already know the answer to and they will not confront themselves with all possible situations of the subject matter. The feedback they receive on their actions, therefore, does not include all situations, and especially not those situations they do not know the answer to or situations they are uncertain about, that is situations that are critical to develop new knowledge. Students not guided through the subject matter will profit less from the material presented, independently whether they have the possibility to manipulate or not, as the advantage of manipulation is supposed not to reveal without guidance. In sum, it is supposed that students who are guided will more easily develop new knowledge, especially when they have the possibility to manipulate objects in the world. Students not guided will profit least from the material presented, independent of the possibility to manipulate.

5.2 Method

5.2.1 Participants

The participants were 72 first- and second-year social science students (18 male, 54 female; mean age 19.0 years, *SD* 1.1). They volunteered for the experiment for which they were paid a fee of 50 Dutch guilders (approximately € 23). None of the students had any experience in computer programming or logic.

5.2.2 Learning environment

The computer-based learning environment Tarski's World 4.1 for Windows (Barwise & Etchemendy, 1992) was used. Tarski's World provides an introduction in first-order logic. In the problems to be solved, a well-defined, simple world of three kinds of geometrical objects (cubes, tetrahedrons and dodecahedrons) was used. Participants could change the size of the objects (small, middle or large) and the position of the objects (to the left of, to the right of, at the back of, in front of, and between). The learning environment consisted of four main components: (a) the world module, in which the students could place the objects of a certain size and shape in the proper position; (b) the sentence module, in which the formal sentences appeared; (c) the keyboard module for constructing sentences in the sentence module; and (d) the inspector module, in which sentences from the sentence module could be checked to verify whether they were well-formed, syntactically correct, and true/false in relation to the world in the world module.
The programme was adapted to fit the experimental design by making four versions that corresponded with the four conditions (see Appendix C). In all four conditions, the students had to translate Dutch sentences into first-order predicate language. These sentences had to be checked in a world. In the manipulation conditions, the geometrical objects in the world could be manipulated and new worlds could be constructed. In the non-manipulation conditions, the given worlds consisted of geometrical objects that could not be manipulated. In the guidance conditions, problems were given that led the students systematically to all different types of basic problem situations that can be derived from the conditional, with the emphasis on those problem situations that were supposed to be difficult and that would lead to the development of new knowledge. Additionally, the problems stimulated the students in (mentally or physically) manipulating the geometrical objects in the world, so that changes from the existing problem situation to a new situation and their effects on the truth or falsity of the logical sentences become clear. In the non-guidance conditions, the students were not guided in their problem solving process to explore what the effect was of the different kinds of connectives, including the conditional, on the truth or falsity of complex sentences combined by these connectives and they were not stimulated to (mentally or physically) manipulate the geometrical objects in the world. They were given problems in which they had the possibilities to create the same problem situations as the students who were guided, with the distinction that the students who were not guided were not especially directed to critical problem situations. Furthermore, students who were not guided were not stimulated to (mentally or physically) manipulate the geometrical objects in the visual representation.

The following changes in Tarski's World were made: (a) the menu bar was made invisible, so that students were not able to give commands themselves; (b) the

programme was translated from English into Dutch, so language could not interfere with the results; (c) worlds and sentences were automatically loaded and saved when starting and finishing a task; and (d) in the non-manipulation conditions worlds were given which could not be manipulated by the students.
The instruction accompanying Tarski's World was provided in the browser of Netscape Communicator 4.06. The changes in the browser were as follows: (a) the menu options were disabled, so that students could not navigate completely freely in the browser nor surf on the internet; and (b) the browser was linked to Tarski's World, so commands in one programme resulted in actions in the other programme.

5.2.3 Learning materials

The learning material comprised the *conditional*. Two sentences p and q can be combined into a new sentence with the symbol of the conditional ($\rightarrow$). The new sentence will look like $p \rightarrow q$; its English counterpart being "If p, then q".

5.2.4 Tests, problems and questionnaires

To measure the students' knowledge of the conditional ($p \rightarrow q$), transfer tests were administered. These tests were divided into three parts with a total of 27 items. The first part, four puzzles consisting of 14 items, tested the knowledge of the students on the conditional within the domain of Tarski's World. The students were given a certain situation which was either a world with several geometrical objects or a sentential description. In addition, several statements in natural language were given. The students' task was to tell whether these statements were true or false in the given situation or whether it was valid or not for certain conclusions to be drawn from the statements. Figure 5-2 shows a typical example of the items given in the first part of the transfer test.
The second part, consisting of nine items, tested the knowledge of the students on the conditional within an everyday life setting. The students were given a statement and a conclusion about an everyday life situation and they had to decide whether the conclusion was true or false or whether you could not tell from the information given. Figure 5-3 shows a typical example of the items given in the second part of the transfer test.
The third part of the test consisted of four items testing the conditional. These items involved the Wason selection task (Wason, 1966). Figure 5-4 shows a typical example of an item of this task, as used in the experiment. One abstract Wason (card)task, two concrete, non-arbitrary Wason tasks and one near-transfer Wason task in a Tarski's World setting were used in the third part of the transfer test.
The students had to complete three comparable versions of the transfer test, namely a pre-, post- and retention test. These tests were designed to measure the

knowledge gained after the various instructions. All students received the same tests. Appendix D gives the complete version of the pre test.

Given are two **true** statements:

1. If there is a cube, then this cube is large.
2. If there is a tetrahedron, then this tetrahedron is small.

A. Given is that there is a cube.

- Is this cube large?

  ○ yes, this cube is large

  ○ no, this cube is not large

B. Given is that there are no small objects.

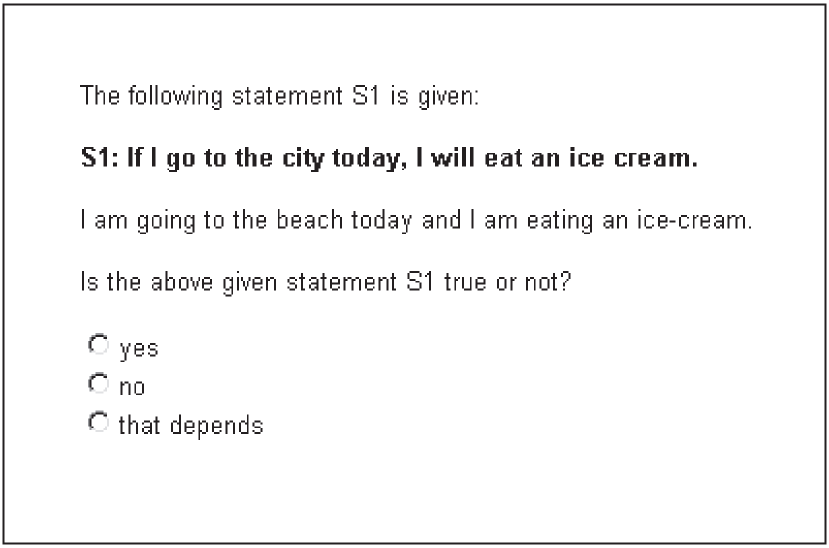
- Is there a tetrahedron?

  ○ yes, there is a tetrahedron

  ○ no, there is no tetrahedron

Figure 5-2. ***A typical example of the items given in the first part of the transfer tests measuring knowledge about the conditional within the domain of Tarski's World. The correct answers are "yes, this cube is large" and "no, there is no tetrahedron".***

During the experiment, students were given four different types of problems: (a) investigation problems type I: problems in which the students had to investigate all basic combinations of truth values of several self-made sentences in a self-made or given world; (b) investigation problems type II: problems in which the students (mentally or physically) had to investigate all basic combinations of truth values of one self-made sentence in several self-made or given worlds; (c) production problems: problems in which the students had to produce a world in which given sentences were true or false or in which the students had to produce sentences that were true or false in a given, fixed world; and (d) modification problems: problems in which the students had to modify (explicitly or implicitly) a given world, in such a way that given sentences would become true or false.


The following statement S1 is given:

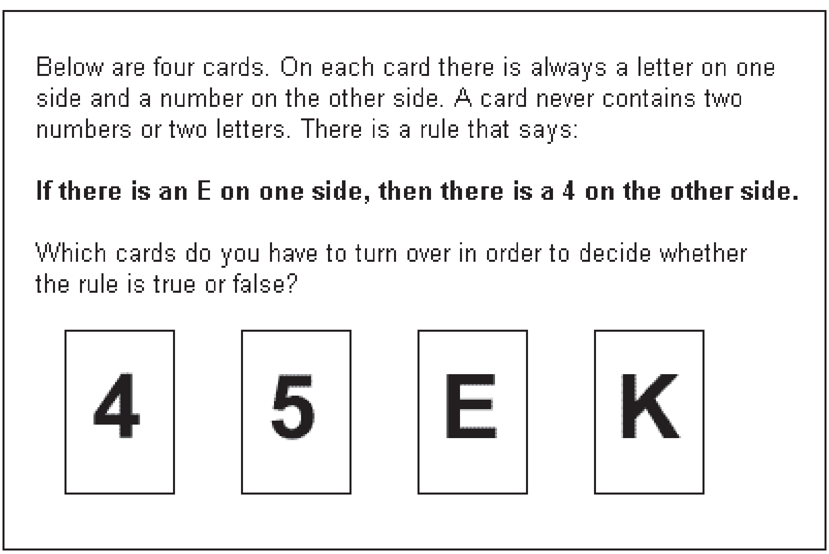
S1: If I go to the city today, I will eat an ice cream.

I am going to the beach today and I am eating an ice-cream.

Is the above given statement S1 true or not?

- yes
- no
- that depends

Figure 5-3. ***A typical example of the items given in the second part of the transfer tests measuring knowledge about the conditional in an everyday life setting. The correct answer is "yes".***


Below are four cards. On each card there is always a letter on one side and a number on the other side. A card never contains two numbers or two letters. There is a rule that says:

If there is an E on one side, then there is a 4 on the other side.

Which cards do you have to turn over in order to decide whether the rule is true or false?

Figure 5-4. ***A typical example of the items given in the third part of the transfer tests measuring knowledge about the conditional with a Wason selection task (Wason, 1966). The correct answer is to turn over the cards with the E and the 5 written on them.***

Furthermore, two questionnaires were administered. The first concerned the former education of the students on mathematics, programming and logic. The second questionnaire was an evaluation of the instruction in combination with Tarski's World.
All tests, problems to be solved and questionnaires were administered on the computer.

5.2.5 Log files

Two log files were generated during the experiment. The first log file logged all the actions of the students while working in Tarski's World. It logged the status of the sentences and the matching world every time the students checked this combination on well-formed formula (WFF), sentence and/or truth. The second log file logged all the actions of the students while working in the browser. This, among others, concerned the answers of students on two questionnaires, the answers of students on the transfer tests, and the time registration.
From these log files, the mean number of checks per student could be computed. As the subject matter concerned the conditional with the general format $p \rightarrow q$, these checks could be divided into the four possible truth-falsity combinations ($1 \rightarrow 1$, $1 \rightarrow 0$, $0 \rightarrow 1$, $0 \rightarrow 0$, in which 1 = true, and 0 = false).

5.2.6 Design and procedure

Experimental conditions

A full-factorial design of four conditions was administered (see Appendix C): (a) a manipulation / guidance condition Man+Guid+; (b) a manipulation / non-guidance condition Man+Guid-; (c) a non-manipulation / guidance condition Man-Guid+; and (d) a non-manipulation / non-guidance condition Man-Guid-. The subjects were randomly assigned to one of the four conditions. In all conditions, the students were given Dutch sentences, which they had to translate in first order predicate language. In the Man+Guid+ condition, the students had to construct graphical worlds in which they had to check the sentences to be true or false. The problems presented led them through the subject matter to all different types of basic problem situations that can be derived from the conditional, especially to those problem situations in which the students were supposed to develop new knowledge. Moreover, the students were specifically directed towards manipulating certain objects in order to see what effect this change had on the truth values of the sentences. In the Man+Guid- condition, the students had to construct graphical worlds in which they had to check whether the sentences were true or false. The students were given problems that instructed them to manipulate objects in the world to see what happened with the truth values of the sentences. By doing this the students were free to explore and manipulate whatever they wanted without guiding them in this process. In the Man-Guid+ condition, the students had to check whether the sentences were true or false in a given, fixed graphical world. Problems presented led them through the subject matter to all different types of basic problem situations that can be derived from the conditional, especially to those problem situations in which the students were supposed to develop new knowledge. Moreover, the students were specifically directed towards imagining what would happen with the truth values of the sentences, if they changed the world in a certain way. The students were also given feedback on this. In the Man-Guid- condition, the students had to check whether the sentences were true or false in a given, fixed graphical world. The students were given problems in which they were instructed to construct sentences that could be checked in the given world. By giving this possibility the students were free to explore whatever they wanted without being guided in this process and without the possibility to manipulate objects.

Procedure

The experiment was held in three consecutive sessions; the first (pre test, instruction and general training) and the second (exercises and post test) on two successive days and the third (retention test) two weeks later.
The first session started with an introduction after which the students had to complete a questionnaire about their previous education in mathematics and logic. This questionnaire was followed by a pre test consisting of three parts with a total of 27 items measuring the knowledge of the subjects of several aspects of the conditional (see Appendix D). Successively, the subjects received a verbal instruction of 30 minutes in which they had an introductory course into first-order logic, as used in Tarski's World. This instruction gave the subjects an idea of what logic can be used for, what Tarski's World can do, what logic operators and quantifiers are available, how these operators and quantifiers can be used and what truth and falsity means. This, together with some examples was the knowledge the students were equipped with. After the instruction, the students received a general training of about two hours depending on the condition to which they were assigned. During the training, the students had to complete eleven assignments in which they learnt to work with Tarski's World and with the logic operators. For this, model progression was used, an idea introduced by White and Frederiksen (1990). One of the general principles of model progression is to structure the rich information source and to keep the environment manageable by not introducing too many ideas at one time. Model progression entails starting with a simplified version of a model and gradually offering more complex versions of the model. In this case, the model was the field of predicate logic. The concepts were introduced in the following order: (a) predicates and constants, (b) connectives and parentheses, (c) quantifiers and

variables, and (d) conditional. If the students had any questions, assistance was given by one of the experimenters present.

The second session started with twelve exercises that had to be solved by the students in the four conditions. The first two problems were presented to refresh the knowledge acquired the day before. Consequently, ten exercises addressed the conditional. In all the exercises students were asked to translate Dutch sentences into first-order logic and to check the truth of the sentences in the (given or self-constructed) world of geometrical objects. After the students had completed the exercises, a post test was administered. In this post test, the students were again tested on their knowledge of the conditional. The post test consisted of the same type of items as used in the pre test. The second questionnaire in which the instruction and Tarski's World was evaluated, was also administered during this second session. Two weeks after the experiment, the students had to return for the retention test. This test consisted of comparable items as were used in the pre- and post test.

5.3 Results

5.3.1 Reliability

The reliability of pre-, post- and retention test, as measured with Cronbach's α, was $\alpha = .43$; $\alpha = .74$; and $\alpha = .78$ respectively. Deleting items from the test did not lead to significant higher reliabilities.

5.3.2 Pre-, post- and retention tests

Table 5-1 shows the means and standard deviations on the pre-, post- and retention test for the four conditions Man+Guid+, Man+Guid-, Man-Guid+ and Man-Guid-. The maximum score was 27. A significant mean effect was found from pre- to post- to retention test ($F(2, 67) = 21.71$, $p < .001$) and an interaction was found between the three tests and the four conditions ($F(6, 136) = 5.02$, $p < .001$). The scores of the students in condition Man+Guid+ increased significantly from pre- to post test ($F(1, 17) = 18.35$, $p < .001$) and from post- to retention test ($F(1, 17) = 9.03$, $p < .01$). The scores of the students in condition Man-Guid+ increased significantly from pre- to post test ($F(1, 17) = 19.17$, $p < .001$), but decreased significantly from post- to retention test ($F(1, 17) = 5.28$, $p < .05$). Scores on pre-, post- and retention tests did not increase significantly for condition Man+Guid- ($F(2, 34) = 1.54$, $p > .05$) and condition Man-Guid- ($F(2, 34) = 1.01$, $p > .05$).

Table 5-1. ***Means and standard deviations for each condition on pre-, post- and retention tests***

	Condition			
Test	Man+Guid+	Man+Guid-	Man-Guid+	Man-Guid-
Pre test				
M	13.22	14.17	14.11	14.50
SD	2.18	3.37	2.14	1.98
Post test				
M	17.50	14.49	18.22	15.50
SD	3.90	4.17	4.15	3.17
Retention test				
M	19.22	15.33	17.33	15.11
SD	4.07	4.43	4.21	3.31

Note. Maximum score = 27. Man+Guid+ = manipulation, guidance condition; Man+Guid- = manipulation, non-guidance condition; Man-Guid+ = non-manipulation, guidance condition; Man-Guid- = non-manipulation, non-guidance condition.

The scores of the students were divided into the scores on the three different parts of the transfer tests (Tarski's World items, everyday life items and Wason tasks, see Figure 5-5). In part 1, the Tarski's World items, a significant mean effect was found from pre- to post- to retention test ($F(2, 67) = 13.29$, $p < .001$) and an interaction was found between the three tests and the four conditions ($F(6, 136) = 2.91$, $p < .05$). The scores on pre- to post test increased significantly for the two guidance conditions, Man+Guid+ ($F(1, 17) = 13.73$, $p < .01$) and Man-Guid+ ($F(1, 17) = 13.91$, $p < .01$). All other differences were non-significant.

In part 2, the everyday life items, again a significant mean effect was found from pre- to post- to retention test ($F(2, 67) = 13.38$, $p < .001$) and an interaction was found between the three tests and the four conditions ($F(6, 136) = 4.34$, $p < .001$). Scores on pre- to post test increased significantly for the two guidance conditions, Man+Guid+ ($F(1, 17) = 11.98$, $p < .01$) and Man-Guid+ ($F(1, 17) = 18.98$, $p < .001$). Scores on post- to retention test again increased for condition Man+Guid+ ($F(1, 17) = 8.12$, $p < .05$), but decreased for condition Man-Guid+ ($F(1, 17) = 5.39$, $p < .05$). The scores in the two non-guidance conditions, Man+Guid- and Man-Guid-, did not differ from pre- to post- to retention test.

In part 3, the Wason tasks, a significant mean effect was found from pre- to post- to retention test ($F(2, 67) = 3.60$, $p < .05$) and an interaction was found between the three tests and the four conditions ($F(6, 136) = 2.19$, $p < .05$) caused by a significant increase of scores from pre- to post- to retention test for condition Man+Guid+. No other differences were found for the other conditions between scores on pre- to post- to retention tests.

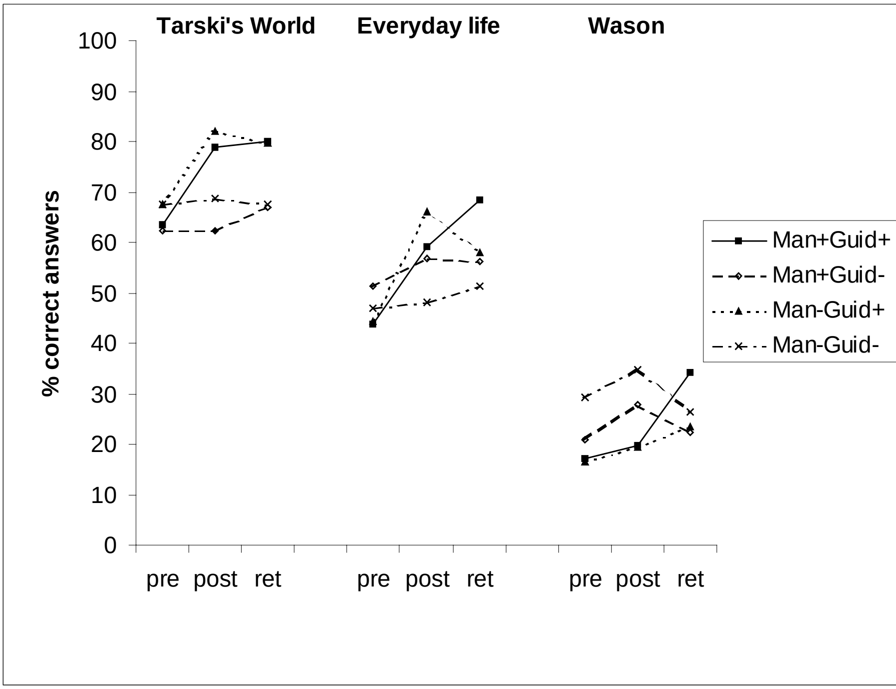


Figure 5-5. ***The scores of the students on the pre-, post and retention test of the Tarski's World items, the everyday-life items and the Wason items as a function of the conditions of the students.***

5.3.3 Process data

In Table 5-2, a summary is given of the mean number of checks of students working on the experimental exercises in Tarski's World. As can be seen, students in condition Man+Guid+ used significantly more checks on all sentences than students in the condition Man+Guid- ($F(1, 34) = 12.07$, $p < .001$), students in the condition Man-Guid+ ($F(1, 34) = 14.67$, $p < .001$) and students in the condition Man-Guid- ($F(1, 34) = 93.50$, $p < .001$). Students in condition Man-Guid- used significantly fewer checks on all sentences than students in the condition Man+Guid- ($F(1, 34) = 38.13$, $p < .001$) and students in the condition Man-Guid+ ($F(1, 34) = 97.57$, $p < .001$).
Furthermore, students in the non-guidance conditions, Man+Guid- and Man-Guid-, checked most in the situations to which they already knew the answer ($1 \rightarrow 1$ and $1 \rightarrow 0$). In unfamiliar situations ($0 \rightarrow 1$ and $0 \rightarrow 0$), they checked significantly less (a ratio of 3.0 to 1 and 4.8 to 1 respectively). Students in the guidance conditions, Man+Guid+ and Man-Guid+, spread their checks equally over all possible situations (a ratio of 1.3 to 1 and 1.0 to 1 respectively).

Table 5-2. ***Summary of process data of students working in Tarski's World***

	Condition			
	Man+Guid+	Man+Guid-	Man-Guid-	Man-Guid-
# checks				
all sentences[a]	73.7	55.4	56.9	31.4
$1 \rightarrow 1$[b]	25.3	24.7	18.8	14.4
$1 \rightarrow 0$[b]	16.7	17.6	10.2	11.6
$0 \rightarrow 1$[b]	13.2	4.7	10.2	2.3
$0 \rightarrow 0$[b]	13.7	7.4	10.9	2.6
0 (na) $\rightarrow x$[b]	4.7	1.1	6.8	0.5

Note. Man+Guid+ = manipulation, guidance condition; Man+Guid- = manipulation, non-guidance condition; Man-Guid+ = non-manipulation, guidance condition; Man-Guid- = non-manipulation, non-guidance condition.

[a] *mean number of checks over all sentences during the experimental exercises.*

[b] *mean number of checks per student made on five possible checks of the conditional ($p \rightarrow q$) during the experimental exercises (1= true; 0 = false; 0 (na) = empty set; x = unknown truth value).*

5.4 Discussion

5.4.1 Reliability

The fact that the reliability of the pre test is lower than the reliabilities of the post- and retention test is explained by the small number of correct answers on the pre test ($M = 14.07$, $SD = 2.52$). Apparently, the pre test was difficult, so that the students may have been guessing when answering the items, which has a negative influence on the reliability of the test.

5.4.2 Overall learning results

The overall scores on the pre-, post- and retention tests clearly show that students who received problems in which they were directed to situations in which new knowledge could be developed (i.e., students in the conditions Man+Guid+ and Man-Guid+), profited from this instructional support. These students were able to solve significantly more logic problems correctly on the post- and retention tests in comparison to the pre test. Students having the possibility to manipulate even increased their scores from post- to retention test, whereas the scores of the students having no possibilities to manipulate decreased again after two weeks, but were still higher than the scores on the retention test of the students who were not guided. The latter students (i.e., students in the conditions Man+Guid- and Man-Guid-) did not develop new knowledge concerning the conditional, whether they had the possibility to manipulate or not. Having the possibility to manipulate is shown to be only effective when combined with guidance. The process data confirm the hypothesis that this is caused by the students not using the possibilities of manipulation when they are given freedom to explore whatever they want. Only guidance will induce students to explore all possible situations, so they also experience what happens with the truth values of the sentences in situations they do not know the answer to or they are uncertain about. The results show that guidance in combination with the possibility to manipulate objects in a domain supports the development of new knowledge and skills in that domain. Guidance leads to the students seeing all possible situations including the ones in which new knowledge can be developed, which has positive effects directly after instruction. The possibility to manipulate leads to the students playing around in such a way that they (systematically) experience the results of their actions by the feedback given on the same problems in different worlds. This 'seeing with your own eyes' and (physically) experiencing what happens with the truth values of the logical expressions when the world is changed is assumed to lead to making inferences, to lead to a better memorisation of the subject matter and to a continuing reflection on the logic problems in the interval between administering the post- and retention test. The students needed time for integrating their newly developed knowledge with their existing knowledge, as the positive effect of this variable appears after two weeks of non-instruction.

5.4.3 Tarski's World puzzles: near-transfer

The scores of the students on the items in the first part of the transfer tests, the Tarski's World puzzles, show that near-transfer is possible if (and only if) the students are given guidance, independently of the possibility to manipulate. The scores of the students who were guided (i.e., scores of the students in the conditions Man+Guid+ and Man-Guid+) increased significantly from pre- to post test, showing that, in near-transfer tasks, students directly profit from this form of guided instruction. In addition, the scores of the students who were guided, remained at the same level from post- to retention test, showing that, in near-transfer tasks, the subject matter learnt is well integrated into the students' knowledge base to correctly solve near-transfer tasks after two weeks of non-instruction, but that the subject matter is not reflected upon anymore after instruction. A possible cause for this finding is that, during the near-transfer test, the students recognise the subject matter from the material presented during instruction, but after instruction they are not confronted anymore with similar situations. Furthermore, the scores of the students in the non-manipulation conditions did not differ from the scores in the comparable manipulation conditions. Apparently, manipulation did not influence the scores of the students on the near-transfer tests.

5.4.4 Everyday-life puzzles: medium-transfer

The scores of the students on the items in the second part of the transfer tests, the everyday-life puzzles, show that medium-transfer is possible if (and only if) the students are given guidance and they show that this transfer continues after an instruction-free period when combined with manipulation. The increase of the scores from pre- to post test of the students who were guided (i.e., students in the conditions Man+Guid+ and Man-Guid+) show that students also directly profit from this form of guided instruction in medium-transfer tasks. However, from post- to retention test this guided instruction interacts with the possibility to manipulate. Students having this possibility continue reflecting upon the subject matter in the instruction-free period. They start to apply their newly developed knowledge in everyday life situations, which results in increasing scores from post- to retention test on the everyday-life puzzles in the experiment. The scores of students who were guided but who did not have the possibility to manipulate decreased from post- to retention test. The first possible explanation for this is that the students did not reflect upon the subject matter in the instruction-free period and that they did not completely retain all knowledge acquired, because they did not have the possibility to manipulate and thus could not (physically) experience their actions. The second explanation is that the students did reflect upon the subject matter in the instruction-free period, but that they did not fully retain the newly learnt knowledge caused by the lack of possibilities to manipulate. This may then have led to confusion resulting from an interplay between the newly acquired knowledge and the existing everyday life knowledge coming forward again in everyday life situations. As the authors cannot explain why the possibility to manipulate would lead to reflection and the lack of the possibility to manipulate would not lead to reflection, they assume that the second explanation is the most probable one.

5.4.5 Wason tasks: far-transfer

The scores of the students on the items in the third part of the transfer tests, the Wason tasks, show that far-transfer is still difficult, even after instruction in which the students are guided in their problem solving process. The scores of the students in all conditions are low, although there is a trend from post- to retention test of increasing scores of the students in the Man+Guid+ condition. It is possible that students in this condition continued reflecting on these puzzles, but that integration of their newly developed knowledge was not yet completely established for far-transfer tasks.

5.4.6 Main findings and conclusion

Students who were not guided during instruction were not able to control their problem solving process themselves. They stayed in familiar situations and did not confront themselves with situations they (thought they) did not know the answer to or they were uncertain about. The result was that the students did not develop new, or change existing knowledge concerning the conditional.
Students who were guided through the subject matter were confronted with all different types of basic problem situations that could be derived from the conditional, especially those situations that were supposed to lead to the development of new knowledge. The students could infer knowledge about the conditional resulting in confirming existing everyday life knowledge that was correct and learning and developing new knowledge. Apparently, the students were able to use this knowledge directly after instruction when solving near-transfer and middle-transfer problems. Far-transfer problems still yielded difficulties. It is assumed that, in the two-week instruction-free period, the students tried to cognitively elaborate on their newly developed knowledge resulting in different outcomes. First of all, the students were most probably not confronted anymore with near-transfer problems. Therefore, they did not reflect on these problems, but the knowledge resulting from these problems became integrated well enough to yield similar scores on the retention test compared to the post test. Secondly, the students were still confronted with the middle-transfer problems in their everyday life. Students who, in addition to being guided, had the possibility to manipulate during instruction retained their newly learnt knowledge correctly and continued to reflect on this knowledge. Students who did not have this possibility might not have retained the newly learnt knowledge correctly, leading to confusion between this newly acquired knowledge and the incorrect everyday life knowledge, resulting in lower scores on the retention test. Thirdly, students were still not able to solve the far-transfer problems although it seems that the students in the guided, manipulation condition also reflected on these problems leading to a small increase of scores on the Wason tasks.

The findings of this study give reasons to reflect on the design of learning environments. Questions arise, such as, how 'open' should the problem solving situations be? To what extent can learners be left on their own? To what extent do they need at least some kind of instructional support or guidance? Can learners be held responsible at all for their own development of knowledge and skills? Is the extent of required guidance age-bound? What influence do these open learning environments have on the quality and quantity of the subject matter learnt? And what about the effects in the long run? All these (and other) questions are important for the design of learning environments and for the design of instruction. The results of the present study can be used as a start in an attempt to answer some of them, of course with the restriction in mind that no social interaction was possible during the experiment. This study shows that giving learners complete freedom to explore in an open problem solving environment does not easily lead to new knowledge development. In such open situations, learners continue to use their (often incorrect) existing knowledge and they do not get any feedback on the way they approach the subject matter. Consequently, they do not confront themselves with all different types of basic problem situations and thus do not receive any feedback on these critical situations in which new knowledge could be developed. Therefore, learners need to be guided in their problem solving processes. In addition, guidance should stimulate learners to operate on objects so that the correctness of their existing knowledge is (physically) experienced in the world given. This results in making inferences about what was correct and what was incorrect leading to changes of incorrect knowledge or development of new knowledge and integrating this with correct existing knowledge, leading to better long-term effects.

6

Overview and reflection

Abstract

An overview and reflection is given for the experiments described in Chapter 3 up to Chapter 5. First, the instructional design is discussed. Then, the instructional variables which were used, a concrete context, visualisation, manipulation and guidance, are investigated followed by a discussion about the design, the methodology, the subjects and the transfer tests. These all changed from experiment to experiment. The second part concerns implications which can be drawn from the experiments for further research as well as for educational practice.

6.1 Instructional design

In Chapter 1, instruction was defined as a communicative activity which is intended to initiate and support the development of knowledge and skills. Instruction is part of a communication between an expert and a novice and is conveyed through a medium. In instructional design, three major questions are asked: (a) what are the goals of the instruction; (b) what should the instructional communication look like; and (c) how will the learning results be measured. The answer to the first question leads to the description of knowledge, skills and attitudes which concern the goals of the instruction and which should be developed or learnt. A learning process should lead to knowledge development, which can be used and transferred to other situations and which will be remembered for a long time. In the research described in the preceding chapters, the aim was the development of logic concepts, and in particular, learning the truth values of the conditional in first-order predicate logic.

The main issue dealt with in this thesis was the design of the instructional communication. In order to address this issue several decisions have to be made. First, the decision must be made whether the reality (i.e., objects and situations in a domain) will be used for instruction or whether a medium will be used to depict or describe the reality. If the reality itself is used, it is called macroscopic and perceptible with the senses. If a medium is used, the function of the medium should be considered. A medium can represent the object or situation, it can depict a change in the situation or it can represent a process, often in an iconic way. Furthermore, the type of representation (e.g., language, pictures, sounds or combinations of these) should be chosen, as well as the extent to which the medium resembles the reality (e.g., homomorphism or isomorphism) and the way of depiction (e.g., iconic by means of natural mapping principles or symbolic by means of arbitrary mapping principles). Moreover, the number of representations should be chosen. Sometimes a single representation suffices, whereas in other cases multiple representations provide more effective instructions (e.g., by providing more than one perspective of the problem situation). In the three experiments which were reported in this thesis, a medium was chosen to represent the reality. For this purpose, language representations as well as visual representations were chosen. If a representation or depiction of the reality is chosen, it should be decided which type of medium in the technical sense is most effective. This decision depends on the type of problems the learners have to solve, which depends on the difficulties learners encounter in connecting the new knowledge to their existing knowledge. In the three experiments that were performed, all these decisions led to a computer with a browser and a microworld as the medium for the instructional communication.

The last question of instructional design concerns the type of assessment and the performance criteria. The purpose of assessment is to determine to which extent learners have achieved the learning goals. In this thesis, transfer tests were used in which the knowledge of the students on the conditional was measured in near-transfer situations, medium-transfer situations and far-transfer situations.

6.2 Reflections on experiments and results

The main objective of this thesis was to examine whether instructional variables, often used in learning empirical sciences, can also be used to facilitate the development of knowledge and skills in the formal sciences, particularly in learning logic. In this section, this research question will be answered by addressing the instructional variables used. Furthermore, the methodology, the experimental design, the transfer tests and the subjects, issues which changed from experiment to experiment, will be reflected upon.

6.2.1 The instructional variables

Concrete context

In learning the content and problem solving methods of the empirical sciences, it is assumed that embedding abstract rules and concepts in a concrete context facilitates learning (e.g., White, 1993). From these concrete objects or situations, abstract rules can be induced and predictions can be made (Piaget, 1973). It was supposed that embedding the abstract rules and concepts of logic into a concrete context would facilitate learning logic, especially because different studies have shown that learners assign their difficulties in learning logic to the abstractness and formality of the rules of logic (Fung, O'Shea, Goldson, Reeves & Bornat, 1994; Goldson, Reeves & Bornat, 1993). Successively, the analysis of representations used in everyday life reasoning and formal reasoning as described in Chapter 1, led to the idea to use a well-defined concrete context in which subtle differences in the meaning of the connectives, ambiguity of words, internal relations within sentences, the context of expressions and user conditional aspects could not interfere with the subject matter. In the experiments, a well-defined context of geometrical objects was used (Barwise & Etchemendy, 1992). Although the effect of this instructional variable was not studied in an experimental setting, learners stated that the concrete context made it easier to work with formulae.

Visualisation

Instructions to learn the content of the empirical sciences often use visualisation techniques. Paivio (1986) and Clark and Paivio (1991) stated that memory for pictures is better than memory for texts, because texts are processed and encoded in the verbal systems, whereas pictures are processed both in the image and in the verbal systems. Other theories of cognition state that graphical representations are more effective in communicating material, which explains their advantage (e.g., Koedinger & Anderson, 1994; Kulpa, 1995; Larkin & Simon, 1987; Stenning & Oberlander, 1995). These advantages can also contribute to learning logic. Visualisations give the possibility to visually check the reasoning process, to retain the reasoning steps made and to receive feedback from the situation. In the second experiment, this instructional variable, visualisation, was studied. The results of this experiment showed that no differences were found between learners who were presented with visual depictions of the situations and learners who were presented with verbal descriptions of the situations. It is possible that visualisation did not facilitate the development of knowledge and skills in learning logic. However, the process data showed that the majority of the learners who were presented with verbal descriptions drew pictures themselves. As this resulted in visual support of the verbal condition, this led to the conclusion that a possible effect of visualisation was cancelled out resulting in no differences between the visual and the verbal condition. However, the result that learners who were not presented with visualisations drew pictures themselves, showed that, apparently, these learners needed the visualisations to come to the same results as the learners who had these visualisations to their disposal. This strengthened the idea that visualisations facilitate learning logic.

Manipulation

In learning the content of the empirical sciences, the possibility to manipulate objects is considered to be an important instructional variable. In studying physics, for example, this means using devices and doing experiments. When trying to solve problems, learners can induce general rules. By acting on objects, learners can see what effect their actions have. These interactions result in observations and experiences. Several experiences can lead to hypotheses and predictions and these can be tested. The activities lead to the development of new knowledge or the change of existing, incorrect knowledge. In the end, the subject matter is better understood and remembered. As learning logic involves learning to make predictions by reasoning whether sentences are true or false, manipulation of objects is assumed to also facilitate the development of knowledge and skills of logic. Piaget (1973) called this learning by logic-mathematical experience. This hypothesis was tested in the second and the third experiment. In the second experiment, no differences were found in the knowledge development between learners who were given the possibility to manipulate objects and learners who were not given this possibility. However, the process data showed that learners having the possibility to manipulate did not use this possibility at all. Their working method did not differ from learners not having the possibility of manipulation, cancelling out possible effects.

Therefore, a third experiment was designed in which learners were compelled to manipulate objects by giving them instructions to do so. In this experiment, differences were found between learners who had the possibility to manipulate the objects and learners who did not have this possibility, but only in the long run. There were no immediate effects, but after two weeks, the learners who were stimulated to manipulate objects remembered more of the subject matter than learners who could not manipulate the objects. Therefore, manipulation of objects has shown to be a helpful component to include in instruction in order to facilitate learning logic.

Guidance

In the last experiment, the instructional variable, guidance, was studied as well. Guiding the learners in their learning process was operationalised by stimulating the learners to manipulate objects in such a way that they were confronted with problem situations that could not be solved with their existing knowledge. In such situations, existing knowledge has to be changed or new knowledge has to be developed. This instructional variable, guidance, which also appeared to be essential in empirical sciences (e.g., de Jong & van Joolingen, 1998), gave clear results. Learners not guided in their problem solving process did not learn, whereas learners guided in their problem solving process did learn, in the short as well as in the long run. This leads to the conclusion that including guidance to those problem situations that are critical for learning is an essential part of the instruction.

Conclusion

From the results of the experiments summarised above, the research question whether instructional variables originating from empirical sciences can also be used to learn logic, is answered positively. Learners stated to be positive about embedding abstract rules in a concrete context. Apart from the possible facilitating effect on learning, this also involves a motivational effect. Therefore, embedding the abstract rules of logic in a concrete context is a good instructional design decision. The concrete context chosen in the experiments described in this thesis involved a completely pre-defined context of geometrical objects. It is possible that worlds consisting of everyday life objects give the same results, but only when such a world is well defined. Only then can differences in truth and user conditional aspects that play a role in everyday life be controlled, so that the chance of interference with the learning process is small or even absent.
Furthermore, the experiments showed that learners who were presented with verbal descriptions of the problem situation apparently needed additional graphical representations to reach the same results as learners who were given these visualisations. In the experiments, worlds of objects were used in which the objects had spatial relations to each other (e.g., to the left of, behind, between). It is possible

that these spatial relations automatically evoked mental images, so that learners could not reason without these visualisations. However, it is assumed that other (concrete) relations than spatial relations would have evoked mental images as well. Therefore, in the design of instructions, visualisations (as a medium) should be included.

The advantage of manipulation of objects was shown in the last experiment. Learners who had the possibility to manipulate objects remembered the subject matter better after two weeks compared to learners who did not have this possibility to manipulate. This instructional variable is regarded as useful to teach logic, but only in combination with the stimulation to manipulate, as learners do not manipulate on their own accord. The learners should be directed to problem situations in which their knowledge is insufficient to solve the problems and from which they can develop new knowledge.

The advantage of guiding the students in their problem solving process was shown in the last experiment. The results clearly showed that learners not guided in their problem solving process did not learn, whereas learners guided in their problem solving process did learn. From this experiment, the conclusion is drawn that guidance to the situation in which knowledge can be changed is an important instructional variable when the learners are expected to generalise the rules of reasoning from specific cases in a computer-based learning environment. Guidance can be conceptualised, though, in more than one way. In this study, problems consisted of assignments which directed the students to certain situations critical for knowledge development. Furthermore, the environment and instruction were structured by giving the learners problems instead of letting them explore the environment themselves. These problems were ordered from simple to complex and new concepts were introduced one by one, that is model progression was used (White & Frederiksen, 1990). Other ways of regulating the problem solving process of the learners, such as registering all experiments which the learner performed (de Jong & van Joolingen, 1998), can possibly lead to facilitating effects as well. In Tarski's World, this would mean registering all checks of combinations of sentences and worlds and their truth values, so that the learners have an overview of what they have been doing.

6.2.2 Experimental design

The experimental designs used in the three experiments differed from experiment to experiment. In the first experiment, three conditions were compared, one of which was the control condition. The remaining two conditions differed in the number of representations, so that the added value of the extra representation (the formal, symbolic expression) could be studied. In the second experiment, two variables, visualisation and manipulation, were studied. As manipulating objects cannot occur without a visualisation, this resulted in three conditions (instead of four). In the third experiment, two variables, manipulation and guidance, were compared in a full-factorial design. This resulted in four conditions so that both variables could be studied independently and in combination.

6.2.3 Methodology

In all experiments, a 'pre-post-retention test' design was used in which there was an instructional period from pre- to post test and in which there was a retention period from post to retention test. The instructional period consisted of (a) an introduction in which basic elements of first-order predicate logic were described and in which learners were shown how to use Tarski's World; (b) a general training depending on the conditions the learners were assigned to, in which they learnt to work with Tarski's World and the logical connectives; and (c) the experimental part depending on the conditions the learners were assigned to, in which learners had to solve problems in order to generalise the rules of logic from specific cases.

In the first experiment, the introduction was presented on screen. The learners had to read this and were given the possibility to reread it and, then, had to continue without having the possibility to go back. As there was no check on whether the students had indeed read this part and as this was important for the experiment, the introduction used in the second and third experiment was presented in a verbal, class form supported by overhead sheets. This introduction into the basic elements of logic took about 30 minutes and students were allowed to make notes on the slides and questions could be asked. Presenting the introduction in a verbal, class form was seen as an important improvement. The positive effect of the class introduction, which described the basic elements of first-order predicate logic, is shown by the result that students made fewer errors in the translation process after a class introduction than after an introduction on the computer.

Moreover, the whole experimental environment was changed from a hypercard environment during the first experiment to a browser environment during the second and third experiment. This gave the learners the possibility to go back and forth to other sections and problems within the general training and within the experimental part, whereas this was not possible in the hypercard environment. Yet, students were advised to go through the exercises in the order indicated, as the exercises increased in difficulty. The increase of freedom in the environment was seen as an improvement, as all students kept to the advice of the indicated order, but they used the possibility to see what they had done earlier and to compare problems and solutions, too.

Moreover, during the general training the learners in the second and third experiment were allowed to use the introductory slides, containing the most important issues dealt with during the class introduction. This measure was assumed to enable learners to better acquire the basic knowledge necessary for the experimental part.

Furthermore, the problems and the number of problems presented to the learners changed from experiment to experiment. During the experiments, it became clear that some problems were too simple whereas others were too complex. Besides this, the problems were adapted to the different goals and thus to the conditions of the experiments. Moreover, the conditions within an experiment had to be comparable. This, together with the number of problems presented, had an effect on the length of time of the instructional period.

6.2.4 Participants

The participants in the first experiment were first-year computer science students, as this experiment was a replication of the experiment of Van der Pal and Eysink (1999). These students took part in an obligatory course in which they were given an introduction into logic. As part of the course, the students could volunteer to participate in the experiment. They were not paid for this. In the second and third experiment, the participants were students who were in their first- or second-year of a social science study. These students were preferred to, as they were supposed to have more difficulties in learning logic than computer science students (Fung, O'Shea, Goldson, Reeves & Bornat, 1994; Goldson, Reeves & Bornat, 1993). They were paid for their participation and did not know that the subject matter concerned logic. When they were informed about this at the beginning of the experiment, some students expressed their discontent and reluctance. During and after the experiment, however, students stated that their view of logic had changed. They became very motivated to solve the problems and to finish the experiment with success. In the second experiment, 82% of all students stated that they enjoyed working with Tarski's World. In the third experiment, this number was 81%.

6.2.5 Transfer tests

In order to measure the development of knowledge after instruction, a 'pre-post-retention' test design was administered. These transfer tests were designed in order to assess the knowledge of the learners concerning the conditional. The tests evolved from experiment to experiment.

First experiment

In the first experiment, the transfer tests consisted of four Wason selection tasks. The reliabilities of these tests were high (α = .93; α = .96; and α = .97 respectively). This can be explained by the fact that all items were similar problems. Learners that were in the complete-insight state of the model of Johnson-Laird and Wason (1970) will consistently have answered all items correctly, whereas learners in the no-insight or partial-insight state will either have consistently given the same, but incorrect answer to all problems or they will have been guessing. Yet, these tests had a major disadvantage. The mean score on the tests immediately after instruction was still only 1.40, whereas the maximum score was 4. Although the Wason selection task measures the knowledge of the conditional, Wason selection tasks have proved to be difficult, and therefore, it is possible that students understood the conditional within the domain of Tarski's World, but that they were not able to transfer their newly developed knowledge to a complicated task like the Wason selection task. To solve this, the test was extended with items that could be more easily transferred to, resulting in a new transfer test for the second experiment.

Second experiment

The second experiment used this new test consisting of three Wason selection tasks, one Wason selection task in a Tarski's World setting, one Reduced Array Selection Task (i.e., a Wason selection task without a p-card), two items to be solved best by using sets and four logical deduction problems in an 'everyday life' setting. The reliabilities of these tests decreased (α = .49; α = .68; and α = .75 respectively), but they gave a better idea of the knowledge of the students concerning the conditional. Yet, there were some flaws in this test. Interpretation of the effect of the Reduced Array Selection Task appeared to be difficult. In this task, the same instructions are given as in a Wason selection task with the difference that the (easy) p-choice is omitted. Students are forced to give an answer, so they have to continue thinking about an answer even though they think the p-choice is the only correct answer. Therefore, it is possible that this task yields a learning effect by solving it. Moreover, the items solved best by using sets were not seen as suitable items to solely measure knowledge about the conditional. Furthermore, some near-transfer tasks within Tarski's World, similar to problems presented during the experimental part, were supposed to give a better view of the development of knowledge.

Third experiment

Taking the considerations of the transfer tests as used in the second experiment into account, this test was changed and extended for the third experiment. A complete version of the pre test is given in Appendix D. It consisted of three parts: (a) near-transfer items testing the knowledge of the students on the conditional within the domain of Tarski's World; (b) medium-transfer items testing the knowledge of the students on the conditional by logical deduction problems within an everyday life setting; and (c) far-transfer items testing the knowledge of the students on the conditional by means of Wason selection tasks. The reliability of the pre test was not high (α = .43), but this was explained by the fact that the test was difficult for students not having knowledge about the conditional. The students may have been guessing when answering the items and this influences the reliability of a test negatively. The reliabilities of the post- and retention tests, however, showed a reasonable internal consistency (α = .74; and α = .78 respectively). By dividing the transfer test into three parts, it became clear how well students were able to apply their newly developed knowledge in different situations and, thus, how well they understood this knowledge. The pattern found shows that students had difficulties in applying their newly developed knowledge to the Wason selection tasks, but that they were able to solve more items correctly on near- and medium transfer tasks. This shows the importance of including near- and medium transfer items in a knowledge test.

6.3 General implications

The overall conclusion, which can be drawn from the results gained from the three experiments, is that instructional variables often used in learning empirical sciences can also be used to develop knowledge and skills in first-order predicate logic. In the following part, some possible implications of these findings for further research and for the educational field will be discussed. Suggestions for further research concern the transfer tests and the manipulation variable. Implications for logic teaching concern solving the difficulties learners encounter when learning logic, the importance of learning the semantics of logic, embedding a learning environment in a social context, choosing an instructional method, choosing appropriate problem situations, and being aware of the consequences of open learning environments.

6.3.1 Implications for further research

Elaboration of transfer

One of the main general assumptions behind the current studies was to facilitate learning logic by relating the abstract concepts of logic to concrete situations. By doing this, the learners were supposed to be able to induce general (abstract) rules from concrete objects and situations leading to a better understanding of the rules. In the experiments, the learners were tested on their ability to transfer their newly developed knowledge. Although the transfer tests differed in the distance between the learning situation and the testing situation, the situations in which the knowledge was tested were all concrete situations: concrete objects in Tarski's World, concrete situations in everyday life and concrete (except for one) Wason selection tasks (see Appendix D). Learners were tested whether they understood the rules of logic and whether they could apply these rules in *concrete* situations. The learners were not tested whether they were able to use these rules in *abstract* situations. Examples of such contexts can be situations in which the learners are asked to construct the truth tables, or situations in which the learners must tell whether a certain (abstract) conclusion (e.g., p) can logically be deduced from some given abstract premises (e.g., $p \rightarrow \neg q$ and q). It is possible that the learners were still not able to transfer their knowledge to abstract situations and that they only had an implicit idea about the rules of logic without being able to explicate these in symbolic formulae. Therefore, further research can focus on studying the extent to which learners can apply the rules in abstract contexts, and the extent to which they can make these rules explicit.

Elaboration of manipulation

The last experiment showed the advantageous effects of manipulation of objects for the development of knowledge. This manipulation was operationalised as manipulation of geometrical objects in a well-defined world. By manipulating these objects, the learner could see what effects certain actions had on the truth values of the sentences. However, manipulation can also be conceptualised as manipulation of sentences. Again the learner can see what effects certain actions have on the truth values of the sentences in a certain world. But does this difference between manipulation of objects and manipulation of sentences lead to the same type and amount of knowledge development? Manipulating objects shows that the same sentences are true or false in more than one world. Manipulating sentences, however, shows that within one world more than one sentence can be true or false. In both cases, the learners perceive the relation between a logical expression and a world. In both cases, the learners can be confronted with counter-intuitive situations in which incorrect, existing knowledge must be changed or new knowledge has to be developed. It is hypothesised, however, that manipulation of objects in a world would lead to better learning results than manipulation of logical expressions. Humans are used to manipulating objects in everyday life, whereas manipulation of symbols with which they are not familiar is a new and complex activity. It would be interesting to investigate whether this hypothesis is correct, whether these two types of manipulation lead to the same or different results and whether this has any influence on the type and amount of knowledge development.

6.3.2 Implications for logic teaching

Resolving the difficulties of logic

In consequence of the analysis of differences between representations used in everyday life reasoning and representations used in formal reasoning, two difficulties could be distinguished in learning logic: (a) the formality and abstractness of logic, and (b) the differences between logic in an everyday life context and logic in a formal context and the interference of the first when learning the latter. This thesis has shown that resolving these difficulties is important in order to facilitate learning logic, and suggestions have been put forward as to how this can be realised. In the current work, these difficulties have been resolved by embedding abstract rules in a completely pre-defined context of concrete objects. Moreover, different components of instruction were designed, such as visualisation of these objects, manipulation of these objects and guiding the learners through their problem solving process. The instructional components that were designed and developed at certain moments of the research project have shown to be effective.

Learning the semantics of logic

One conclusion of this study is that learning logic should be more than learning how to manipulate formulae. Learners should learn the semantics of logic. In other words, a connection should be made between symbolic expressions and reality, so that the learners experience the notion of truth. Learning these semantics can be promoted by the instructional variables used in the experiments described in this thesis. These instructional variables were especially chosen with the difficulties in mind which emerged from the analysis of differences between everyday life reasoning and formal reasoning. The experiments have shown that the chosen instructional variables were effective.

However, teaching has to be adapted to the learners. The present research aimed at facilitating learning logic for learners who were supposed to have difficulties in learning logic. The components of instruction have shown to be effective for this type of learners. Yet, it is possible that learners who do not have any difficulties in learning logic perceive such an instruction as redundant and not necessary. It is possible that they do not need these instructional variants to learn the semantics of logic and that they are able to understand the rules of logic from abstract concepts. Thus, learning logic should include learning the semantics of logic, but this should depend on the learner characteristics.

Embedding an environment in a social context and choosing an instructional method

When deciding to make use of a computer-based learning environment for logic teaching, this environment should be embedded into a social context. Using a learning environment as a research environment differs from using the same environment for learning. When learning is the goal, social interaction comes into play. Learners can ask questions and clear goals including performance criteria are set. Furthermore, the question should be asked how the environment will be used and what the goal of the environment is. Will it be used in combination with a book or does it stand on its own? Will the environment be used as a means to give examples and let the learners practise with the subject matter learnt or will it be used to let the learners induce the general prescriptions of valid reasoning themselves? These issues coincide with the instructional method that will be used. When the choice is made for deductive learning, a book and/or lectures can teach the learners the general rules, which can then be applied to problems given in the learning environment to practise with these rules. Domains in which the rules to be learnt are very complex are supposed to be suitable for deductive learning, as inducing these complex rules will be too difficult for learners without any prior knowledge. An example is learning a new language. First, some words to communicate about and various rules of grammar should be given. This knowledge can then be applied in a learning environment in which the learners can practise. When the choice is made for inductive learning, the learning environment can be used on its own to let the learners induce general rules from specific cases. Inductive learning is suitable for domains in which the learner already has enough prior knowledge to be able to induce general rules of grammar. An example is learning communication skills in different situations. The learners already have an idea about the best way to communicate in different situations and the environment can provide them with situations in which the learners experience that their strategy is not the best one. In this case, the learning environment is used in such a way that learners have to induce general rules from specific cases.

Guiding learners by choosing appropriate problem situations

The data of the experiments show that learners should be guided in their learning process, in such a way that they can be directed to those situations that are critical or conditional for knowledge development. By first giving the students formal but concrete problems (i.e., a well-defined situation such as a situation consisting of geometrical objects, in which user conditional aspects cannot play a role), which can be solved correctly by their existing, everyday life knowledge, this existing knowledge is confirmed. Giving them then some formal but concrete problems which are counter-intuitive (i.e., which cannot be solved by their existing everyday life knowledge) will cause at least some confusion and doubt. Some of the students will ignore the problem, some of them will try the same answer in the same situation again and some will see what happens in another, but similar situation. But eventually, after several times of the same counter-intuitive feedback, they will extend their existing knowledge with the notion that, apparently in logic, the concerning situation evokes a different answer than their everyday life knowledge would suggest. Finally, giving the same counter-intuitive situation in an everyday life setting should lead to the realisation that, apparently, the same counter-intuitive answer does not only hold good for formal logic but also for everyday life situations, which eventually must lead to a change of their existing, everyday life knowledge.

Consequences of open learning environments

The results of the experiments showed that giving learners the possibility to manipulate and thus giving them the possibility to explore the environment on their own accord did not work. Learners stayed in familiar situations to which they already knew the answer. They did what they were asked to do, but nothing more than that. Possibilities to explore the environment were not used at all. In other words, learners needed guidance to direct them to non-familiar and counter-intuitive situations which they did not explore on their own. Manipulation was thought to be a good instructional variable, but the fact that learners did not manipulate on their own accord was not taken into account. Therefore, teachers should be aware of the fact that learners do not always behave in the way they are supposed to. Apparently, learners need some kind of support in open learning environments. This, however, does not mean that complete answers should be given. There must be a balance between the presentation of structured information and freedom (e.g., de Jong & van Joolingen, 1998; van Merriënboer & Krammer, 1990). Although this issue was not the aim of the present research, it gives rise to the question to what extent learners can be left on their own and to what extent they need at least some kind of instructional support or guidance. This immediately evokes the question whether learners can be held responsible for their own development of knowledge and skills or not. If they do not know how they should behave in an open learning environment and if they are not told how they can improve their learning process, this does not seem to be a fair thing to do. In the domain of interactive graphical representation systems, Reimann (1999) stated that students should first acquire specific strategic skills in graph processing, before any significant learning effects can be expected. In the domain of simulations, De Jong and Njoo (1992) acknowledged the importance of discovery learning skills, which they divided into transformative processes such as analysis, hypothesis generation, testing, interpretation and evaluation, and regulative processes such as planning, verifying and monitoring. Taking all the types of processes necessary in open learning environments into account, the learners should possess several skills in order to learn and to make a successful learning process possible. The question is when these skills develop during life, whether they are developed at a certain age and whether the extent of the required support is age-bound. A question at least as important as the amount of support is the influence of open learning environments on the quality and quantity of the subject matter learnt and the effects in the long run. A learning process should lead to knowledge development, which can be used and transferred to other situations and which will be remembered for a long time. The present study has shown that, with a careful analysis of the instructional communication including difficulties that learners encounter, effective instruction can be designed and developed.

References

Barwise, J., & Etchemendy, J. (1992). *The language of first-order logic: Including the Microsoft Windows program Tarski's World 4.0 for use with IBM-compatible computers.* Stanford, CA: CSLI.

Barwise, J., & Etchemendy, J. (1994). *Hyperproof.* CSLI Lecture Notes. Chicago: Chicago University Press.

Barwise, J., & Etchemendy, J. (1998). Computers, visualization, and the nature of reasoning. In T. W. Bynum & J. H. Moor (Eds.), *The digital phoenix: How computers are changing philosophy* (pp. 93-116). London: Blackwell.

Begg, I. (1987). Some. *Canadian Journal of Psychology, 41*(1), 62-73.

Begg, I., & Harris, G. (1982). On the interpretation of syllogisms. *Journal of Verbal Learning and Verbal Behavior, 21,* 595-620.

Bender, J. (1987, February). *Teaching introductory logic in the self-paced Keller format.* Paper presented at the Conference on Critical Thinking in College Academic Enrichment Programs. Los Angeles, CA.

Bernsen, N. O. (1994). Foundations of multimodal representations: A taxonomy of representational modalities. *Interacting with Computers, 6*(4), 347-371.

Blackburn, P., & Bos, J. (1997). *Representation and inference for natural language: A first course in computational semantics.* Lecture Notes for ESSLLI97, Aix-en-Provence.

Blackwell, A. F., & Engelhardt, Y. (in press). A meta-taxonomy for diagram research. In P. Olivier, M. Anderson & B. Meyer (Eds.), *Diagrammatic representation and reasoning.* London: Springer-Verlag.

Burbules, N. C., & Reese, P. (1984, April). *Teaching logic to children: An exploratory study of "Rocky's Boots".* Paper presented at the American Educational Research Association (AERA) meeting, New Orleans, LA.

Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18,* 293-328.

Chinn, A. C., & 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), 1-49.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149-210.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31,* 187-276.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58,* 1-73.

Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction, 9*(4), 343-363.

Cox, R., Stenning, K., & Oberlander, J. (1994). Graphical effects in learning logic: Reasoning, representation and individual differences. In A. Ram & Eiselt, K. (Eds.), *Proceedings of the 16th Annual Conference of the Cognitive Science Society* (pp. 237-242). Mahwah, NJ: Lawrence Erlbaum.

Cox, R., Stenning, K., & Oberlander, J. (1995). The effect of graphical and sentential logic teaching on spontaneous external representation. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society, 2*(4), 56-75.

Dijkstra, S. (2000). Epistemology, psychology of learning and instructional design. In J. M. Spector & T. M. Anderson (Eds.), *Integrated and holistic perspectives on learning, instruction and technology: Understanding complexity* (pp. 213-232). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Dijkstra, S. (2001). The design of multimedia-based training. In S. Dijkstra, D. Jonassen & D. Sembill (Eds.), *Multimedia learning: Results and perspectives* (pp. 17-40). Berlin: Lang.

Dijkstra, S., Jonassen, D., & Sembill, D. (2001). The use of multimedia in education and training. In S. Dijkstra, D. Jonassen & D. Sembill (Eds.), *Multimedia learning: Results and perspectives* (pp. 3-13). Berlin: Lang.

Dobson, M. (1998). Toward decision support for multiple representations in teaching early logic. In M. W. Van Someren, P. Reimann, H. P. A. Boshuizen & T. de Jong (Eds.), *Learning with multiple representations. Advances in learning and instruction series* (pp. 87-101). Oxford: Elsevier Science, Ltd.

Dobson, M. (1999). Information enforcement and learning with interactive graphical systems. *Learning and Instruction, 9*(4), 365-390.

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

Eco, U. (1985). Producing signs. In M. Blonsky (Ed.), *On signs* (pp. 176-183). Baltimore, MD: John Hopkins University Press.

Edwards, L. D. (1993). Microworlds as representations. In A. A. diSessa, C. Hoyles & R. Noss (Eds.), *Computers and exploratory learning* (pp. 127-154). Berlin: Springer-Verlag (NATO ASI Series F, Computer and System Sciences, Vol. 146).

Eulenberg, A. (1996). Moreover, besides, what is more: More than just conjunctions. *Online Proceedings of the 1996 Midwest Artificial Intelligence and Cognitive Science Conference.* URL: http://www.cs.indiana.edu/event/maics96/Proceedings/eulenberg.html.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction.* Hove: Lawrence Erlbaum Associates, Publishers.

Eysink, T. H. S. (1998). [Formalisation in the development of logical concepts]. Unpublished raw data.

Eysink, T. H. S., Dijkstra, S., & Kuper, J. (2001a). Cognitive processes in solving variants of computer-based problems used in logic teaching. *Computers in Human Behavior, 17*(1), 1-19.

Eysink, T. H. S., Dijkstra, S., & Kuper, J. (2001b). *The role of guidance in computer-based problem solving for the development of concepts of logic.* Manuscript submitted for publication, University of Twente, Enschede, The Netherlands.

Freudenthal, H. (1991). *Revisiting mathematics education: China lectures.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Fung, P., O'Shea, T., Goldson, D., Reeves, S., & Bornat, R. (1994). Why computer science students find formal reasoning frightening. *Journal of Computer Assisted Learning, 10,* 240-250.

Fung, P., O'Shea, T., Goldson, D., Reeves, S., & Bornat, R. (1996). Computer tools to teach formal reasoning. *Computers Education, 27*(1), 59-69.

Gamut, L. T. F. (1982). *Logica, taal en betekenis 1: Inleiding in de logica* [Logic, language and meaning 1: Introduction into logic]. Utrecht, The Netherlands: Het Spectrum.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review, 98*(2), 254-267.

Goldson, D., Reeves, S., & Bornat, R. (1993). A review of several programs for the teaching of logic. *The Computer Journal, 36*(4), 373-386.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 41-58). New York: Academic Press.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic materials effect in the Wason selection task. *British Journal of Psychology, 73,* 407-420.

Johnson-Laird, P. N. (1989). Mental models. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 469-499). Cambridge, MA: The MIT Press.

Johnson-Laird, P. N., & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology, 1,* 134-148.

Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, M. S. (1972). Reasoning and a sense of reality. *British Journal of Psychology, 63,* 395-400.

Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology: Research and Development, 45*(1), 65-94.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology: Research and Development, 48*(4), 63-86.

de Jong, T., Ainsworth, S., Dobson, M., van der Hulst, A., Levonen, J., Reimann, P., Sime, J., van Someren, M. W., Spada, H., & Swaak, J. (1998). Acquiring knowledge in science and mathematics: The use of multiple representations in technology-based learning environments. In M. W. van Someren, P. Reimann, H. P. A. Boshuizen & T. de Jong (Eds.), *Learning with multiple representations. Advances in learning and instruction series* (pp. 9-40). Amsterdam: Pergamon.

de Jong, T., & Njoo, M. (1992). Learning and instruction with computer simulations: Learning processes involved. In E. de Corte, M. Linn, H. Mandl & L. Verschaffel (Eds.), *Computer-based learning environments and problem solving* (pp. 411-427). Berlin: Springer-Verlag.

de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*(2), 179-201.

Kahneman, D., Slovic, P., & Tverski, A. (1982). *Judgement under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press.

Kahneman, D., & Tverski, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237-251.

Kamp, H., & Reyle, U. (1993). *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory. Studies in Linguistics and Philosophy: Vol. 42.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94,* 211-228.

Koedinger, K. R., & Anderson, J. R. (1994). Reifying implicit planning in geometry: Guidelines for model-based intelligent tutoring system design. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 15-46). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Kuhn, D., Schauble, L., & Garcia-Mila, M. (1992). Cross-domain development of scientific reasoning. *Cognition and Instruction, 9*(4), 285-327.

Kulpa, Z. (1995). Diagrammatic representation and reasoning. *Machine Graphics and Vision, 3*(1-2), 77-103.

Kuper, J. (in press). Meaning, formalisation, and corpus linguistics. *Proceedings of the 5th TELRI Seminar: Corpus Linguistics - How to Extract Meaning from Corpora.* Lubljana, Sept 22-24, 2000.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chigaco Press.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65-99.

Lehman, D. R., & Nisbett, R. E. (1990). A longitudinal study of the effects of undergraduate training on reasoning. *Developmental Psychology, 26*(6), 952-960.

MacEachren, A. M. (1995). *How maps work: Representation, visualization, and design.* New York: Guilford Press.

Markman, A. B. (1999). *Knowledge representation.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*(4), 715-726.

van Merriënboer, J. J. G., & Krammer, H. P. M. (1990). The "completion strategy" in programming instruction: Theoretical and empirical support. In S. Dijkstra, B. H. A. M. van Hout-Wolters & P. C. van der Sijde (Eds.), *Research on instruction* (pp. 45-61). Englewood Cliffs, NJ: Educational Technology Publications.

Merrill, M. D. (2001, April). *First principles of instruction.* Paper presented at the American Educational Research Association (AERA) meeting, Seattle, WA.

Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1978). Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology, 30,* 395-406.

Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P .W. (1987). Teaching reasoning. *Science, 238,* 625-631.

Norman, D. A. (1988). *The psychology of everyday things.* New York: BasicBooks.

O'Brien, D. P., Braine, M. D. S., Connell, J. W., Noveck, I. A., Fisch, S. M., & Fun, E. (1989). Reasoning about conditional sentences: Development of understanding of cues to quantification. *Journal of Experimental Child Psychology, 48,* 90-113.

Paivio, A. (1986). *Mental representation: A dual coding approach.* New York: Oxford University Press.

van der Pal, J. (1995). *The balance of situated action and formal instruction for learning conditional reasoning.* Doctoral Dissertation, University of Twente, The Netherlands.

van der Pal, J., & Eysink, T. (1999). Balancing situativity and formality: The importance of relating a formal language to interactive graphics in logic instruction. *Learning and Instruction, 9*(4)*,* 327-341.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas.* New York: The Harvester Press Limited.

Papert, S. (1987). Computer criticism versus technocentric thinking. *Educational Researcher,16*(1), 22-30.

Piaget, J. (1973). *Psychologie en kennisleer* [Psychology and epistemology. Translated from: Psychology et epistémology, 1973]. Utrecht, The Netherlands: Spectrum.

Piaget, J. (1976). *Genetische epistemologie: Een studie van de ontwikkeling van denken en kennen* [Genetic epistemology: a study of the development of thinking and knowing. Translated from: Epistémologie génétique, 1970]. Meppel, The Netherlands: Boom.

Pintrich, P. R. (1990). Implications of psychological research on student learning and college teaching for teacher education. In W. R. Houston, M. Haberman, & J. Sikula (Eds.), *Handbook of research on teacher education* (pp. 826-857). New York: MacMillan.

Poletiek, F. H. (2001). *Hypothesis-testing behaviour. Essays in cognitive psychology.* Hove: Psychology Press Ltd.

Reimann, P. (1999). The role of external representations in distributed problem solving. *Learning and Instruction, 9,* 411-418.

Rieber, L. P. (1996). Animation as feedback in a computer-based simulation: Representation matters. *Educational Technology: Research and Development, 44*(1), 5-22.

Robinett, W., & Grimm, L. (1987). Rocky's Boots [computer software]. The learning company.

Rothbart, A. (1998). Learning to reason from Lewis Carroll. *The Mathematics Teacher, 91*(1), 6-10, 96.

Rumain, B., Connell, J., & Braine, M. D. S. (1983). Conversational comprehension processes are responsible for reasoning fallacies in children as well as adults: If is not the biconditional. *Developmental Psychology, 19*(4), 471-481.

Sanders, W. J., & Antes, R. L. (1988). Teaching logic with logic boxes. *Mathematics Teacher, 81*(8), 643-647.

Seel, N. M., & Winn, W. D. (1997). Research on media and learning: Distributed cognition and semiotics. In R. D. Tennyson, F. Schott, N. Seel & S. Dijkstra (Eds.), *Instructional design: International perspectives: Vol. 1. Theory, research, and models* (pp. 293-326). London: Lawrence Erlbaum Associates, Publishers.

Simon, H. A. (1978). On the forms of mental representation. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science: Vol. IX. Perception and cognition: Issues in the foundations of psychology* (pp. 3-18). Minneapolis: University of Minnesota Press.

Stenning, K. (1999). The cognitive consequences of modality assignment for educational communication: The picture in logic teaching. *Learning and Instruction, 9*(4), 391-410.

Stenning, K., Cox, R., & Oberlander, J. (1995). Contrasting the cognitive effects of graphical and sentential logic teaching: Reasoning, representation and individual differences. *Language and Cognitive Processes, 10*(3/4), 333-354.

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science, 19,* 97-140.

Teodoro, V. D. (1992). Direct manipulation of physical concepts in a computerized exploratory laboratory. In E. de Corte, M. C. Linn, H. Mandl & L. Verschaffel (Eds.), *Computer-based learning environments and problem-solving* (pp. 445-464). Berlin: Springer-Verlag.

Thompson, P. (1987). Mathematical microworlds and intelligent computer-assisted instruction. In G. Kearsly (Ed.), *Artificial intelligence and instruction: Applications and methods* (pp. 83-109). Reading, MA: Addison Wesley.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207-232.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology: Vol. I* (pp. 135-151). Harmondsworth: Penguin.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20,* 273-281.

Wason, P. C. (1969). Regression in reasoning? *British Journal of Psychology, 60,* 471-480.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content.* London: Batsford.

Wason, P. C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology, 23,* 63-71.

White, B. Y. (1993). ThinkerTools: causal models, conceptual change, and science education. *Cognition and Instruction, 10*(1), 1-100.

White, B. Y., & Frederiksen, J. R. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence, 42,* 99-175.

Zhang, J. (2000). The coordination of external representations and internal mental representations in display-based cognitive tasks. In M. Anderson, P. Cheng & V. Haarslev (Eds.), *Diagrams 2000: Theory and application of diagrams. Lecture notes in artificial intelligence 1889* (pp. 6-6). Berlin: Springer-Verlag.

Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science, 18,* 87-122.

Nederlandse samenvatting

(Dutch summary)

Instructie is een deel van een communicatieve activiteit die bedoeld is om de ontwikkeling van kennis en vaardigheden te initiëren en te ondersteunen. Dit kan gerealiseerd worden door de communicatie tussen een expert, bijvoorbeeld een leraar, mentor of instructieprogramma, en een beginner of leerling. Deze instructieve communicatie gaat over objecten en opvattingen over deze objecten en vindt plaats met behulp van een medium (Dijkstra, 2001). Doordat de computertechnologie zich de laatste tientallen jaren sterk ontwikkeld heeft, zijn de mogelijkheden van toepassingen in het onderwijs toegenomen en wordt ook de computer regelmatig als medium voor instructieve communicatie gebruikt. De computer kan objecten bijvoorbeeld op meerdere manieren representeren, de computer kan leerlingen met de betreffende objecten laten interacteren, en de computer kan gerichte feedback geven. Dit alles heeft geresulteerd in een hernieuwde interesse op het gebied van het ontwerpen van instructieve communicatie. Het effect van verschillende typen instructievariabelen op de ontwikkeling van kennis en vaardigheden werd bestudeerd en speciale interesse ontstond voor de effecten van visualisatietechnieken, de effecten van het gelijktijdig aanbieden van meerdere representaties en de mogelijkheden voor de leerlingen om te interacteren met representaties. De onderzoekers richtten hun aandacht op theorieën die dieper ingingen op het oplossen van authentieke problemen en het visualiseren van probleemsituaties. De studies waarin deze theorieën toegepast en op bruikbaarheid getoetst werden, gebruikten hoofdzakelijk lesmateriaal uit de empirische wetenschappen, zoals de biologie, de scheikunde en de natuurkunde. In de empirische wetenschappen worden representaties op zowel iconisch als symbolisch niveau weergegeven. In de formele wetenschappen, zoals de wiskunde en de logica, worden de concepten en operaties echter meestal enkel en alleen beschreven op een zuiver symbolische manier. Voor veel leerlingen leidt instructie met alleen symbolen tot een gebrek aan begrip van de formele regels en operaties. De technieken toegepast op materiaal uit de empirische wetenschappen zouden dan ook een positief effect kunnen hebben op het leren van materiaal uit de formele wetenschappen. De vraag in dit proefschrift gesteld, is dan ook hoe men het beste instructieve communicatie voor formele wetenschappen kan benaderen. Het

onderzoek waarvan in dit proefschrift verslag is gedaan heeft zich daarvoor gericht op de effecten van (a) formalisatie van taal, (b) visualisatie van objecten, (c) manipulatie van objecten, en (d) sturing van leerlingen op de ontwikkeling van concepten van de logica, in het bijzonder de eerste-orde predicaten logica.
Een samenvatting wordt nu gegeven van de inhoud van de hoofdstukken van dit proefschrift.

Hoofdstuk 1

In hoofdstuk 1 wordt de hypothese, dat de gekozen vorm van een representatie van formele objecten van invloed is op de concepties die leerlingen uiteindelijk ontwikkelen over deze objecten, theoretisch onderbouwd. Het hoofdstuk is opgedeeld in drie secties: (a) representaties, (b) redeneren met representaties, en (c) leren redeneren met representaties.

Het eerste deel bespreekt (externe) representaties in het algemeen. Representaties kunnen worden beschreven door een drievoudige relatie (bv. Seel & Winn, 1997; Markman, 1999; MacEachren, 1995) tussen: (a) een gerepresenteerd object (de werkelijkheid); (b) een representerend medium (de representatie zelf); en (c) een contextuele interpretatie. Het gerepresenteerde object en het representerende medium zijn gerelateerd aan elkaar door een verzameling van afbeeldingsprincipes (Bernsen, 1994). Deze afbeeldingsprincipes kunnen variëren in de mate van overeenkomst tussen het representerende medium en het gerepresenteerde object (homomorfisme of isomorfisme), de manier van afbeelden van het representerende medium ten opzichte van het gerepresenteerde object (bv. een iconische afbeelding door middel van natuurlijke afbeeldingsprincipes of een symbolische afbeelding door middel van arbitraire afbeeldingsprincipes) en de functie van de afbeelding van het representerende medium ten opzichte van het gerepresenteerde object (bv. letterlijke overeenkomst of metaforische overeenkomst). De keuze voor bepaalde afbeeldingsprincipes en dus de keuze voor een bepaald type representatie is van invloed op de manier waarop leerlingen de informatie tot zich nemen. In dit eerste deel van hoofdstuk 1 worden vervolgens kenmerken van visuele representaties en representaties in taal besproken, hun voor- en nadelen en de manier waarop deze representaties meestal in de logica gebruikt worden. Visuele representaties worden in de logica meestal gebruikt om de werkelijkheid af te beelden; representaties in taal worden meestal gebruikt om beweringen over de realiteit te doen en om te redeneren over die realiteit. Beide typen representaties kunnen variëren in abstractieniveau.

Het tweede deel van hoofdstuk 1 behandelt het redeneren met representaties. Denken omvat redeneren met representaties. Dit redeneren kan zowel impliciet als expliciet gebeuren. Ervaringen uit het verleden kunnen worden gegeneraliseerd en deze kennis kan gebruikt worden in nieuwe situaties om conclusies te trekken. Echter, het redeneren in het dagelijks leven gebeurt niet altijd volgens de strikte regels van de logica. Gebruikersaspecten, zoals bijvoorbeeld de context waarin een redenatie zich afspeelt, en heuristieken om het redeneren te vergemakkelijken interfereren met het redeneerproces volgens de logische regels. Omdat bij het leren van logica leerlingen hun naïeve concepties opgedaan door ervaringen in het alledaagse leven meenemen, is het belangrijk te weten wat de overeenkomsten en de verschillen zijn tussen het redeneren in het dagelijks leven en de formele logica. In dit tweede deel van hoofdstuk 1 is geprobeerd deze verschillen in kaart te brengen. Deze analyse laat zien dat er drie hoofdverschillen zijn tussen alledaags redeneren en formeel redeneren. Het eerste verschil betreft verschillen in abstractieniveau. De analyse onderscheidt hier vier typen waarvan de logica abstraheert, maar die in de natuurlijke taal wel voor kunnen komen: subtiele betekenisverschillen, ambiguïteit, interne relaties (contextueel, temporeel, causaal) en context. Het tweede verschil betreft waarheidsconditionele aspecten, dat wil zeggen dat zinnen in de natuurlijke taal soms andere waarheidswaardes hebben dan dezelfde zinnen die puur logisch benaderd worden. Tevens kan een zin in de natuurlijke taal onzin zijn, geen waarheidswaarde hebben of er kan onzekerheid bestaan over het waar of niet waar zijn van deze zin. Het derde verschil betreft gebruikersconditionele (ofwel pragmatische) aspecten, zoals bijvoorbeeld het conversationele coöperatie-principe (Grice, 1975) dat ervan uit gaat dat iedereen in een conversatie zich zo gedraagt dat het gezamenlijke doel van de communicatie wordt bereikt op een zo optimaal mogelijke manier.

Het derde deel van hoofdstuk 1 behandelt het leren redeneren met representaties. Volgens Piaget (1973) moet logica geleerd worden door middel van logisch-mathematische ervaringen. Door veelvuldig te interacteren met objecten kan kennis worden geabstraheerd van die interacties. Op een gegeven moment is de interactie met fysieke objecten echter niet meer nodig en zijn de logisch-mathematische operaties geïntegreerd in symbolische operatoren, die in verschillende contexten gebruikt kunnen worden. Wat er dan overgebleven is, is pure logica waarvoor geen ervaring meer nodig is. Dit leidt dus tot de conclusie dat het leren van logica gebeuren moet door te interacteren met concrete objecten, zodat de leerling vanuit zijn ervaringen uiteindelijk de betekenis van de logische operatoren kan abstraheren.

Hoofdstuk 2

Dit hoofdstuk probeert een brug te slaan tussen de theorie en de experimenten door de beslissingen die genomen zijn voor de experimenten te verantwoorden met behulp van de theorie van hoofdstuk 1. Mensen ontwikkelen ideeën over de regels van de logica door ervaringen in het alledaagse leven. Wanneer deze 'alledaagse regels' overeenkomen met de logische regels zullen zij geen problemen ondervinden. Die problemen ontstaan op het moment dat er verschillen zijn tussen beide soorten regels. In hoofdstuk 1 werd duidelijk dat er waarheidsconditionele aspecten en gebruikersconditionele aspecten gebruikt worden in het redeneren van alledag, maar dat de logica alleen de waarheidsconditionele aspecten erkent en abstraheert van de gebruikersconditionele aspecten. Echter, er zijn ook verschillen tussen het alledaags redeneren en het formele redeneren wat betreft die waarheidsconditionele aspecten. Conclusies die 'logisch' lijken in het alledaagse leven, blijken opeens incorrect te zijn in de logica, situaties die contra-intuïtief worden genoemd. Verder kunnen logische abstracties ook voor misconcepties zorgen. Eén van de hoofddoelen van een introductiecursus in de eerste-orde predicaten logica is het leren van de condities waaronder zinnen waar zijn en de condities waaronder zinnen onwaar zijn, de zogenaamde waarheidsconditionele aspecten. In de experimenten is ervoor gekozen de aandacht op deze waarheidsconditionele aspecten te richten. Contra-intuïtieve situaties en misconcepties vormen daarvoor een interessant beginpunt, terwijl situaties waarin gebruikersconditionele aspecten voorkomen vermeden dienen te worden, zodat de leerlingen zich compleet op de waarheidsconditionele aspecten kunnen richten zonder enige interferentie van andere, onduidelijk pragmatische aspecten. Dit betekent dat objecten en situaties uit het alledaagse leven niet gebruikt kunnen worden. Daarnaast heeft het gebruik van abstracte objecten als nadeel dat er geen integratie in de context is en dat begrip moeilijk te realiseren is. Daarom is gekozen voor een instructie met concrete, maar volledig kenbare objecten, namelijk geometrische objecten die bepaalde relaties tot elkaar hebben, zoals "de grote kubus staat rechts van de kleine tetraëder". Alle operaties zijn mogelijk en de concrete objecten zijn volledig gedefinieerd, zodat irrelevante kenmerken van de context geen rol kunnen spelen. Abstracte principes kunnen worden gerelateerd aan concrete objecten, zodat betekenisvol begrip kan worden bereikt. Tevens hebben onbedoelde gebruikersconditionele aspecten geen invloed, zodat de leerlingen zich volledig op het leren van de waarheidsconditionele aspecten kunnen richten.

Naast het gebruik van concrete, geometrische objecten is in de experimenten gebruik gemaakt van het oplossen van problemen, omdat dit door vele onderzoekers gezien wordt als de meest effectieve manier van leren (bv. Dijkstra, 2000, 2001; Jonassen, 1997, 2000; Merrill, 2001). Een computer-gestuurde leeromgeving waarin dit mogelijk is, is Tarski's World (Barwise & Etchemendy, 1992). Tarski's World is geschikt voor het leren van waarheidsconditionele aspecten van de eerste-orde predicaten logica. Het maakt gebruik van geometrische objecten die bepaalde relaties tot elkaar kunnen hebben en waarop operaties toegepast kunnen worden. In de experimenten is tevens gebruik gemaakt van meerdere representaties, te weten beweringen in natuurlijke taal, beweringen in eerste-orde predicaten logica, verbale beschrijvingen van de werkelijkheid in natuurlijke taal en/of grafische representaties van de werkelijkheid waarin de betreffende beweringen waar of onwaar zijn.

Hoofdstuk 3

In dit hoofdstuk wordt het eerste experiment beschreven. Dit experiment is een gedeeltelijke replicatiestudie van het onderzoek van Van der Pal en Eysink (1999). In dit experiment is het effect bestudeerd van de symbolische, logische expressie op de kennisontwikkeling van logische concepten. Het doel was te bepalen of een visuele representatie toegevoegd aan een representatie in natuurlijke taal voldoende was om kennis over logische concepten en regels te ontwikkelen, of dat een extra symbolische representatie noodzakelijk was om deze kennis te ontwikkelen. Om dit te kunnen onderzoeken, zijn 49 eerstejaars technische studenten getest, zodat drie condities vergeleken konden worden: (a) een conditie S met zinnen in natuurlijke taal en een grafische representatie, (b) een gecombineerde conditie SF met zinnen in natuurlijke taal en een grafische representatie gecombineerd met een formele, symbolische representatie, en (c) een controle conditie C, waarin studenten naast een algemene training geen instructie ontvingen. Om de kennisontwikkeling van de studenten te meten, is een voor-, na-, en retentietoets gebruikt. Deze toetsen bestonden uit vier Wason taken (Wason, 1966). De resultaten laten zien dat de studenten in de gecombineerde SF-conditie veel moeite hadden met het vertalen van de natuurlijke taalzinnen in formele expressies. Daardoor zal het vertaalproces geïnterfereerd hebben met het leren van de waarheidsconditionele aspecten van de logica, wat heeft geleid tot minder goede resultaten voor de gecombineerde SF-instructie dan voor de S-conditie en de C-conditie.

Hoofdstuk 4

Hoofdstuk 4 beschrijft het onderzoek van Eysink, Dijkstra en Kuper (2001a) dat de gecombineerde grafische en formele conditie uit het eerste experiment verder bekijkt. Deze conditie bevat in feite twee variabelen, namelijk het gebruik van een grafische realiteit en het manipuleren van objecten in deze realiteit. In dit tweede onderzoek is nagegaan wat de invloed van beide variabelen, visualisatie en manipulatie, is op de kennisontwikkeling van logische concepten en regels. Om het effect van visualisatie te kunnen bepalen is een conditie waarin een verbale beschrijving in natuurlijke taal werd gegeven vergeleken met een conditie waarin een visuele representatie werd gegeven. Om het effect van manipulatie te kunnen bepalen is een conditie waarin een visuele representatie werd gegeven die niet gemanipuleerd kon worden vergeleken met een conditie waarin deze visuele representatie wel gemanipuleerd kon worden. Omdat de conditie met een visuele representatie overeenkwam met de representatie waarin gemanipuleerd kon worden, leidde dit tot een vergelijking van drie condities: (a) de verbale, niet te manipuleren conditie SN, (b) de grafische, niet te manipuleren conditie GN, en (c) de grafische, te manipuleren conditie GM. Om het effect van beide variabelen te kunnen onderzoeken zijn 66 eerstejaars studenten van diverse sociale wetenschappen gelijkelijk verdeeld over de drie condities. De kennisontwikkeling is ook in dit onderzoek gemeten met een voor-, na-, en retentietoets. De toetsen zijn, vergeleken met het eerste experiment, uitgebreid met enkele nieuwe en makkelijker op te lossen items, zodat transfer van kennis beter gemeten kon worden. De resultaten laten zien dat er wat betreft de kennisontwikkeling geen verschil optreedt tussen de studenten in de visuele conditie en de studenten in de verbale conditie. Echter, een groot deel van de studenten die alleen verbale beschrijvingen kregen, maakten zelf visuele afbeeldingen van de beschrijvingen. Dit betekent dat de studenten de voorkeur gaven aan visuele afbeeldingen boven verbale beschrijvingen, maar het betekent bovenal dat studenten in de verbale conditie de visuele afbeeldingen blijkbaar nodig hadden om tot hetzelfde resultaat te komen als de studenten in de visuele conditie. Verder laten de data zien, dat studenten de mogelijkheden om te manipuleren helemaal niet gebruikten. Ze bleven in bekende situaties waar ze het antwoord al op wisten.

Hoofdstuk 5

Dit hoofdstuk beschrijft de studie van Eysink, Dijkstra en Kuper (2001b) waarin de manipulatie-variabele verder onderzocht is. Om een conclusie te kunnen trekken ten aanzien van de invloed van de manipulatie-variabele op de ontwikkeling van kennis over logische concepten en regels, moeten de studenten gestimuleerd worden om de mogelijkheden tot manipulatie daadwerkelijk te gebruiken. Daarom is voor dit derde experiment een nieuwe variabele geïntroduceerd, namelijk sturing. Deze variabele moest ervoor zorgen dat studenten naar alle mogelijke, en met name naar de moeilijke, contra-intuïtieve situaties geleid werden door ze te stimuleren objecten te manipuleren. Om het effect van manipulatie te kunnen bepalen is een conditie, waarin een visuele representatie werd gegeven die niet gemanipuleerd kon worden, vergeleken met een conditie, waarin deze visuele representatie wel gemanipuleerd kon worden. Om het effect van sturing te bepalen is een conditie, waarin problemen werden gegeven die de studenten naar alle verschillende typen probleemsituaties leidden en die hen stimuleerden om de objecten van de visuele representatie (fysiek of mentaal) te manipuleren, vergeleken met een conditie, waarin deze sturing niet werd gegeven. Dit leidde tot vier condities: (a) de te manipuleren, gestuurde conditie Man+Guid+; (b) de te manipuleren, niet gestuurde conditie Man+Guid-; (c) de niet te manipuleren, gestuurde conditie Man-Guid+; en (d) de niet te manipuleren, niet gestuurde conditie Man-Guid-. In totaal 72 eerste- en tweedejaars studenten van diverse sociale wetenschappen zijn gelijkelijk verdeeld over de drie condities. Kennisontwikkeling is gemeten door een voor-, na-, en retentietoets. De toetsen zijn, vergeleken met de eerste twee experimenten, uitgebreid zodat ze uit 27 items bestonden verspreid over drie delen: (1) een 'near-transfer test' met 14 items, die de kennisontwikkeling binnen Tarski's World toetste; (2) een 'middle-transfer test' met 9 items waarin de kennisontwikkeling gemeten werd door middel van logisch-deductieve problemen in een alledaagse situatie; en (3) een 'far-transfer test' met 4 items waarin de kennisontwikkeling gemeten werd door middel van Wason taken. De resultaten laten zien dat de studenten die niet werden gestuurd in hun oplossingsproces niet profiteerden van instructie, aangezien zij óf geen mogelijkheid tot manipuleren hadden óf uit zichzelf geen gebruik maakten van die mogelijkheid. De studenten die gestuurd werden in hun probleem-oplossingsproces profiteerden wel van de gegeven instructie, onafhankelijk van het feit of ze de mogelijkheid hadden om te manipuleren of niet. Het meest opvallende resultaat, echter, kwam duidelijk naar voren na twee weken: de studenten die werden gestuurd maar niet de mogelijkheid tot manipulatie hadden, onthielden minder van de geleerde stof, terwijl de studenten die waren gestuurd en wel manipulatie-mogelijkheden hadden hun scores zelfs nog verbeterden.

Hoofdstuk 6

Hoofdstuk 6 begint met een reflectie op de drie uitgevoerde experimenten. Deze reflectie begint met de bespreking van de instructievariabelen. De invloed van een concrete context is niet expliciet onderzocht, maar de studenten gaven in het tweede experiment duidelijk te kennen dat ze de voorkeur gaven aan de concrete context in vergelijking met een abstracte context. De invloed van visualisatie is in het tweede experiment onderzocht. Er werd geen verschil gevonden tussen de conditie waarin de realiteit gerepresenteerd werd door middel van een visuele afbeelding en de conditie waarin deze realiteit gerepresenteerd werd door een verbale beschrijving. De studenten maakten echter zelf visualisaties indien een verbale beschrijving gegeven was. Dit had tot effect dat het verschil tussen de verbale conditie en de visuele conditie niet duidelijk meer aanwezig was, hetgeen een verklaring kan zijn voor het uitblijven van verwachte resultaten. Echter, het feit dat de studenten zelf visualisaties maakten, betekent dat de studenten in deze conditie de visuele representaties blijkbaar nodig hadden om tot dezelfde resultaten te komen als de studenten in de visuele conditie. De invloed van manipulatie is onderzocht in zowel het tweede als het derde experiment. In het tweede experiment maakten de studenten niet of nauwelijks gebruik van de mogelijkheden om te manipuleren. Dit had tot effect dat de manipulatie-conditie en de niet-manipulatie-conditie nauwelijks van elkaar verschilden, hetgeen een verklaring kan zijn voor het uitblijven van een positief effect van de manipulatie-variabele. De invloed van manipulatie werd echter

wel duidelijk in het derde experiment, waar de studenten aangespoord werden om te manipuleren. De resultaten laten zien, dat de studenten die manipuleerden het lesmateriaal na twee weken instructie-vrije periode beter onthielden dan studenten die niet konden manipuleren. Manipulatie wordt daarom gezien als een goede instructievariabele voor het leren van logica, maar alleen in combinatie met ondersteuning die ervoor zorgt dat studenten ook daadwerkelijk gebruik maken van de mogelijkheden om te manipuleren. Het positieve effect van sturing werd duidelijk in het derde experiment. De studenten die niet gestuurd werden in hun probleem-oplossingsproces leerden niet, terwijl de studenten die wel gestuurd werden een duidelijke uitbreiding van hun kennis lieten zien.
Al met al kan geconcludeerd worden dat de instructievariabelen die vaak gebruikt worden in de empirische wetenschappen ook gebruikt kunnen worden voor het leren van logica en wellicht voor het leren van alle formele wetenschappen.

Na de reflectie op de instructievariabelen, worden achtereenvolgens de designs, de methodologie, de proefpersonen en de transfertoetsen behandeld, die per experiment verschilden. Met name de ontwikkeling van de transfertoetsen bleek een goede onwikkeling te zijn, omdat een beter inzicht werd verkregen in het wel of niet kunnen toepassen door leerlingen van de nieuw ontwikkelde kennis op verschillende typen situaties.

Het tweede deel van hoofdstuk 6 bespreekt de algemene implicaties die de onderzoeken hebben voor het uitvoeren van vervolgonderzoek en voor de onderwijspraktijk op het gebied van de logica. In vervolgonderzoek zou de transfer van het geleerde verder bestudeerd kunnen worden, zodat duidelijk wordt of de leerlingen ook daadwerkelijk algemene, abstracte regels hebben ontwikkeld en of deze regels toegepast kunnen worden op alle, dus ook abstracte situaties. Een tweede suggestie voor vervolgonderzoek is om de manipulatie-variabele verder te onderzoeken. Zo zou een vraag bijvoorbeeld kunnen zijn of het voor een positief leereffect noodzakelijk is dat leerlingen objecten in een wereld manipuleren of dat de manipulatie van zinnen ook zou kunnen volstaan. Het zou interessant zijn om te onderzoeken of deze twee en wellicht nog andere soorten van manipulatie hetzelfde effect op kennisontwikkeling hebben of dat ze leiden tot een ander soort of een andere hoeveelheid kennis.

De onderzoeken kunnen tevens aanwijzingen geven voor de onderwijspraktijk. Ten eerste is het belangrijk om de moeilijkheden die de logica veroorzaakt op te lossen. Dat wil zeggen dat de abstracte regels van de logica geïntegreerd moeten worden in een concrete context, van waaruit de leerlingen de abstracte regels kunnen abstraheren. Tevens moeten gebruikersconditionele aspecten die een rol spelen in het alledaags redeneren niet kunnen interfereren met het leren van formele redenaties, waar deze pragmatische aspecten helemaal geen rol spelen.

Ten tweede moet het leren van logica meer inhouden dan slechts het manipuleren van formules. Het gaat om de semantiek van de logica, dus de betekenis van de logische aspecten en hun relatie tot een bepaalde werkelijkheid. In de onderzoeken die beschreven zijn in dit proefschrift werden aan de hand van de moeilijkheden, die studenten kunnen ondervinden bij het leren van logica, instructievariabelen zoals concretisering en visualisering gebruikt. Echter, wanneer er sprake is van leerlingen die begrip kunnen ontwikkelen aan de hand van de formele, symbolische concepten en operaties, dan moeten deze leerlingen niet lastig gevallen worden met deze instructievariabelen.

Ten derde moet een leeromgeving ingebed worden in een sociale context en moet de vraag gesteld worden hoe de omgeving gebruikt gaat worden en wat het doel van de omgeving is. Het gebruik van een leeromgeving als onderzoeksomgeving verschilt bijvoorbeeld van het gebruik van dezelfde omgeving met als doel te leren. In het laatste geval gaat de sociale context een veel grotere rol spelen, kunnen vragen gesteld worden door de leerlingen en is het doel voor hen duidelijker. Tevens kan een omgeving gebruikt worden om een boek te ondersteunen met problemen die de studenten moeten oplossen om zich zo de stof meer eigen te maken (een voorbeeld van deductief leren), of kan de omgeving gebruikt worden als op zichzelf staand lesmateriaal waaruit de leerlingen algemene regels kunnen abstraheren (een voorbeeld van inductief leren).

Een vierde implicatie betreft de sturing van leerlingen. Leerlingen moeten naar die situaties gestuurd worden die kritisch kunnen zijn voor de kennisontwikkeling. Van tevoren moet worden vastgesteld op welke manier deze sturing aangeboden wordt en welke mate van sturing voor de leerlingen geschikt en nodig is.

Dit laatste punt hangt direct samen met een vijfde punt waar rekening mee moet worden gehouden, namelijk het besef van de consequenties, die open leeromgevingen met zich mee brengen. Vragen moeten worden gesteld, zoals waar de balans ligt tussen het 'open' zijn van een omgeving en de sturing binnen zo'n omgeving, of deze balans afhangt van de karakteristieken van de leerling, of leerlingen verantwoordelijk kunnen worden gehouden voor hun eigen leerproces, en wat de effecten op de kwantiteit en de kwaliteit van de kennisontwikkeling is op zowel korte als lange termijn.

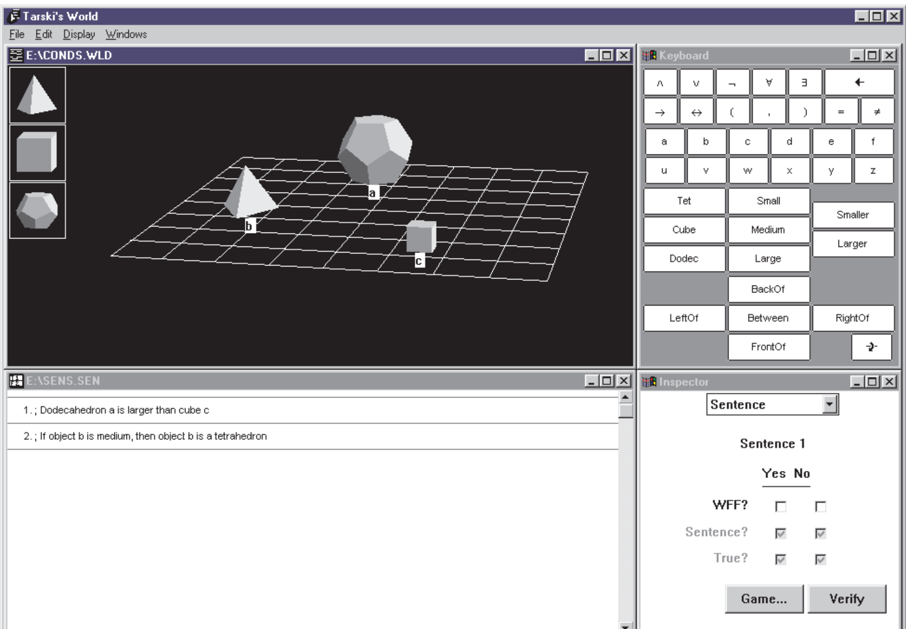
Appendix A

Conditions in the first experiment: S and SF[1]

[1] Examples were translated from Dutch into English.

Condition S

In most assignments, the students in this condition are given an empty world and several natural language sentences. The students have to construct a world in which the natural language sentences are true or false and they have to check this. See Figure A-1, for a typical example of a problem given to the students in the S-condition. In this figure, the student already constructed a world.



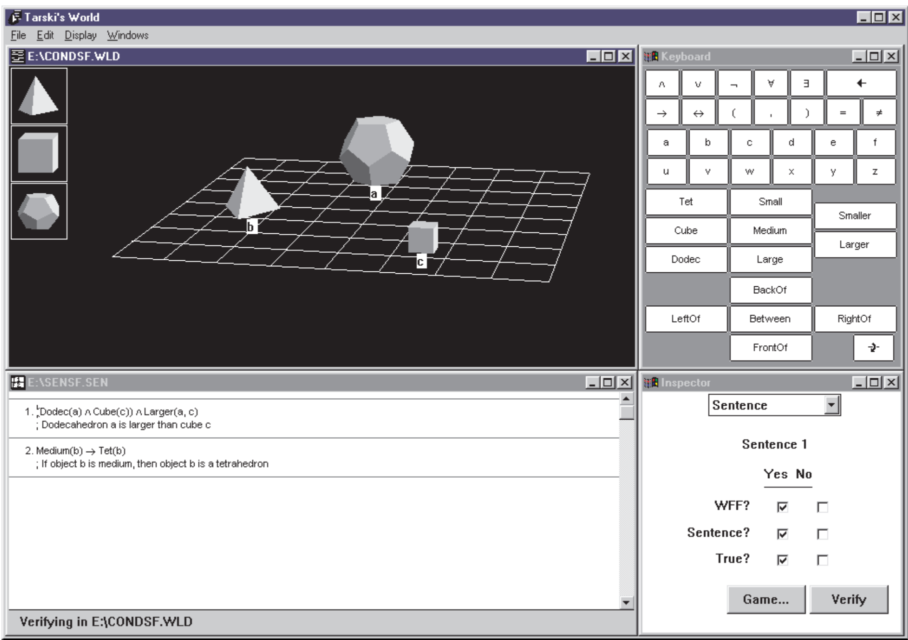
Assignment:

1. Construct a world in which both natural language sentences are true.
2. Check whether the sentences are indeed true. Correct if necessary.

Figure A-1. ***A typical example of a problem given to students in the condition S.***

Condition SF

In most assignments, the students in this condition are given an empty world and natural language sentences. First, they have to translate the natural language sentences into correct formal expressions. Then, they have to construct a world in which the sentences are true or false and they have to check this. For a typical example of a problem given to the students in the SF-condition, see Figure A-2. In this figure, the student already constructed a world.



<u>Assignment:</u>

1. Translate the two given natural language sentences into correct logical expressions.
2. Construct a world in which both natural language sentences are true.
3. Check whether the sentences are indeed true. Correct if necessary.

Figure A-2. ***A typical example of a problem given to students in the condition SF.***

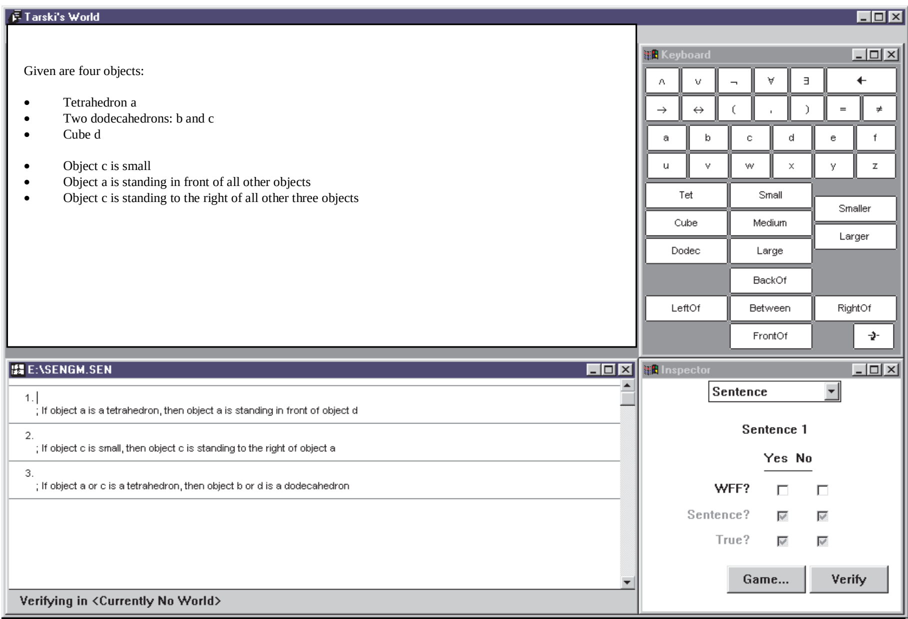
Appendix B

Conditions in the second experiment: SN, GN and GM[1]

[1] Examples were translated from Dutch into English.

Condition SN

The students in this condition are given a verbal description of a situation and several natural language sentences. First, they have to translate the given natural language sentences into correct formal expressions. Then, they have to check whether the sentences are true or false given the verbal description. For a typical example of a problem given to the students in the SN-condition, see Figure B-1.



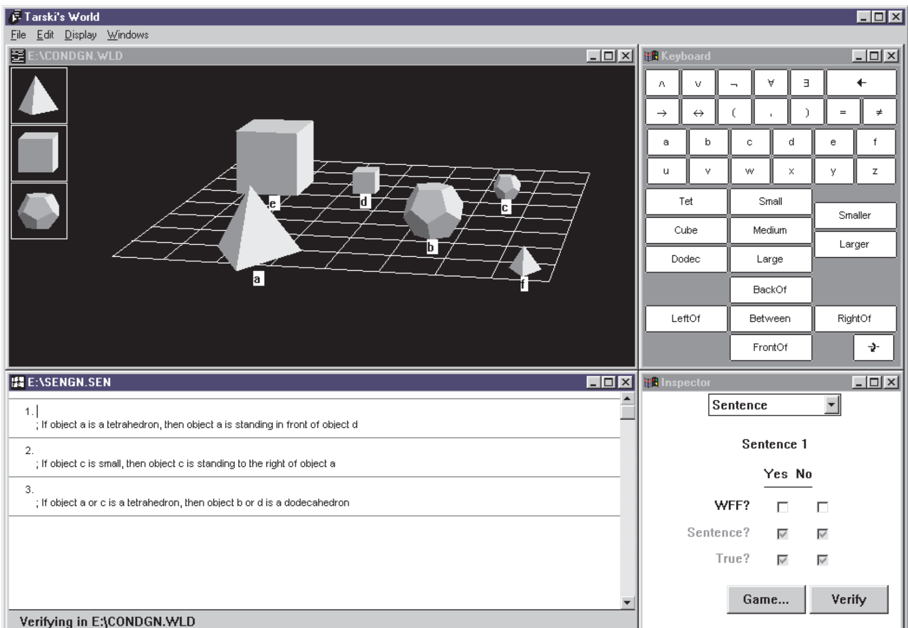
<u>Assignment:</u>

1. Translate the three given natural language sentences into correct logical expressions.
2. Predict whether the sentences are true or false given the description of the situation and check your prediction.

Figure B-1. ***A typical example of a problem given to students in the condition SN.***

Condition GN

The students in this condition are given a graphical depiction of a situation, that is a world with objects placed in it, and several natural language sentences. First, they have to translate the given natural language sentences into correct formal expressions. Then, they have to check whether the sentences are true or false given the objects in the world. The objects in the world cannot be manipulated. For a typical example of a problem given to the students in the GN-condition, see Figure B-2.



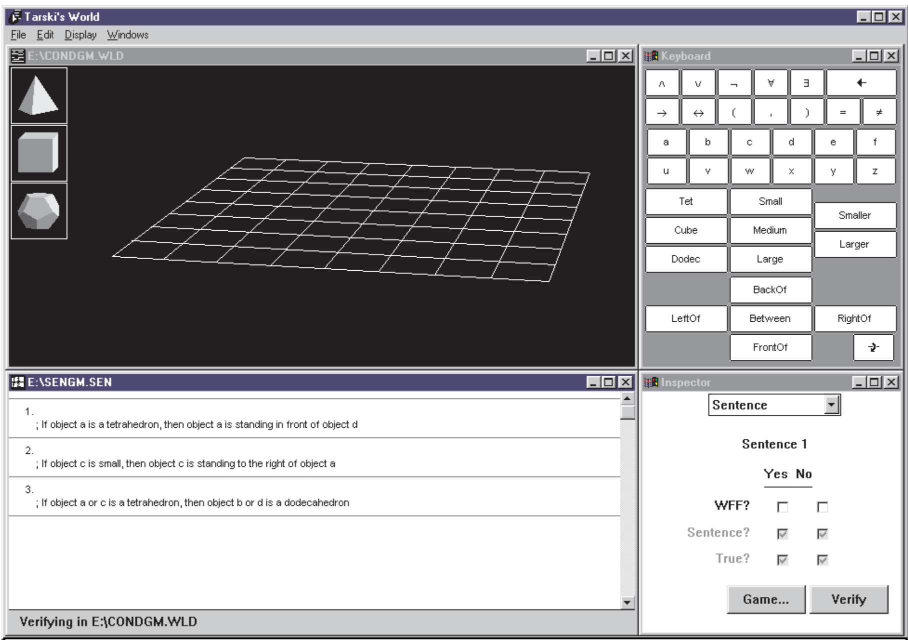
<u>Assignment:</u>

1. Translate the three given natural language sentences into correct logical expressions.
2. Predict whether the sentences are true or false in the given world and check your prediction.

Figure B-2. ***A typical example of a problem given to students in the condition GN.***

Condition GM

In most assignments, the students in this condition are given an empty world and several natural language sentences. First, they have to translate the given natural language sentences into correct formal expressions. Then, they have to put objects in the world, in such a way that the sentences are true or false in their constructed world. Then, they have to check whether the sentences are indeed true or false. For a typical example of a problem given to the students in the GM-condition, see Figure B-3.



Assignment:

1. Translate the three given natural language sentences into correct logical expressions.
2. Construct a world in which all natural language sentences are true.
3. Check whether the sentences are indeed true. Correct if necessary.

Figure B-3. ***A typical example of a problem given to students in the condition GM.***

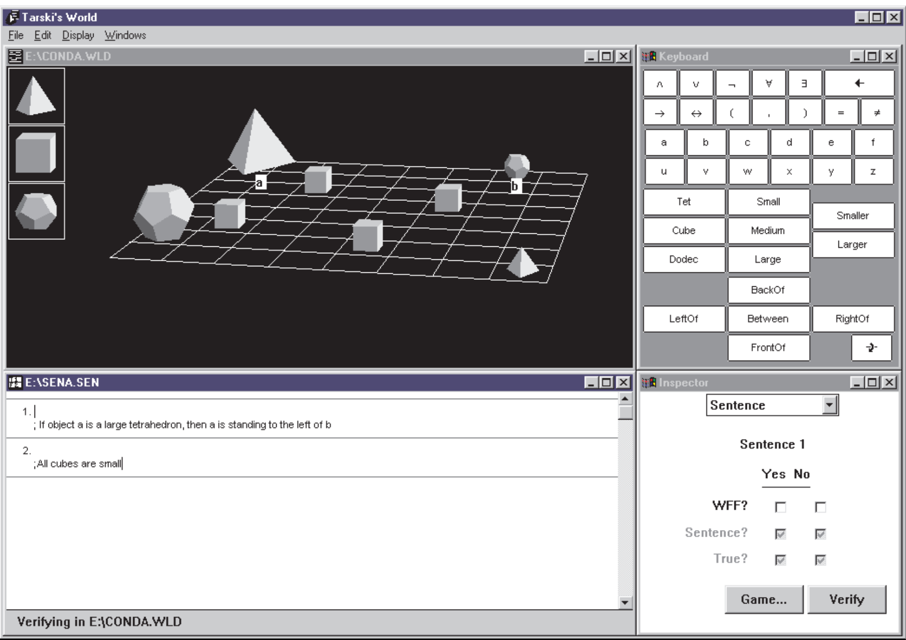
Appendix C

Conditions in the third experiment: Man+Guid+, Man+Guid-, Man-Guid+, Man-Guid-[1]

[1] Examples were translated from Dutch into English.

Condition Man+Guid+

In most assignments, the students in this condition are given a world with objects placed in it and some natural language sentences. First, they have to translate the given natural language sentences into correct formal expressions. Then, they have to predict whether the sentences are true or false in the world given and they have to check this. Then, the students have to change the world in a certain prescribed way. They have to predict what effect this change has on the truth values of the sentences and they have to check this. For a typical example of a problem given to the students in the Man+Guid+ condition, see Figure C-1.



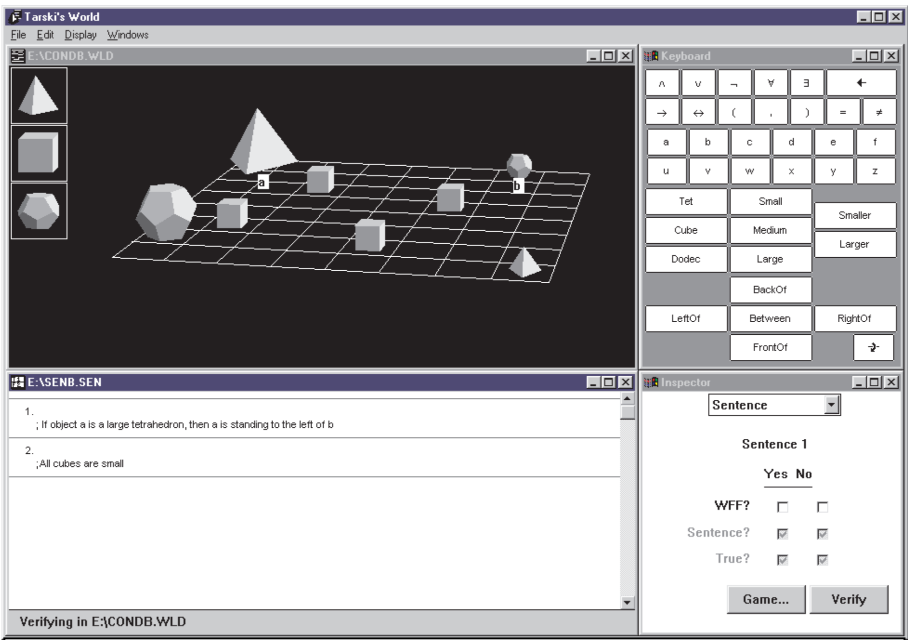
Assignment:

1. Translate the two given natural language sentences into correct logical expressions.
2. Predict whether the sentences are true or false in the world given and verify this.
3. Change the world in such a way that:
   - no cubes are left
   - all tetrahedrons are changed into dodecahedrons
4. Predict again whether the sentences are true or false in the changed world.

Figure C-1. ***A typical example of a problem given to students in the condition Man+Guid+.***

Condition Man+Guid-

In most assignments, the students in this condition are given a world with objects placed in it and some natural language sentences. First, they have to translate the given natural language sentences into correct formal expressions. Then, they have to predict whether the sentences are true or false in the world given and they have to check this. Then, the students have to change the world in whatever way they like. They have to predict what effect their change has on the truth values of the sentences and they have to check this. For a typical example of a problem given to the students in the Man+Guid- condition, see Figure C-2.



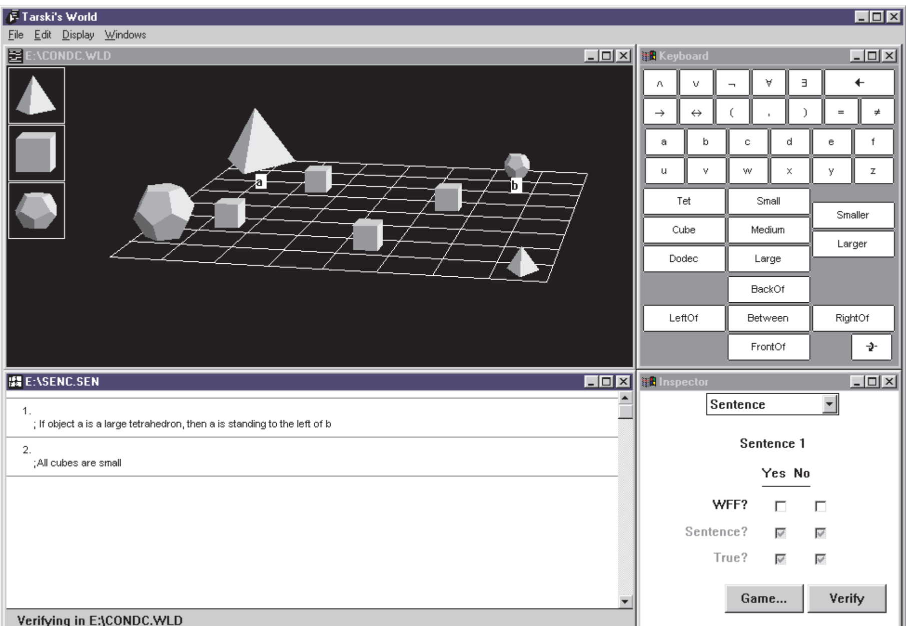
Assignment:

1. Translate the two given natural language sentences into correct logical expressions.
2. Predict whether the sentences are true or false in the world given and verify this.
3. Now change the world and predict again whether the sentences are true or false in your new world.

Figure C-2. ***A typical example of a problem given to students in the condition Man+Guid-***

Condition Man-Guid+

The students in this condition are given a world with objects placed in it and some natural language sentences. First, they have to translate the given natural language sentences into correct formal expressions. Then, they have to predict whether the sentences are true or false in the world given and they have to check this. Then, the students have to imagine that the world is changed in a certain prescribed way. They have to predict what effect this change would have on the truth values of the sentences and they have to check this. For a typical example of a problem given to the students in the Man-Guid+ condition, see Figure C-3.



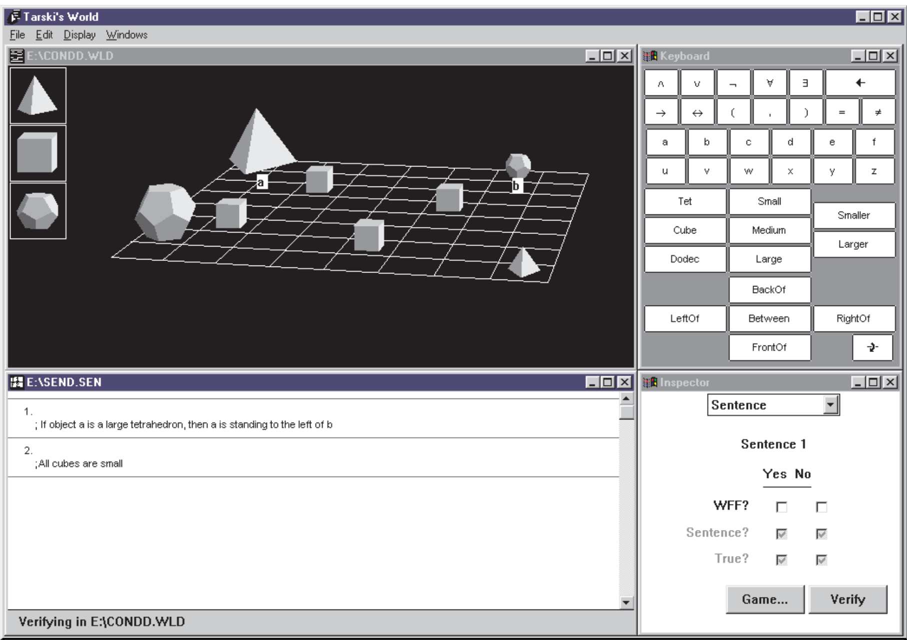
<u>Assignment:</u>

1. Translate the two given natural language sentences into correct logical expressions.
2. Predict whether the sentences are true or false in the world given and verify this.
3. Now imagine, that the world is changed in such a way that:
    - no cubes are left
    - all tetrahedrons are changed into dodecahedrons
4. Predict whether the sentences will be true or false in this imagined world.

Figure C-3. ***A typical example of a problem given to students in the condition Man-Guid+.***

Condition Man-Guid-

The students in this condition are given a world with objects placed in it and some natural language sentences. First, they have to translate the given natural language sentences into correct formal expressions. Then, they have to add some other sentences to the sentences given and they have to translate these, too. Then, they have to predict whether the sentences are true or false in the world given and they have to check this. For a typical example of a problem given to the students in the Man-Guid- condition, see Figure C-4.



Assignment:

1. Translate the two given natural language sentences into correct logical expressions.
2. Add to these two sentences two new sentences and translate these into correct logical expressions, too.
3. Predict whether the four sentences are true or false in the world given and verify this.

Figure C-4. ***A typical example of a problem given to students in the condition Man-Guid-.***

Appendix D

Transfer tests of the third experiment[1]

[1] The test presented is the pre test of the third experiment translated from Dutch into English. The correct answers are indicated. The post- and retention tests were similar to this test.

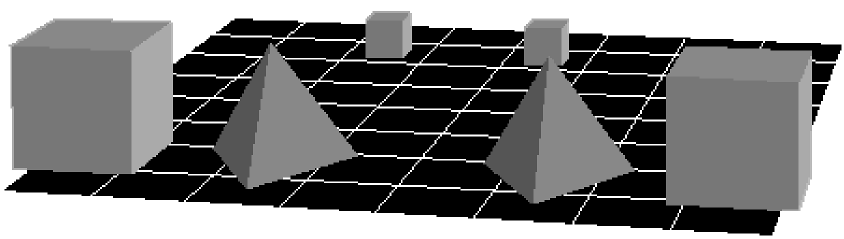
Part 1

Problem 1

A table can contain a number of objects with a certain shape (tetrahedron, cube or dodecahedron) and a certain size (small, medium, large).

Given is the table below containing:

- two small cubes
- two large cubes
- two large tetrahedrons
- no dodecahedrons

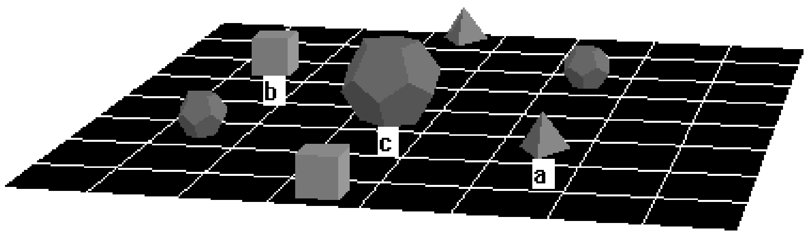


Indicate for each of the sentences below whether they are true or false for the table given.

		true	false
1.	All cubes are large.	○	⦿
2.	All dodecahedrons are large.	⦿	○
3.	All tetrahedrons are large.	⦿	○

Problem 2

Given is a table containing a number of objects. Except for the medium object *c*, all objects are small.



Indicate for each of the sentences below whether they are true or false for the table given.

		true	false
1.	If object *a* is a tetrahedron, then *a* is standing to the right of a small cube.	◉	○
2.	If object *c* is standing between object *a* and object *b*, then object *c* is small.	○	◉
3.	If there is a medium tetrahedron, then object *a* is small.	◉	○
4.	A large cube is standing in front of a large tetrahedron if object *b* or object *c* is a large cube.	◉	○
5.	If there is an object left of *b*, then this object is large.	◉	○

Problem 3

Given are four sentences:

1. There is one large cube.
2. Object *c* is standing in front of object *d.*
3. If object *a* is a large cube, then object *b* is a small tetrahedron.
4. If there is a large object, then tetrahedron *c* is standing in front of dodecahedron *d*.

A. Given is that sentence 1 is **false**.

- What does this tell you about the truth of **sentence 3**?

  ◉ sentence 3 is true
  ○ sentence 3 is false

- What does this tell you about the shape and size of object *b*?

  ○ object *b* is a small tetrahedron
  ○ object *b* is not a small tetrahedron
  ◉ a conclusion about the shape and size of object *b* cannot be drawn from the data given

B. Given is that sentence 2 is **true**.

- What does this tell you about the truth of **sentence 4**?

  ◉ sentence 4 is true
  ○ sentence 4 is false

- From the sentences given, is it possible to conclude whether there are large objects in the world?

  ○ yes, it is possible to conclude that
  ◉ no, it is not possible to conclude that

Problem 4

Given are two **true** sentences:

1. If there is a cube, then that cube is large.
2. If there is a tetrahedron, then that tetrahedron is small.

A. Given is that there is a cube.
   - Is this cube large?

     ◉ yes, this cube is large
     ○ no, this cube is not large

B. Given is that there are no small objects.
   - Is there a tetrahedron?

     ○ yes, there is a tetrahedron
     ◉ no, there is no tetrahedron

Part 2

Problem 1

Given is statement S1. Statement S1 is **true**.

S1: **If today is the first Friday of the month, then I go to the hairdresser's.**

Today is Wednesday. Is the conclusion that I am not going to the hairdresser's correct?

- ○ yes, this conclusion is correct
- ◉ no, this conclusion is not correct

Problem 2

Given is statement S2.

S2: **If I go on holiday to Italy, then I go by plane.**

This year I'm going to Spain and not to Italy. Is, in this case, statement S2 true, false or can't you know?

- ◉ Statement S2 is true
- ○ Statement S2 is false
- ○ You cannot know whether statement S2 is true or false

Problem 3

Given is statement S3.

S3: **All giraffes in this room can type.**

Is statement S3 true, false or can't you know?

- ◉ Statement S3 is true
- ○ Statement S3 is false
- ○ You cannot know whether statement S3 is true or false

Problem 4

Given is statement S4.

S4: **If I go to the city today, I will eat an ice-cream.**

I am not going to the city today, but I will go to the beach and I will eat an ice-cream there. Is statement S4 true, false or can't you know?

- ◉ Statement S4 is true
- ○ Statement S4 is false
- ○ You cannot know whether statement S4 is true or false

Problem 5

Given is statement S5.

S5: **If I eat a piece of cake, then I feel sick.**

I'm not eating a piece of cake and I'm feeling fine. Is statement S5 true, false or can't you know?

- ◉ Statement S5 is true
- ○ Statement S5 is false
- ○ You cannot know whether statement S5 is true or false

Problem 6

Given is statement S6. Statement S6 is **true**.

S6: **If I climb the Eiffeltower, I will suffer from fear of heights.**

I am climbing the Eiffeltower. Does this mean that I will suffer from fear of heights?

- ◉ yes
- ○ no

Problem 7

Given is statement S7.

S7: **If I go to bed, then I read a book.**

I'm reading a book. Can you conclude that I have gone to bed?

- ○ yes, you can conclude this
- ◉ no, you cannot conclude this

Problem 8

Given is statement S8. Statement S8 is **true**.

S8: **If my parakeet does not whistle, it is hungry.**

My parakeet isn't hungry. Can you conclude that it will whistle?

- ◉ yes, you can conclude this
- ○ no, you cannot conclude this

Problem 9

Given is statement S9.

S9: **If I sport, I will loose weight.**

I have lost 5 kilos. Is, in this case, statement S9 true, false or can't you know?

- ◉ true
- ○ false
- ○ you cannot know

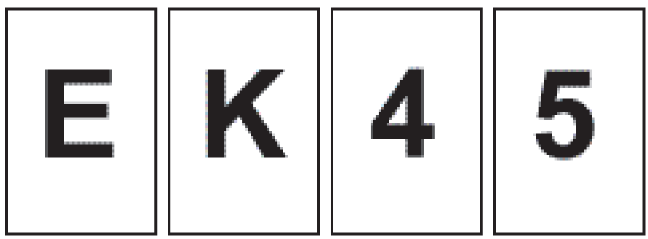
Part 3

Problem 1

Below are four cards. On each card there is always a *letter* on one side and a *number* on the other side. a card never contains two numbers or two letters.

Statement: **If there is an E on one side, then there is a 4 on the other side.**

Which card(s) do you have to turn over in order to decide whether the rule is true or false? Turn over only the card(s) necessary.



Problem 2

Imagine, you're in the supermarket and in front of you there are four jars. The front of each jar tells what the contents of that jar are; the price of the jar is put on the back. Two jars are standing back-to-front on the shelf. The flyer of the supermarket says:

40% discount on all jars of jam

However, you wonder whether this statement is true, because you have noticed a lot of errors in this supermarket. You decide to check this. Which jar(s) do you have to turn round in order to see whether the statement in the flyer is correct or not?

Problem 3

Imagine, you're a biologist and you're in doubt about the following statement:

If a bird on the island Tripon has purple blots under its wings, then that bird builds nests on the ground.

You are doing a small research on four individual birds: a, b, c and d. Which of these birds do you have to examine more closely to know whether this statement is true or not? Examine only the bird(s) necessary.

- ❑ Bird a (has purple blots under its wings)
- ❑ Bird b (does not have any purple blots under its wings)
- ❑ Bird c (builds nests on the ground)
- ❑ Bird d (builds nests in trees)

Problem 4

On a table are four objects. These objects are wrapped in either a *cylinder* or a *bag*. When a cylinder, you know that it contains a cube or a tetrahedron, but you do **not** know what size it is. When a bag, you do not know what is in it, but you **do** know what the size of the object is.

Attention! There are four objects on the table: a **cube** and a **tetrahedron** of which you do not know how small or big they are, and a **large** and a **small** object of which you do not know the shape.

Which of the four objects do you have to look at more closely to be able to tell with certainty that the following statement is true or false?

Statement: **All cubes are large.**