NEW PERSPECTIVES ON LEARNING AND INSTRUCTION

Use of Representations in Reasoning and Problem Solving

Analysis and improvement

Edited by Lieven Verschaffel, Erik De Corte, Ton de Jong and Jan Elen

Earli

ROUTLEDGE

Use of Representations in Reasoning and Problem Solving

Within an increasingly multimedia-focused society, the use of external representations in learning, teaching and communication has increased dramatically. Whether in the classroom, university or workplace, there is a growing requirement to use and interpret a large variety of external representational forms and tools for knowledge acquisition, problem solving and to communicate with others.

Use of Representations in Reasoning and Problem Solving brings together contributions from some of the world's leading researchers in educational and instructional psychology, instructional design and mathematics and science education to document the role that external representations play in our understanding, learning and communication. Traditional research has focused on the distinction between verbal and non-verbal representations, and the way they are processed, encoded and stored by different cognitive systems. The contributions here challenge these research findings and address the ambiguity about how these two cognitive systems interact, arguing that the classical distinction between textual and pictorial representations has become less prominent. The contributions in this book explore:

- how we can theorise the relationship between processing internal and external representations
- what perceptual and cognitive restraints can affect the use of external representations
- how individual differences affect the use of external representations
- how we can combine external representations to maximise their impact
- how we can adapt representational tools for individual differences.

Using empirical research findings to take a fresh look at the processes that take place when learning via external representations, this book is essential reading for all those undertaking postgraduate study and research in the fields of educational and instructional psychology, instructional design and mathematics and science education.

Lieven Verschaffel is Professor at the Faculty of Psychology and Educational Sciences, Katholieke Universiteit Leuven, Belgium.

Erik De Corte is Emeritus Professor of the Faculty of Psychology and Educational Sciences, Katholieke Universiteit Leuven, Belgium.

Ton de Jong is Professor of Educational Psychology, University of Twente, Netherlands.

Jan Elen is Professor at the Faculty of Psychology and Educational Sciences, Katholieke Universiteit Leuven, Belgium.

New Perspectives on Learning and Instruction

New Perspectives on Learning and Instruction is published by Routledge in conjunction with EARLI (European Association for Research on Learning and Instruction). This series publishes cutting-edge international research focusing on all aspects of learning and instruction in both traditional and non-traditional educational settings. Titles published within the series, take a broad and innovative approach to topical areas of research, are written by leading international researchers and are aimed at a research and postgraduate student audience.

Also available:

Transformation of Knowledge Through Classroom Interaction
Edited by Baruch Schwarz, Tommy Dreyfus and Rina Hershkowitz

Contemporary Perspectives on Reading and Spelling
Edited by Clare Wood and Vincent Connelly

Researching and Understanding Educational Networks
Edited by Robert McCormick, Alison Fox, Patrick Carmichael and Richard Proctor

Learning Across Sites: New Tools, Infrastructures and Practices
Edited by Sten Ludvigsen, Andreas Lund, Ingvill Rasmussen and Roger Säljö

Use of Representations in Reasoning and Problem Solving

Analysis and improvement

Edited by

Lieven Verschaffel, Erik De Corte, Ton de Jong and Jan Elen

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2010
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
270 Madison Avenue, New York, NY 10016

Routledge is an imprint of the Taylor & Francis Group, an informa business

This edition published in the Taylor & Francis e-Library, 2010.

To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Use of representations in reasoning and problem solving:
analysis and improvement/edited by Lieven Verschaffel [et al.].
1. Interaction analysis in education. 2. Thought and thinking.
3. Problem solving—Study and teaching. 4. Learning, Psychology of.
I. Verschaffel, Lieven, 1957-
LB1060.U84 2010
370.15′24—dc22 2010001812

ISBN 0-203-84782-2 Master e-book ISBN

ISBN13: 978-0-415-55673-6 (hbk)
ISBN13: 978-0-415-55674-3 (pbk)
ISBN13: 978-0-203-84782-4 (ebk)

Contents

Contributors

Ana Acevedo Nistal	University of Leuven, Belgium
Nicola Ariasi	University of Padova, Italy
Christiane Baadte	University of Koblenz-Landau, Germany
Geraldine Clarebout	University of Leuven, Belgium
Dirk De Bock	University of Leuven & Hogeschool-Universiteit Brussel, Belgium
Erik De Corte	University of Leuven, Belgium
Ton de Jong	University of Twente, The Netherlands
Michiel Doorman	Utrecht University, The Netherlands
Paul Drijvers	Utrecht University, The Netherlands
Jan Elen	University of Leuven, Belgium
Gijsbert Erkens	Utrecht University, The Netherlands
Tessa H. S. Eysink	University of Twente, The Netherlands
Sven Gemballa	University of Tübingen, Germany
Peter Gerjets	Knowledge Media Research Center, Tübingen, Germany
Koeno Gravemeijer	Technical University Eindhoven, The Netherlands
Brian Greer	Portland State University, U.S.A.
Birgit Imhof	Knowledge Media Research Center, Tübingen, Germany
Tomi Jaakkola	University of Turku, Finland
Jeroen Janssen	Utrecht University, The Netherlands
Liesbeth Kester	Open University of the Netherlands
Paul A. Kirschner	Open University of the Netherlands
Bas Kolloffel	University of Twente, The Netherlands
Tim Kühl	Knowledge Media Research Center, Tübingen, Germany
Erno Lehtinen	University of Turku, Finland
Lucia Mason	University of Padova, Italy
Andreas Müller	University of Koblenz-Landau, Germany
Sami Nurmi	University of Turku, Finland
Vanessa Pfeiffer	University of Tübingen, Germany

Renate Rasch	University of Koblenz-Landau, Germany
Catrin Rode	University of Oregon, U.S.A.
Katharina Scheiter	University of Tübingen, Germany
Michael Schneider	ETH Zürich, Switzerland
Wolfgang Schnotz	University of Koblenz-Landau, Germany
Elsbeth Stern	ETH Zürich, Switzerland
Xenia Vamvakoussi	National and Kapodistrian University of Athens, Greece
Wim Van Dooren	University of Leuven, Belgium
Jeroen J. G. van Merriënboer	Open University of the Netherlands & Maastricht University, The Netherlands
Lieven Verschaffel	University of Leuven, Belgium
Stella Vosniadou	National and Kapodistrian University of Athens, Greece
Sandra A. J. Wetzels	Open University of the Netherlands

Use of representations in reasoning and problem solving

An overview*

Lieven Verschaffel, Erik De Corte, Ton de Jong, and Jan Elen

Background and aims

As a result of recent developments in information and communication technology (ICT), the use of (external) representations in information processing, communicating, and learning and teaching has increased dramatically. Nowadays, learners must be able to interpret and use a large variety of (external) representational forms and tools both for their own reasoning, problem solving and learning, and for communicating with others.

Overviews of state-of-the-art research on the nature and use of (external) representations can be found in recent (hand)books, including *Technology-enhanced Learning: Principles and Products* edited by Balacheff et al. (2009), the third edition of the *Handbook of Research on Educational Communication and Technology* edited by Spector et al. (2007), and the *Cambridge Handbook of Multimedia Learning* edited by Mayer (2005).

In the first and second of these books the issue of (external) representations is essentially addressed as one element of educational information and communication technology, along with many other issues such as the impact of various kinds of educational software (e.g., drill-and-practice programs) on students' learning, teachers' beliefs about and attitudes towards various kinds of educational technology, and the implementation of new forms of information and communication technology in educational settings. Mayer's (2005) handbook focuses much more on the (educational) use of (external) representations, but this topic is approached from a number of theoretical perspectives that have been typically developed and used within the field of multimedia learning, especially cognitive load theory (Sweller, 1999) and theories of dual coding (Paivio, 1990). Furthermore, given Mayer's specific definition of multimedia learning

*This publication project was supported by Grant GOA 2006/01 'Developing adaptive expertise in mathematics education' from the Research Fund of the Katholieke Universiteit Leuven, Belgium and by the International Scientific Network WO.002.07N on 'Stimulating Critical and Flexible Thinking' of the Research Foundation – Flanders.

and instruction, the book has a strong emphasis on the issue of textual versus pictorial information.

In one chapter in Mayer's book, Atkinson (2005) notices that the 'subject-matter perspective' is typically missing in current multimedia research, and therefore makes a plea for more research on multimedia learning that is deeply embedded in specific curricular domains such as mathematics or science. In a more recent reflection on the state-of-the-art in multimedia research, Mayer (2010) comments that this research has to address three kinds of questions, namely 'what works?', 'when does it work?', and 'how does it work?'; but he also argues that so far research has principally addressed the first two questions and that therefore more work is needed on the analysis and description of the perceptual, thinking, and learning processes that underlie the effectiveness of multimedia materials and techniques.

Compared with the above-mentioned volumes, the present book puts a stronger emphasis on the issue of (external) representations *as such*, paying ample attention to the similarities and differences between various kinds of (external) representations and to the relationship between external and internal representations. Furthermore, it looks at this representational issue not only from the above-mentioned theoretical perspectives that have been typically applied within the context of multimedia learning, but also from other theoretical perspectives, such as general theories of problem solving and conceptual change, or domain-specific theories of mathematics and science learning.

By strictly focusing on (external) representations and by including these additional theoretical perspectives, we have a dual aim. First, we aim to contribute to a better understanding of how representational forms and tools can – either alone or in combination with others – foster or hinder thinking and learning processes in particular subject-matter domains and instructional settings, that is, the third question emphasised by Mayer (2010). Secondly, we intend to explore how these findings on the relations between (external) representations, the associated thinking and learning processes, and the learning outcomes can be translated into effective and efficient instructional guidelines and methods.

In line with this dual goal, Part 1 addresses the *analysis* of psychological processes involved in working with (external) representations when reasoning and solving problems, and Part 2 the *development* of external representational tools and learning environments aimed at the enhancement of the intended reasoning and problem-solving processes.

The book grew out an international workshop held in Leuven on September 9–12, 2008, organised by the international scientific network 'Stimulating critical and flexible thinking', sponsored by the Research Foundation – Flanders.

Brief overview of the book

Part 1 commences with a chapter in which Schnotz et al. analyse the interplay of external and internal representations in creative thinking and problem solving from the perspective of semiotics and cognitive psychology, using examples from science and mathematics education. Creative thinking and problem solving are analysed from the perspective of Gestalt psychology and the psychology of information processing, emphasising the roles of structures and procedures. The authors make a distinction between two basic kinds of representations, namely descriptive and depictive, which differ in both representational and inferential power. Their analysis of the use of representations in contexts of science and maths education shows that a close interaction between descriptive and depictive representations is needed in order to make the best use of both kinds of representation for successful thinking and problem solving.

In Chapter 2, Vosniadou argues that research on the comprehension of text and pictures has failed to consider an important distinction between pictures that are perceptually based depictions, on the one hand, and those that represent conceptual models, on the other hand. She addresses differences between these two kinds of external representations and presents some of the difficulties students have when faced with conceptual models. These difficulties arise, she argues, because understanding a conceptual model is an interpretive process that can be seriously hampered by students' lack of essential domain-specific knowledge and realistic epistemic beliefs. Vosniadou concludes with some recommendations about how pictures representing conceptual models can be helpful in the teaching of science and mathematics.

The next chapter by Mason and Ariasi addresses the role of external representations in reasoning by examining the epistemic processing of texts and pictures about a biology topic presented on multiple Internet pages. Based on the objective measurement of visual attention through eye-fixations, their study revealed indirect evidence of epistemic processing, that is, processing that takes into account the source, reliability, and accuracy of the informational content. University students were asked to read four web pages differing in authoritativeness, which provided various types of information. The findings showed that participants allocated different amounts of visual attention to different texts and pictures, *within* and *across* web pages, according to source credibility. In addition, students' individual differences regarding prior knowledge, epistemic beliefs, and argumentative reasoning played to some extent a role in epistemic processing.

In Chapter 4, Acevedo Nistal et al. report a study in which they examined students' ability to make adaptive or flexible representational choices while solving linear-function problems. Two secondary school classes solved problems under a choice condition, where they could choose a table, a graph, or a formula to solve each problem, and three no-choice conditions where a predetermined representation (respectively a table, a graph, or a formula) had to be used. Data concerning representational efficiency (extracted from the no-choice conditions) and frequency of representational choice (extracted from the choice condition) were analysed. Students' representational flexibility was assessed using two conceptualisations of flexibility. In a purely task-based conceptualisation choices were considered flexible if they took into account only task characteristics. In a task × student conceptualisation, student characteristics were also brought into the equation. The results obtained from the two approaches are compared.

Chapter 5 by Greer, De Bock, and Van Dooren discusses the role of (external) representations in mathematical proof using the 'Isis problem' as a central case. The Isis problem asks: 'Find which rectangles with sides of integral length (in some unit) have area and perimeter (numerically) equal, and prove the result.' The problem is notable for the variety of proofs available (empirically grounded, algebraic, geometrical) and the associated representations; moreover, it provides an instrument for probing students' ideas about proof. First, the authors set out a variety of approaches leading to proofs, showing thereby how proofs can rely on substantially different mathematical representations, each having its affordances differentially clarifying particular aspects of the mathematical situation. They also argue that being involved in making transitions from one representation to another, and linking various representations can provide deeper insight. In the second part of the chapter, they discuss a study with nine American and 39 Flemish future mathematics teachers who first attempted to solve the problem, then studied five given proofs and commented upon them. The results highlight a preference of the more mathematically advanced students for algebraic proofs over empirically grounded and visual (geometric) proofs.

In the next chapter, Schneider, Rode, and Stern address the availability and activation of diagrammatic strategies for learning from texts in secondary school students. The authors' starting point is that diagrams are powerful tools for learning and reasoning, and that people frequently do not use diagrams even in situations in which they would be very helpful. In two experiments they investigated whether the reason for this is either a lack of *availability* or a lack of *activation* of diagrammatic representation strategies. A group of seventh graders and ninth graders read texts which could be summarised by a diagram as well as by keywords. Students were asked to take down notes. The experimental conditions differed as to whether the instructions for note taking explained the diagrammatic strategy or whether they explicitly requested its use. Results revealed that neither availability nor activation was well developed in students. Instructions aimed at increasing availability or activation led to increased diagram use, better memorising of facts, and better inferences. Spontaneous diagram use improved with grade level, but still remained insufficient even in grade nine. The authors argue that instruction should encourage students to use diagrams based on their specific advantages.

Part 2 begins with a chapter by Jaakkola, Nurmi, and Lehtinen who investigated, using video data, the simultaneous use of a computer simulation and real electrical circuits (a hybrid environment). The central question is why simultaneous use in a hybrid environment promotes students' conceptual understanding of electrical circuits more effectively than the use of the simulation alone. Elementary school students learnt about electrical circuits in a simulation-alone or a hybrid condition. No differences were found in the amount of cognitive conflicts and self-explanations between the two conditions. However, the video data transcripts from the hybrid environment suggested that analogical encoding of two information resources can improve schema abstraction and deepen students' conceptual understanding of electrical circuits. The authors conclude that, overall, it seems to be beneficial to try to promote students' conceptual understanding of electrical circuits at the early elementary school level, because they do not yet have deeply rooted misconceptions that could hamper teaching and learning.

Gerjets et al. provide in Chapter 8 an overview of four studies that compared static and dynamic visualisations in the context of the biological domain of fish locomotion. The different learning objectives addressed in these studies comprise: (1) understanding the physical principles underlying fish locomotion, (2) classifying different fish locomotion patterns, and (3) identifying different fish species based on important static and dynamic features. The results demonstrate that for all three learning objectives dynamic visualisations were superior to static ones. These findings were obtained in laboratory settings as well as in the highly situated learning scenario of using mobile devices during a snorkelling excursion. The authors conclude that their results clearly yield the recommendation to use dynamic instructional visualisations instead of static ones for supporting the comprehension of complex dynamic phenomena in the natural sciences.

The next chapter, by Kolloffel, Eysink, and de Jong, reports two studies in which the effects of external representations on learning combinatorics and probability theory in an inquiry-based learning environment were investigated. In the first study, the effects of five representational formats used to present the domain to students were compared: Tree diagram, Arithmetic, Text, Text + Arithmetic, or Tree diagram + Arithmetic. The main finding was that students in the Text + Arithmetic condition obtained the best learning results. Tree diagrams were found to negatively affect learning and to increase cognitive load. The second study examined the effects of providing support tools students could use to construct domain representations. Three formats of such tools (conceptual, arithmetical, or textual) were compared, both in an individual and collaborative learning setting. Format influenced students' inclination to use a tool, with arithmetical representation being the least popular among the students. Furthermore, the collaborative students obtained better learning outcomes, but if individuals used the support, their learning outcomes equalled those of collaborating students.

In Chapter 10, Gravemeijer, Doorman, and Drijvers argue that the problematic character of symbolic representations in mathematics education is tied to what is called the 'dual nature of mathematics' – which is procedural as well as structural. Historically, procedural conceptions precede structural conceptions, whereas mathematics education often starts at a structural level, using concrete representations to introduce those structural conceptions. According to the authors, this is problematic because these representations derive their meaning from structural conceptions that the students still have to appropriate. In the alternative they propose, bottom-up learning processes in which symbols and meaning co-evolve are fostered. They elaborate such an approach known as the 'emergent modelling instructional design heuristic' – for the topic of algebraic functions. A brief sketch of a teaching experiment on early algebra elucidates this alternative and suggests that information technology can actually support the transition from a procedural to a structural conception of functions.

In Chapter 11, Vamvakoussi reviews a series of studies investigating secondary students' understanding of the density of numbers, and attempting to bring this notion within the grasp of students. She presents empirical evidence demonstrating the adverse effect of the multiple symbolic representations of rational numbers, as well as the limited, sometimes adverse, effect of the number line, on students' judgements on the number of numbers in an interval. Vamvakoussi argues that students' difficulty with the notion of density relates to a more general problem of conceptual change in the development of the number concept. Cross-domain mapping between number and the line is proposed as a mechanism that could facilitate the restructuring of students' number concept. This claim is supported by empirical evidence showing that the number line, a representation grounded on the 'numbers are points' analogy, can facilitate students' understanding of density if purposefully employed in instruction.

Next, Wetzels, Kester, and van Merriënboer outline a theoretical framework providing insights into the use of external representations of low sophistication during prior knowledge activation in the science domain. This framework distinguishes representations that *prompt* (i.e., initiate) prior knowledge activation from representations that *reinforce* (i.e., facilitate) the activation process. Prompts that consist of pictorial representations (e.g., pictures, animations) are regarded as more suitable than verbal representations for activating structural and causal models important for science learning. Furthermore, external representations may reinforce the activation process. There are limits to the amount of information that can be activated simultaneously because of humans' limited working memory capacity. Self-constructed representations (e.g., note taking) might offload working memory while activating prior knowledge. It is argued that the strength of the prompting and reinforcing effects of external representations during prior knowledge activation is mediated by learners' level of prior knowledge. An empirical study that provides support for the framework is reported.

The final chapter explores the visualisation of argumentation as a shared activity. Erkens, Janssen, and Kirschner's starting point is that the use of argumentation maps in computer-supported collaborative learning does not always provide students with the intended support for their collaboration. They compare two argumentation maps from two research projects, both meant to support the collaborative writing of argumentative essays based on external sources. In the COSAR-project, students could use a Diagram-tool to specify positions, pro-arguments, con-arguments, supports, refutations and conclusions in a *free* graphical format to write a social studies essay. The tool was highly appreciated by students and teachers, but did not result in better essays. In the CRoCiCL-project, a Debate-tool allowed students to do the same things but in a *structured* graphical format, meant to visualise the argumentative strength of the positions. This resulted in better history essays. The difference in representational guidance between the tools might explain these distinct effects, with the Debate-tool stimulating students to attend to the justification of positions and their strengths. Implications for research and instruction are discussed at the chapter end.

Acknowledgements

Finally, we express our sincere thanks to all those who have contributed to the production of this book.

First, we thank all the authors for agreeing so enthusiastically to contribute a chapter, for the smooth cooperation during all stages of the production process, and for the high quality of their final products.

We also express our thanks to the editor-in-chief and the other members of the editorial team of the new EARLI book series New Perspectives on Learning and Instruction for the encouragement and support during the production process, as well as to the following colleagues who acted as reviewer of one of the book chapters: A. Renkl, C. Selter, I. Elia, J. Ainsworth, J. Deprez, J. Star, M. Crahay, M. Ebersbach, M. Valcke, T. Schellens, T. Seufert, T. Van Gog, Z. Mevarech, and Z. Zacharia.

Furthermore, we are grateful to Mrs B. Van den Bavière and Mr M. Vlecken for their editorial assistance in the preparation of the final book manuscript and with the correction of the galley proofs.

We are also indebted to the Research Foundation – Flanders and the Research Fund of the Katholieke Universiteit Leuven, Belgium. Their financial support allowed us to organise the workshop on which this book is based and to prepare the final manuscript.

References

Atkinson, R. K. (2005). Multimedia learning of mathematics. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 393–408). Cambridge, UK: Cambridge University Press.

Balacheff, N., Ludvigsen, S. R., de Jong, T., Barnes, S., & Lazonder, A. W. (Eds.). (2009). *Technology-enhanced learning: principles and products.* Berlin: Springer Verlag.

Mayer, R. E. (Ed.). (2005). *Cambridge handbook of multimedia learning.* Cambridge, UK: Cambridge University Press.

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics [Commentary]. *Learning and Instruction*, 20(2), 167–171.

Paivio, A. (1990). *Mental representations: A dual coding approach.* New York: Oxford University Press.

Spector, J. M., Merrill, M. D., van Merriënboer, J. J. G., & Driscoll, M. P. (Eds). (2007). *Handbook of research on educational communication and technology* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Sweller, J. (1999). *Instructional design in technical areas.* Camberwell, UK: ACER Press.

Part 1

Theoretical and empirical analyses of psychological processes in thinking and learning with representations

Chapter 1

Creative thinking and problem solving with depictive and descriptive representations

Wolfgang Schnotz, Christiane Baadte, Andreas Müller, and Renate Rasch

Introduction

Creative thinking and problem solving are higher-order epistemological activities of humans, which play also a major role as educational objectives in our schooling system. We consider thinking and problem solving as creative if they go beyond the reproductive application of previously learned routines. International assessment studies in learning mathematics and science education such as TIMSS or PISA have emphasised the relevance of flexible thinking and problem solving for education and for living in a complex society. We consider flexibility in thinking and problem solving as the amount of different views and approaches an individual can create to solve a task or a problem. Although the relevance of flexibility in thinking and problem solving is widely acknowledged, not very much is known yet about the nature of this flexibility and how it may be improved. We will analyse in this chapter the role of representations in flexible thinking and problem solving. More specifically, we will investigate how the construction, manipulation and usage of different forms of external and internal (i.e., mental) representations, and the interaction of these representations, can contribute to these higher-order cognitive processes.

First, we will describe different views on creative thinking and problem solving in psychology. Second, we will analyse more deeply two basic kinds of representations, which we will call descriptive and depictive representations. Third, we will investigate the functionality of these different kinds of representations in terms of their representational and inferential attributes. Fourth, we will focus on the interplay between external representations and internal (mental) representations in thinking and problem solving and we will differentiate between different kinds of mental representations. Fifth, we will analyse the role of different representations in thinking and problem solving, with examples from science education and mathematics. Finally, we will draw some conclusions for the process of teaching in order to foster creative thinking and problem solving.

Different views on thinking and problem solving

Gestalt psychology: finding the right representation

Among the different theoretical schools in psychology about productive thinking and problem solving, the Gestalt psychology and the psychology of information processing were (and probably still are) the most influential ones. In Gestalt psychology, the main requirement of problem solving was considered to be finding the right representation of the problem. In psychology of information processing, the main requirement of problem solving was considered to be finding the right way through a complex problem space. We will briefly elaborate on these different views of problem solving in the following.

Conforming with the great attention that it devoted to the study of visual perception, Gestalt psychology also considered problem solving mainly as a matter of perception (see Wertheimer, 1938). We know that perception is often constrained by experience. For example, repeated experience of perceiving an object in one way, such as a pair of tongs as a tool for handcraft, may hinder its perception in another way, such as seeing the pair of tongs simply as a possible ballast (see Maier, 1931). Gestalt psychologists coined the term 'functional fixedness' for this phenomenon (Duncker, 1935). According to this view, problem solving takes place when the perception of the situation is suddenly reorganised. The new perception (provided it is the 'right' one) makes the solution immediately obvious or, in other words, the solution can be read off immediately from the new perception. Accordingly, problem solving is something that takes place all of a sudden. Contrary to a behaviourist view of solving problems through a gradual approximation of a goal via a process of trial and error, Gestalt psychology considered problem solving as a matter of sudden insight, often accompanied by a subjective 'aha experience' (Bühler, 1907).

Numerous examples have been cited in the literature to illustrate this view on problem solving. One of them is the so-called 'cheap necklace problem' (Silveira, 1971). Problem solvers are presented with four separate pieces of chain (A–D) each consisting of three connected links (see Figure 1.1, left side). These four chains form the initial state of the problem. The goal state (illustrated on the right side of Figure 1.1) is a chain that consists of a total of 12 interconnected links.

Additionally, the following instruction is presented: 'You are given four separate pieces of chain that are each three chains in length. It costs 2¢ to open a link and 3¢ to close a link. All links are closed at the beginning of the problem. Your goal is to join all 12 links of chain into a single circle at a cost of no more than 15¢' (Silveira, 1971).

According to the figure and the instruction, one obvious representation of the goal state is the one that directly refers to the depiction of the four separate chains, each of which consists of three interconnected links. That is, the problem

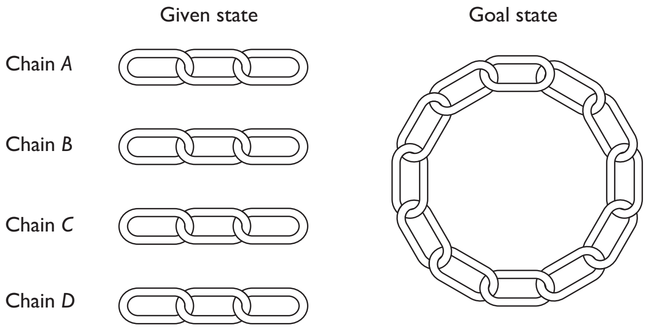


Figure 1.1 The cheap necklace problem.

solver is likely to construct a representation of the given state in which each of the four chains is considered as a distinct entity. As a result of this specific representation, long-term knowledge is activated that provides the problem solver with information (operators) on how to connect the four single chains, such as open link 1 of chain A, connect it with link 3 of chain B and close the link, then open link 1 of chain B, connect it with link 3 of chain C and close the link, etc. Note that in applying such a strategy, the four chains, as the original entities, are still maintained during the process. The problem solver may at some point become aware that this strategy does not actually meet the constraints of the problem situation because the achievement of the goal state would exceed the total cost of 15¢ as incorporating the four chains into one closed chain comprises four openings (8¢) and four closings (12¢) of the respective links, which results in a total cost of 20¢. One of these approaches to solving the cheap necklace problem is to disconnect one of the four short chains into three single links that can subsequently be used to interconnect the remaining three chains. This procedure requires the problem solver to open and to close three links respectively leading to a total cost of 15¢ which meets exactly the monetary constraint of the problem situation.

In other words, a perception of the cheap necklace problem as four chains to be interconnected leaves the problem unsolvable. However, if the individual manages to find the following new perception of the problem, its solution becomes straightforward. One chain (say chain D) can simply be seen as a set of three elements that can be used to connect the remaining three chains (A, B and C) instead of seeing D as a chain that has to remain intact. The problem is perceived in a new way, which means that a new representation of the given state has been created. To summarise: in the Gestalt psychological tradition problem solving is predominantly regarded as a matter of a sudden perceptual reorganisation, namely, seeing the problem situation from a new perspective. If the right representation of the problem has been found, the solution becomes obvious immediately. Finding a new representation, that is, finding a new perception of the problem, basically means a restructuring of the problem representation. For example, in finding the solution of a problem, the problem solver might suddenly become aware of new relations between elements of the given material by mentally changing, amplifying or restructuring the material (Montgomery, 1988). The outcome of such reorganisation is an altered mental representation of the problem which often enables the problem solver to directly read off the correct solution. As this solution apparently becomes immediately obvious, the Gestalt psychologists assumed that the problem solver suddenly gains a deep insight into the problem, often accompanied by the so-called 'aha experience'.

Psychology of information processing: finding the right path

In contrast to Gestalt psychology, the information processing approach postulates that problem solving is a matter of searching for a path from one location to another location in a so-called problem space. This space is constituted by the total set of possible states that can be created during working on the problem (represented by locations in the space) and the operations that transfer one state into another state (represented by connections between locations). According to this approach, problem solving has at least four prerequisites: first the problem solver must have a representation of the initial state at his/her disposal. Second, he/she should have a rather concrete idea about the goal state. Third, he/she must be able to set up a space (at least a partial one at a specific time) of the problem at hand. Fourth, the problem solver should have some operators available that can be applied to the initial state in order to transform it stepwise into some intermediate states until the goal state is reached. Heuristics such as reducing the difference to the goal state (so-called hill climbing), means-end analysis or avoidance of loops can help the problem solver to find his/her way through the problem space, that is, to decide what might be the best next step.

A frequently used example to illustrate this approach is the so-called Tower of Hanoi (see Figure 1.2). Within a system of three rods, a set of discs of different size have to be transferred from one rod (the given state) to a specific other rod (the goal state). The operation rules are as follows: only one disc can be moved at any time; a larger disc can never be put onto a smaller disc. The problem space increases when more discs are introduced: with 2 discs, the space includes 9 states; with 3 discs, it includes 27 states; with 4 discs, it includes 63 states, and so forth.[1] Accordingly, with an increasing number of discs the transformation of the given state into the goal state quickly becomes a non-trivial task which clearly falls into the category of problem solving. Good problem solving means finding a short way rather than a long-winded way.

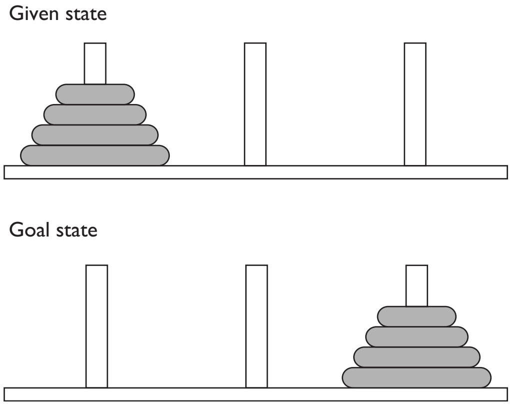


Figure 1.2 Example of a Tower of Hanoi problem.

Synthesis: combining structure and process

At first glance the approach of Gestalt psychology and the approach of psychology of information processing seem to be based on fundamentally different theoretical concepts of human thinking and problem solving. As Ohlsson (1984a, 1984b) delineates, the central concept of the Gestalt theory is restructuring that leads to insight in the problem and hence to its solution. In contrast, the central concept of the information processing approach is the search for possible paths that lead from the initial state to the goal state.

However, on closer inspection, these two approaches do not necessarily act on different assumptions concerning the topic of problem solving. In fact, they merely focus on different aspects of the problem-solving process: whereas Gestalt theory emphasises the significance of the problem's 'good' structure for the solution, information processing theory highlights the procedures that are applied to the material of the initial state in order to achieve specific intermediate states that finally result in the goal state. Hence, Ohlsson (1984b, 1992) proposes a theory that consolidates both perspectives in that he attempts to incorporate some of the Gestalt theory's central assumptions into the information processing approach.

Ohlsson (1992) delineates in his Representational Change Theory some core assumptions that interlink both theoretical approaches by specifying the processes that lead to insight and thus to the solution of the problem. A central aspect of Ohlsson's theory deals with the question of how the problem to be solved (the initial state) is represented and structured in the mind of the problem solver. As the representation is supposed to serve as a memory probe that activates long-term knowledge via spreading activation, a 'valid' representation of the initial state leads to the activation of knowledge (operators or actions) that is relevant for the solution of the problem, whereas an 'invalid' representation of the given situation activates knowledge from long-term memory that is not beneficial for the solution of the problem. Hence, the problem solver has to strive for the construction of a 'valid' representation by changing the 'invalid' representation in order to solve the given problem. According to Ohlsson (1992) there are several ways to change a representation, such as elaborating or adding new problem information to the current representation, relaxing constraints in order to permit new aspects to unfold or re-encoding, which leads to the reinterpretation of the problem at hand. As a consequence of the reinterpretation, new long-term knowledge structures can be activated including operators that may be supportive for goal achievement.

We will illustrate this with the cheap necklace problem described above (Silveira, 1971). Remember that the problem solver might first consider connecting the four single chains by opening link 1 of chain A, connecting it with link 3 of chain B and closing the link. Similarly, he/she might open link 1 of chain B, connect it with link 3 of chain C, and close the link, etc. Then, he/she might become aware that this strategy requires four openings and four closings and therefore exceeds the total cost of 15¢. In this case, the path to the goal state seems to be blocked. According to Ohlsson's Representational Change Theory (1992), such an impasse or block can be overridden if the problem solver successfully manages to alter the representation of the problem situation (the initial state). One way of doing so is a constraint relaxation in which inhibitions on what is regarded as an acceptable solution are removed from the representation of the problem. In the case of the 'cheap necklace problem' such a situational constraint may be the assumption that each of the four chains has to be maintained as an entity and be integrated as an ensemble into the necklace. This misinterpretation of the problem situation basically resembles the 'functional fixedness' described by Duncker (1935) which prevents us from solving a problem because of our knowledge about specific characteristics of some elements in the problem situation. Constraint relaxation in the 'cheap necklace problem' is achieved if the problem solver abandons the assumption that each of the four short chains must be maintained as an ensemble. This reinterpretation of the problem situation allows disconnecting one of the four short chains into three single links that can subsequently be used to interconnect the remaining three chains.

Generally speaking, problem solving requires representations with an adequate structure as well as processes that operate on this structure. Gestalt psychology focused primarily on structures, whereas psychology of information processing focused primarily on processes. Whereas Gestalt psychology did not sufficiently take into account that insight problems do also have a problem space and that they do also require operations within this space, psychology of information processing did not sufficiently take into account that problem solving does sometimes not only require an individual to find a path through a given problem space, but also to transform the given problem space into another problem space (i.e., restructuring the problem representation), which would enable an easier sequencing of operations.

Both aspects can again be illustrated with the cheap necklace problem. As long as the individual views the problem as a set of four chains, which should be interconnected one after the other, the problem space is structured in a way that requires four openings and four closure operations, and therefore violates the financial constraints that the total cost must not exceed 15¢ (see Figure 1.3a). When the unnecessary constraint that all the four chains should remain intact has been given up, a new problem space with a new structure emerges. In this space, the initial state of four chains is transformed into a new state, in which one chain is totally disassembled by three openings. From this intermediate state, only three closure operations are required to interconnect the remaining three chains into a full necklace, which also satisfies the financial constraints (see Figure 1.3b). It should be noticed that although the cheap necklace problem is considered an insight problem, it is nevertheless associated with a problem space. Insight corresponds to a restructuring (or amendment) of this space, and this restructuring provides the key to the solution. However, besides the restructuring, there is still a sequence of processes required to transform the given state

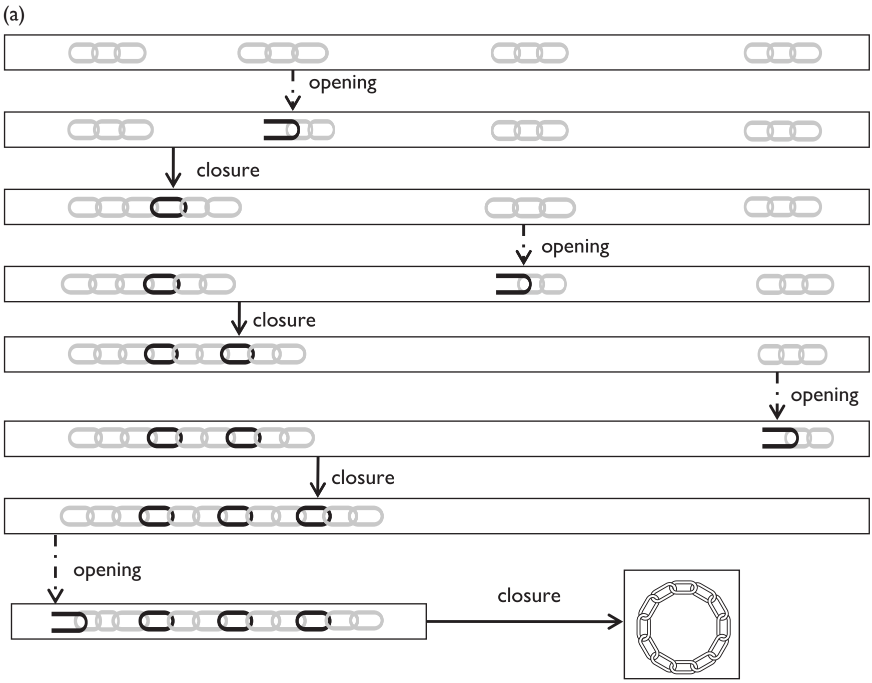


Figure 1.3 Problem spaces of the cheap necklace problem.

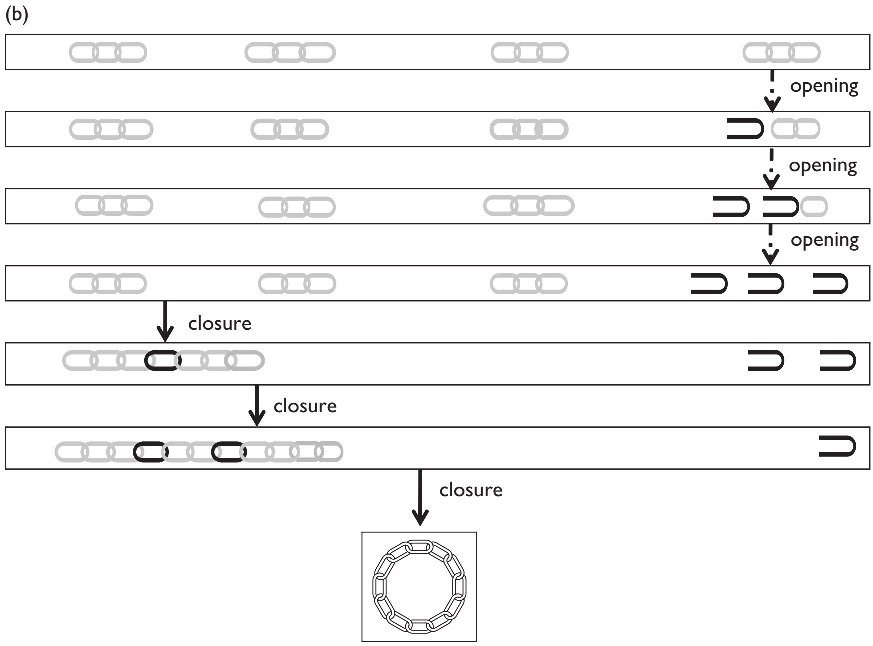


Figure 1.3 Problem spaces of the cheap necklace problem (Continued)

into the goal state. In other words, problem solving does not only require an adequate perception of the problem (which means a problem space structured as simply as possible for the problem at hand). It also requires operations to be performed on this structure.

Problem solving is thus a double-sided process. One side, which was emphasised by Gestalt psychology, is finding or constructing the right representation (i.e., a representation with the right structure). The other side, which was emphasised by psychology of information processing, is operating adequately on this structure by performing the right processes. Accordingly, good problem solving implies finding or creating a representation that enables easy performance of a sequence of operations that transforms a given state into a goal state.

The semiotics of representations

What do we mean by the term 'representations'? We consider a representation of an object or an event that stands for something else. Texts on paper describing *something*, sculptures or pictures of *something* are examples of objects that have a representational function. Stage plays, movies or ceremonies (such as baptism) are examples of events that have a representational function (see

Petersen, 1996). As a representation stands for something else, it refers to it and thus adopts the function of a sign, too. Representations can therefore also be considered as signs and vice versa.

Intentionally produced signs are generally used for communication (see Bühler, 1934). A series of utterances, a written letter, a painting, a photograph, for example, can be produced by a speaker, author, painter or photographer in order to convey knowledge about something (i.e., the represented content) to a recipient. If a sign producer (i.e., a speaker, author, painter or photographer) produces a sign, he/she means something that he/she wants to communicate. In other words, he/she has something in mind about a content to be represented, and he/she expresses his/her view about this content by producing the sign. Hence, meaning is a process that creates an external sign on the basis of what the sign producer knows, that is, what he/she has mentally represented about the content (i.e., the referent of the sign). The recipient of the sign can use it to construct a mental representation of the referent of the sign. In this case, the sign is used in order to reconstruct the knowledge about the represented content that has previously been externalised by the producer of the sign. In other words: the recipient comprehends the sign. Comprehension of signs is therefore a process that creates a mental representation (or knowledge structure) on the basis of the external sign. In successful communication, the sign recipient reconstructs in his/her mind what the sign producer has meant (Hörmann, 1976).

However, intentionally produced signs are not only used for communication with someone else. They can also be produced as external cognitive tools for the sign producer, who uses the external sign to infer afterwards new information about the represented content. This use of external signs can be considered as communicating with oneself: the individual alternately takes the role of the sign producer and the sign recipient by, for example, writing a text and then reading it, or creating a graphic and then observing it, with the aim of gaining new information. Instead of performing all cognitive processes in his/her mind, the individual creates an external representation, which has the advantage of higher stability than the transient mental representations in working memory (Baddeley, 1986). In other words, by creating an external representation, the individual offloads some of the requirements of problem solving to the environment, namely the cognitive load of maintaining a complex mental representation in working memory (Schnotz & Kürschner, 2007; Sweller, 2005; Sweller, van Merrienboer, & Paas, 1998).

Intentionally produced signs can be categorised into two main classes: symbols and icons. The distinction between the two categories of signs was mainly introduced by Peirce (1931). According to Peirce, symbols have an arbitrary structure and are associated with the designated object by a convention. Words and sentences of natural language are examples of symbols. Icons, on the contrary, do not have an arbitrary structure, but are associated with the designated object by similarity (i.e., a form of concrete analogy, in which, besides a structural commonality, the represented attributes and representing attributes are the same) or by a more abstract kind of analogy (i.e., a structural commonality, in which the represented attributes and representing attributes are different).

Despite numerous variants of representations, one can distinguish only two basic forms of representations: descriptions and depictions. A description represents a subject matter with the help of symbols. In a description of an object by natural language, components are referred to by nouns, are specified by adjectives according to their attributes and set in relation to each other with the help of verbs and prepositions. Besides sentences of natural language, there are also other kinds of descriptive representations. Mathematical expressions such as $V = s^3$ (describing the relation between a cube's size and its volume) or the formula $F = m \times a$ in physics (describing the relation between force, mass and acceleration according to Newton's Second Law) are also descriptive representations. Descriptive representations consist of symbols, that is, signs that have no similarity with their referent.

A depiction, on the contrary, is a spatial configuration (i.e., a set of points in a functional space) that represents a configuration in another space (Kosslyn, 1994). In other words: it presents a subject matter with the help of structural commonalities between the respective configurations. Pictures such as photographs, drawings, paintings and maps are depictive representations. However, there are also other kinds of depictive representations. Miniature models of a building or line graphs are also depictive representations. Depictive representations consist of icons, that is, signs that are associated with their referent by similarity or by another structural commonality (i.e., analogy). Depictions do not describe, but rather show the characteristics of an object.

The representational and inferential power of descriptions and depictions

Descriptive representations and depictive representations have different uses for different purposes. Descriptive representations are relatively general and abstract, whereas depictive representations are more concrete and specific. On the one hand, descriptions are therefore representationally more powerful than depictions. There is no problem in formulating general negations or disjunctions by descriptions as, for example, 'Pets are not allowed' or 'High blood pressure can be caused by nicotine or a lack of movement'. Depictive representations, on the contrary, can only show specific negations (with a specific pet), and they can only show disjunctions with the help of several pictures. Therefore, descriptive representations are more powerful in expressing abstract knowledge.

Depictive representations are always complete with regard to a specific class of information, whereas descriptive representations are more selective. If we draw an object, for example, we draw not only its shape, but, necessarily, also its size and orientation. In a description, on the contrary, one can specify the form of an object without mentioning its size or orientation (Kosslyn, 1994). A map, for example, includes all geometric information of the depicted geographical area, and a picture of a bird of prey eating a mouse does not only include information about the shape of the bird and the shape of the mouse, but, necessarily, also about their size, about their orientation in space, how the bird holds its prey, etc. Depictive representations are therefore more useful to draw inferences. They have a high computational efficiency, because the new information can be read off directly from the representation (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991; Kosslyn, 1994).

The inferential power of depictive representations has already been referred to implicitly by Duncker (1935) in his analysis of productive thinking, when he distinguished between analytic and synthetic reading. A geometric figure such as a triangle, for example, can be described with specific attributes. Let us assume that length of side A is 62 cm, length of side B is 70 cm, and the size of angle Gamma is 30°. With this description, the triangle is fully determined and can be constructed unambiguously. After construction has taken place, the triangle has all the attributes of a triangle, and all these attributes have specific values. With a ruler or a protractor, a user can identify all these values: the length of side C, the size of angles Alpha and Beta, the perimeter of the triangle and, with scaled paper, even the area of the figure. All these values can be identified just by reading from the figure without any computations. In the previous example, reading the length of side A as 62 cm, length of side B as 70 cm, and the size of angle Gamma as 30° would be possible, but it would also be trivial, because this information was known from the beginning. This is what Duncker called 'analytic reading'. Reading the length of side C, the size of angles Alpha and Beta or the perimeter, however, would lead to new information that has been derived from the representation. This is what Duncker called 'synthetic reading'. In other words: synthetic reading is exactly what makes depictive representations inferentially powerful.

Descriptive representations are especially useful for characterising objects and scenarios at higher levels of abstraction, for the explanation or prediction of events, and for directing the individual's attention and cognitive processing. Depictive representations are especially useful for envisioning the appearance of objects, for comprehension of scenarios and events, for reasoning, arguing and problem solving. The construction and manipulation of mental depictive representations – so called mental models – has been described by Johnson-Laird (1983; see also Johnson-Laird & Byrne, 1991) as the core of conditional syllogistic reasoning and reasoning about categories.

External and internal representations in comprehension, thinking and problem solving

Cognitive psychology considers comprehension as a process of constructing internal (mental) representations (Kintsch, 1998). If comprehension takes place as part of communication, it is based on signs such as texts or pictures, which have been written or drawn by someone. In the latter case, comprehension is a process of constructing internal representations on the basis of intentionally created external representations. The distinction between descriptive and depictive representations mentioned above applies also to the internal (mental) representations constructed during comprehension. A mental representation of the surface structure of a text, for example, and a propositional representation of the text's semantic content are descriptive representations, as they use symbols to describe the subject matter. A visual image and a mental model, on the contrary, are depictive representations, as they are assumed to have an inherent structure that corresponds to the structure of the subject matter (Johnson-Laird, 1983; Kosslyn, 1994).

The distinction between descriptive and depictive external and internal representations is at the core of a theoretical framework developed by Schnotz and Bannert (2003) for analysing text and picture comprehension, which is usually referred to as the integrated model of text and picture comprehension. The basic structure of the model is shown in Figure 1.4. The model consists of a descriptive (left side) and a depictive (right side) branch of representations.

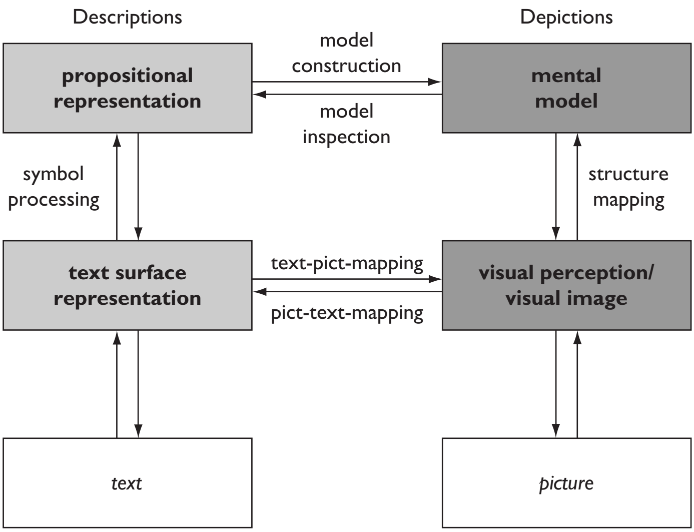


Figure 1.4 Outline of the integrated model of text and picture comprehension (Schnotz & Bannert, 2003).

The descriptive branch comprises the (external) text, the (internal) mental representation of the text surface structure and the (also internal) propositional representation of the semantic content. The interaction among these descriptive representations is based on symbol processing. The depictive branch comprises the (external) picture or diagram, the (internal) visual perception or image of the graphical display and the internal mental model of the depicted subject matter. The interaction among these depictive representations is based on processes of analogue structure mapping.

When an individual understands a text, he/she is assumed to construct three kinds of mental representations: first, he/she forms a mental representation of the text-surface structure, which is not an understanding yet, but allows repetition of what has been read. Second, based on this surface representation, he/she constructs a propositional representation, which grasps the semantic content of the text. Third, the reader constructs a mental model of the text content, a hypothetical internal quasi-object which holds an analogy to the represented subject matter. Thus, the model integrates the basic assumptions of current text comprehension theories (Kintsch, 1998; van Dijk & Kintsch, 1983; see also Graesser, Millis, & Zwaan, 1997; Schnotz, 1994). When a learner comprehends a picture, he/she also constructs multiple mental representations. First, he/she creates a visual perception or image of the graphical display, which is not understanding yet, but only seeing or imaging the display. Second, he/she constructs a mental model, which grasps the relevant structural features of the graphical display and holds an analogy to the represented subject matter. Third, inspection processes are applied to the mental model, which read off new information from it, and the results of the model inspection are encoded in the form of a propositional representation. Thus, the model integrates the basic assumptions of current picture comprehension theories (Kosslyn, 1994; Lowe, 1996; Schnotz, 2001, 2002).

Propositional representations are descriptive, because propositions are complex symbols which are – similar to the sentences of natural language – assembled of more simple symbols according to specific syntactic rules. Propositional representations are descriptions in a hypothetical mental language. They represent the ideas expressed in the text on a conceptual level, which is independent from the specific wording and syntax of the sentence. Mental models are depictive because they are assumed to be hypothetical internal quasi-objects that hold a structural or functional analogy to other objects which they represent on the basis of this analogy (Johnson-Laird, 1983).

Mental models differ in several respects from visual images. First, a mental model is not a sensory-specific form of mental representation. A mental model of a spatial configuration, for example, can be constructed with visual, auditory and haptic information. It is therefore more abstract than a visual image. Second, mental model construction implies a task-oriented thematic selection. The process of structure mapping includes only those parts of the graphical configuration which seem to be relevant for some anticipated tasks. Irrelevant details of the picture, which are included in the visual image, may be ignored in the mental model. Third, mental model construction implies an elaboration of the model based on prior knowledge and therefore the model entails information about attributes and relations that are not included in the picture or diagram. Hence, the mental model contains additional information from prior knowledge that is not included in the visual image.

According to the integrated model, both text comprehension and picture comprehension result in the construction of multiple mental representations: a descriptive propositional representation and a depictive mental model. The propositional representation and the mental model interact continuously via processes of model construction and model inspection. The integrated model assumes that pictures have a privileged access to the visual system, because pictures as depictive representations are encoded more directly into mental models, whereas in text comprehension, a recoding has to take place in order to construct a mental model. Accordingly, a mental model is constructed more easily with the help of pictures than with the help of texts (see Mayer, 2001).

Text and picture comprehension are not only based on external sources of information (i.e., the text and the picture), but also on prior knowledge that is stored in long-term memory as an internal source of information. In text comprehension, prior knowledge about the graphic pattern of written words, about the sound pattern of spoken words and about possible syntax structures is needed for the mental text-surface representation. In picture comprehension, prior knowledge is needed for the perception of the picture. Prior knowledge influences how easily pictorial information is categorised. Objects can be recognised faster and more easily, when they are presented from a typical perspective than when they are presented from an unusual perspective (Palmer, Rosch, & Chase, 1981). Conceptual prior knowledge about the domain is needed in text comprehension as well as in picture comprehension for the construction of a propositional representation and for the construction of a mental model. Prior knowledge can partially compensate for a lack of external information, for lower working memory capacity (Adams, Bell, & Perfetti, 1995; Miller & Stine-Morrow, 1998), and for deficits of the propositional representation (Dutke, 1996; McNamara et al., 1996; Soederberg Miller, 2001). There seems to be a trade-off between the use of external and internal information sources: pictures are analysed more intensively if the content is difficult and the learners' prior knowledge is low (Carney & Levin, 2002).

If the interaction between a propositional and a mental model is further extended, the process of comprehension turns into a process of thinking and problem solving. If the right representation for the problem at hand has already been constructed, then the right sequence of operations has to be applied in order to convert the given representational state into the required representational state. In other words, the mental model has to be transformed from the given state into the goal state (which corresponds to the information processing approach to problem solving) in order to read off the required information and encode it into a propositional format. If the right representation for the problem at hand has not been constructed yet, a restructuring of the mental model and the propositional representation is required (which corresponds to the Gestalt psychology approach to problem solving), before the required sequence of operations can be applied that will solve the problem.

The framework proposed by Schnotz and Bannert (2003) can be used not only for the analysis of text and picture comprehension, that is, for investigating the process of constructing internal (mental) representations from external representations such as texts and pictures. It can also be used in a reversed order, namely for the analysis of producing texts and graphics. For example, if an individual has specific ideas in mind, he/she can create corresponding propositional representations and mental models in his/her working memory. He/she can then externalise these representations by writing down a corresponding text or by drawing an external picture or a graph. The externalisation of knowledge through writing or drawing can be done for the purpose of communicating with someone else. However, the externalisation can also be done for the purpose of thinking and problem solving by the individual him/herself: the individual can reread his/her own text and reconsider his/her own picture or graph, reconstruct mental representations of the content and then further elaborate his/her mental representations beyond the previous representations. In other words: the individual can create external representations as cognitive tools for his/her own problem solving and, in this way, offload parts of the cognitive requirements of thinking and problem solving onto the environment. The process of externalisation of ideas by creating external representations, and the use of these external representations for further comprehension, thinking and problem solving, corresponds to what we have described above as communication of an individual with him/herself: the individual alternately takes the role of the sign producer and the sign recipient, whereby the sign producer uses the external sign to infer new information about the represented content.

In the remaining part of this chapter, we will analyse how different forms of representations can be used in math education and in science education by students of different ages and expertise.

Descriptive and depictive representations in science education and in mathematics

Science education in physics

Textbooks for science education are full of schematic drawings that show essential spatial or topological structures of a scenario, which is further analysed in terms of the thematically scientific concepts. If, alternatively, a descriptive representation of the scenario were given, most students (perhaps not even experts) would not be able adequately to analyse the scenario for answering questions or solving tasks. For example, depictive representations such as schematic topological drawings are often used in the field of electricity, where these drawings illustrate how the objects involved in the electric circuit are interconnected. Consider the following description of a part of an electronic circuit:

(1) Point A is connected via resistance R_1 to point B. Point B is connected via resistance R_2 to point C. Point C is connected via resistance R_3 to point D. Point A is connected with no resistance to C, and point B is connected with no resistance to D. All resistances R1, R2 and R3 have the same amount R.

Let us assume that students should answer the following question: which resistance would be a substitute for the combination of the resistances described above? We can assume that in such a case students would first construct – as an instance of a 'reversed order' usage of the framework of Schnotz and Bannert (2003) – a depictive representation of this resistance combination as shown in Figure 1.5a.

The analysis of this depictive representation, however, is not a trivial task with a straightforward solution, because different students can come up with different and contradicting answers. Let us assume that three students discuss their solutions with the following arguments. One of the students, Michel, says:

(2a) Direct connections between two points in an electric circuit have an extremely small resistance that can be ignored. Therefore, the connections between A and C and between B and D are not important for the computation of the substitute resistance. Accordingly, the structure corresponds to a simple serial combination, and the substitute resistance is 3 R.

Another student, Peter, argues as follows:

(2b) The electric current divides at point A into two partial currents I_1 and I_2. The current searches for the lowest resistance. Thus, it will flow at the one hand

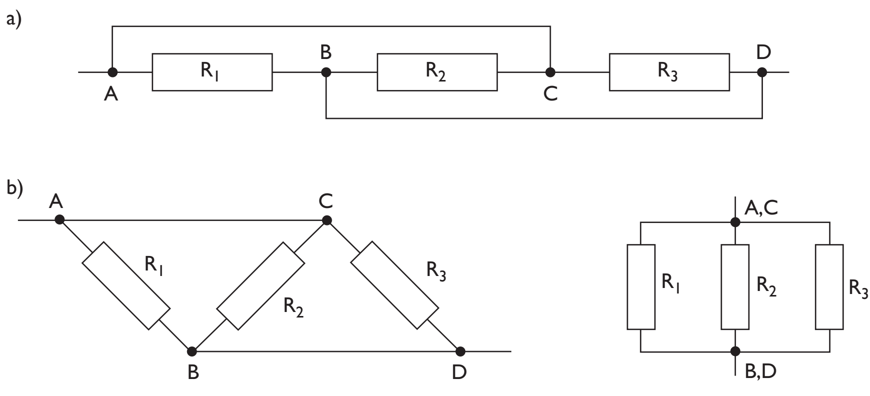


Figure 1.5 Depictive representations of a complex electric resistance problem.

from A to C and from there via R_3 to D, and on the other hand from A via R_1 to B and from there to D. Accordingly, there is a parallel combination of resistances R_1 and R_3, whereas no current flows through R_2. Thus, the substitute resistance for the combination above is R/2.

Finally, Karin says:

(2c) The described structure is a parallel combination of three equal resistances. Therefore, the substitute resistance for the above combination is R/3.

All three arguments seem to have a rational core, but only one of them is correct. The difficulty of finding the right answer in this case is not only due to the formation of a correct depictive representation and the formation of a correct descriptive representation. Instead, the difficulty originates from the specific structure of the depictive representation shown in Figure 1.5a because, due to its visual similarities with other resistance combinations, the representation triggers misinterpretations by erroneously reading attributes from the representation which are not really there.

For example, Michel's view of a simple linear resistance combination (2a) seems to be justified by the linear appearance of the structure in Figure 1.5a. Furthermore, he supports his view by the additional assumption that the connections between A and C and between B and D can be ignored. However, his assumption is not correct: whereas the resistance of these connections *can* be ignored, the connections themselves *cannot*. Thus, a seemingly minor modification of a descriptive representation (i.e., the replacement of 'ignoring of the resistance of connections' with 'ignoring the connections') has led to a fundamental misreading in the conceptual analysis of the depictive representation.

Peter gave a description of the flow of the current, which is also correct to a great extent: the current flows indeed from C to D via R_3 and from A to B via R_1. Therefore, R_1 and R_3 are in fact combined in parallel. However, Peter incorrectly assumes that no current flows through R_2. This seems to be supported by the depiction in Figure 1.5a, where the total current flows from A to D, that is, from left to right. The picture suggests that the current can 'bypass' R_2 on the one hand via the connection A–C and on the other hand via the connection B–D. Therefore, he incorrectly infers that the substitute resistance is R_2.

It should be remembered that a descriptive representation is always more general than a depictive representation, because a depiction has necessarily to be more specific. Accordingly, more than one depictive representation can be created on the basis of one description. In the case of the electronic circuit, the depictive representation shown in Figure 1.5a can be transformed (by reapplying the Schnotz–Bannert framework in reversed order) into the representation shown in Figure 1.5b. The transformation is topologically invariant.

That is, the result of the transformation has the same topological structure as the previous depiction in Figure 1.5a. This transformation follows a simple rule: 'If two points are directly connected, they can be merged into one point', which keeps the topological structure invariant. Contrary to the depictive representation in Figure 1.5a, the representation in Figure 1.5b allows seeing easily that the overall structure formed by the three resistances is in fact a parallel combination. Thus, Karin's answer (2c) was the correct one.

As the example demonstrates, it is not sufficient to construct any depictive representation in order to solve a problem. Instead, it is important that the depictive representation allows reading the relevant features easily. Sometimes, the specific perceptual structure or other perceptual attributes can obscure the relevant structural attributes and, thus, prevent application of correct procedures and stimulate application of other, incorrect procedures.

Mathematics education

Mathematics education in secondary schools emphasises at different points, that mathematical objects such as functions can be represented in different ways. It is well known that algebra has a correspondence with geometry and vice versa. For example, an algebraic function, which maps one variable onto another, can be represented by a description as well as a depiction. For example, if the description of the function is made by the term '$2y + x = 0$', then a corresponding depictive representation can be a straight line in a Cartesian diagram that crosses the origin (0,0) with a negative slope of -0.5. Although algebra and geometry form coherent bodies of knowledge by themselves, which can be considered independently from each other, we suggest that the close relationship between descriptive and depictive representations should be more elaborated in mathematics education.

It can be shown already to young students at the earlier levels of primary school mathematics that there is frequently not only one representation of a mathematical task, but that there exist different options for representations, and that each of these representations can have its own advantages and disadvantages. Rasch (2003) has investigated how young learners from primary school spontaneously find their own representations of complex mathematical text problems. A basic assumption in this project was that knowledge about the spontaneously created external representations of children provides insight into their mathematical thinking and serves as a basis for fostering their mathematical understanding and skills. Let us consider the following task, which was given to students from grades 3 and 4 of primary school:

(3) Little Ant on the Square:
A square has a side length of 200 m. A little ant walks along the sides of the square. In the daytime, the ant travels exactly 200 m. In the night, however, a strong wind blows the ant back one half of the distance that it has travelled

during the day. The ant starts on Monday morning. It starts from A across B, C and D back to A. When will the ant arrive at A?

(A drawing of a square is presented to the students as part of the task description.)

Nearly all students from grade 3 and 4 in primary school participating in the study used the presented square in order to answer the question. However, they used different kinds of mental representations and different strategies to solve the task. Catherine, for example, created the drawing shown in Figure 1.6. She used only a depictive representation and tried to reproduce the path of the ant on the pictorial level. She explained her strategy as follows:

(3a) I went from A to B, and then I went 100 m back. I continued this way until I reached A again. I found that the ant arrives at A after 7 days. So, it arrives at A on Sunday.

Catherine's strategy was totally pictorial, because she reconstructed the path of the ant step by step on the square as a depictive representation. Although this strategy was successful for her, other students who used the same strategy came to other solutions such as '8 days' or '10 days'. The solution included the reconstruction of a series of forward steps and backward steps on a depictive representation, and finally counting the days required to come back to A. Such a strategy is prone to error, because it is easy to miss one step or to miscount the steps.

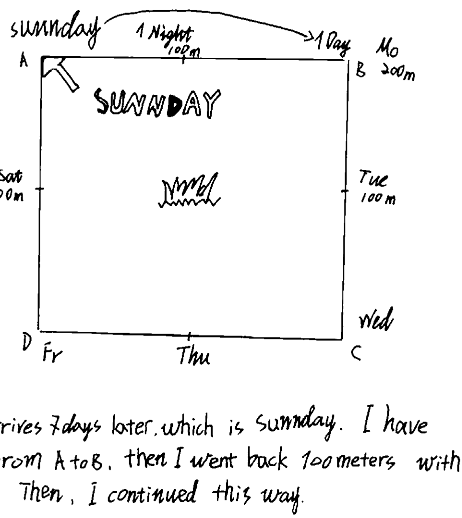


Figure 1.6 The little ant on the square: Catherine's solution.

Other students focused more on the mathematical structure of the task and combined the depictive representation of the square with a descriptive representation of the ant's movement. Robert, for example, described his strategy as follows:

(3b) The square has four sides. The ant has to walk half of the way twice, because the wind blows it back half of the distance it has travelled the day before. So, it needs 4 days for travelling the simple distance: four times 200 m for the total distance of 800 m. Then it needs another 2 days for travelling half of the distance a second time. So, it takes the ant 6 days to arrive at A.

Although Robert tried to solve the task on a mathematical symbolic level, he did not come up with the correct solution. As the verbal protocol revealed, he had created an incorrect descriptive representation of the task: the ant does not have to walk half of the way (which is 400 m) twice, but rather a distance of 600 m. With this correct description, Robert had found that a total distance of 800 m plus 600 m, which equals 1400 m, has to be travelled by the ant. So, if the ant travels 200 m during the day, it needs $1400 \div 200 = 7$ days, which was the correct solution. The example shows that creating a correct descriptive representation of the task is not a trivial requirement for the students, even if the basic structure of the situation is already presented as a depictive representation.

Another student, Rebecca, focused exclusively on the mathematical structure of the task. However, her description of the task differed from Robert's description mentioned above. She said:

(3c) The square has four sides. Each side has a length of 200 m. So the ant has to travel 800 m. The ant travels 200 m in the daytime, but is blown back 100 m at night. It finally moves forward only 100 m per day. Therefore, the ant needs 8 days, because 8 times 100 equals 800.

In Rebecca's strategy, the depictive representation of the square played only a minor role. The picture was used (if it was at all) to verify that the square with four sides of 200 m each has a total perimeter of 800 m. All the rest was done essentially on a descriptive (symbolic) representation: 200 m minus 100 m equals 100 m; 800 m divided by 100 m equals 8. Accordingly, 8 days was considered as the correct answer. The problem in case of Rebecca's strategy is that she does not use a depictive representation as a constraint for her descriptive (symbolic) representation. Although it is true that someone who travels 100 m per day needs 8 days to travel 800 m, the situation is different in the case of the little ant task, because the ant is *more* advanced on its way in the *evening* of a day than it is in the following *morning*. Therefore, the ant arrives at A already by the end of the seventh day (i.e., on Sunday), although it might wake up again 100 m away from A the next morning. Instead, the descriptive symbolic representation was used in a mechanical way, which ignored the fact that the ant's position is more advanced in the evening than it is the next morning. If the descriptive symbolic representation had been mapped on the depictive representation, the evening–morning difference might have become more obvious and perhaps helped Rebecca to find the correct solution also on the symbolic level.

A fourth student, Paul, used a strategy in which a depictive representation was closely integrated with a descriptive representation by a one-by-one mapping of graphical entities onto descriptive terms. This solution is illustrated in Figure 1.7. Paul placed distance tags on the square (100, 200, 300, . . . , 800). Then, he created a table which described the beginning and the end of the distance travelled by the ant for each day: 000–200 for day 1, 100–300 for day 2, 200–400 for day 3, etc. This table is a description (a kind of 'log file') of the ant's journey around the square. The stepwise elaboration of the table finally leads to the entry '600–800 for day 7', which answers the question of the task. Although Paul did not write down a formal procedure, he implicitly applied the following pair of production rules:

(3d)
RULE 1:

IF: Starting point of day i is x_i,

THEN: Endpoint of day i is $y_i = x_i + 200$)

RULE 2:

IF: End point of day i is y_i,

THEN: Starting point of day $i+1$ *is* $x_{i+1} = y_i$ *-100.* These rules were reiterated until the distance tag of the endpoint of the day was 800, which means that point A in the square was reached again. On the one hand, it seems unlikely that the extraction of these rules and their repeated use for generating the table had been possible without an external or internal depictive representation of the square. On the other hand, the entries of the table (including number of

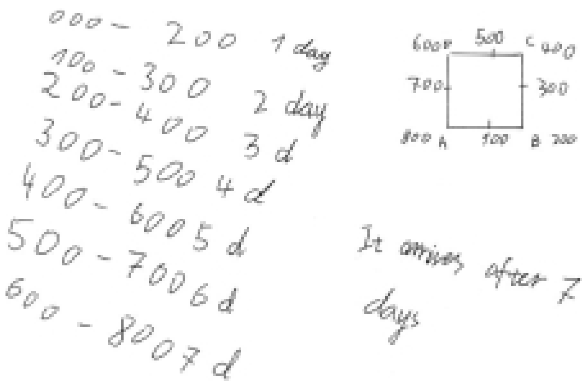


Figure 1.7 The little ant on the square: Paul's solution.

the day, starting point and endpoint) can only be interpreted meaningfully by reference to the depictive representation. Thus, this kind of solution is a nice example of a close interaction between a descriptive representation (the text), a depictive representation (the square elaborated by distance tags) and a descriptive representation in the form of a table, which finally leads to a precise solution based on insight.

External depictive representations of a problem constrain the possible descriptive representations and vice versa. The verbal description of a task by a written text often allows the construction of different depictive representations with specific structures. Sometimes, the depictive representation enables different procedures to operate on them, whereby some procedures are less complex (i.e., require less operations or easier operations) than others. Sometimes, one depictive structure allows operating more easily on it by counting or reading than another depictive structure. Sometimes, one depictive structure allows the creation of a simple descriptive representation that allows quick symbolic processing. Although the task presented above is relatively simple from an adult's perspective, it allows insight into the spontaneous mathematical thinking of primary schoolchildren. It demonstrates that children can be considerably creative in representing mathematical content and that they can use different representational formats with more or less flexibility in order to solve tasks. We consider the combined use of different representations – especially descriptive and depictive ones – as a key concept for teaching mathematics and for thinking and problem solving in mathematics.

Conclusion: teaching for flexibility in thinking and problem solving

The aim of our analysis was to illustrate that flexible thinking and problem solving require a specific interplay of structures and procedures – that is, two 'ingredients' that were formerly investigated by Gestalt psychology and the psychology of information processing. More specifically, flexible thinking and problem solving require representations with an adequate structure and procedures that can operate easily on this structure. A structure of a representation is adequate if it grasps the subject matter correctly and if it allows easy performance of the required procedures. Easy performance of the procedures means that the required operations are not difficult per se and that the sequences of operations are relatively short (see Ohlsson, 1984a, 1984b; Peterson, 1996).

Our analysis has also shown that a close interaction has to take place between descriptive representations and depictive representations and vice versa in order to make the best use of both kinds of representations for successful thinking and problem solving (Schnotz, 2005; Schnotz & Bannert, 2003). However, as we have demonstrated above, the creation of a depictive representation such as a drawing is not sufficient for successful cognitive problem solving, even if the drawing is correct. Sometimes, the specific perceptual structure or other perceptual attributes can obscure the relevant structural attributes and trigger the application of inappropriate procedures. Thus, it is also important that the individual creates the most useful depictive representation and that he/she knows how to operate on it. The depictive representation needs to have a structure that can be used by the individual to apply adequate procedures on it. These procedures are usually guided by descriptive representations of the subject matter. Some procedures can elaborate further the depictive representation. Other procedures can read new information from the depictive representation which is then encoded again in a descriptive format. Finally, it is important that the depictive representations are used as a cognitive tool, that is, as something to operate on as a means of thinking rather than being only a means for illustrating the results of thinking.

Our considerations suggest the following general guidelines for fostering flexible thinking and problem solving. First, learners should be taught systematically how to search for adequate forms of representations for the problems at hand. This also includes the teaching and learning of how to create adequate depictive representations that allow easy task-oriented performance. Although the corresponding guidelines might be to some extent domain-specific, there could also exist general heuristics of how to create and operate with different kinds of representations. Second, learners should be taught systematically how to analyse depictive representations and how to operate on them in a task-oriented manner. Operating on depictive representations implies a close interaction between description and depiction, because procedures on a representation are usually guided by a descriptive representation and because reading information from a depiction leads, in turn, to a description (or an elaboration of an already existing description). Accordingly, learners should be taught to closely interconnect descriptive and depictive representations when trying to solve problems. Finally, students need to learn how to transform one representation into another, informationally equivalent representation. By training the students' representational flexibility, they may become more skilled in finding the best representation for solving specific tasks or problems and, thus, become more creative and flexible in their thinking and problem solving by combining the right structures with the right procedures.

Note

1. The number of states in a Tower of Hanoi problem is 9 times the sum of 2^{i-1} whereby i increases (by incremenseparatosets of 1) from 1 to the number of discs minus 1.

References

Adams, B. C., Bell, L., & Perfetti, C. (1995). A trading relationship between reading skill and domain knowledge in children's text comprehension. *Discourse Processes*, *20*, 307–323.

Baddeley, A. (1986). *Working memory*. Oxford, UK: Oxford University Press.

Bühler, K. (1907). Tatsachen und Probleme zu einer Psychologie der Denkvorgänge. *Archiv für Psychologie*, *9*, 297–365.

Bühler, K. (1934). *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena: Fischer.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, *14*, 5–26.

Duncker, K. (1935). *Zur Psychologie des produktiven Denkens*. Berlin: Springer.

Dutke, S. (1996). Generic and generative knowledge: Memory schemata in the construction of mental models. In W. Battmann & S. Dutke (Eds.), *Processes of the molar regulation of behavior* (pp. 35–54). Lengerich, Germany: Pabst Science Publishers.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, *48*, 163–189.

Hörmann, H. (1976). *Meinen und Verstehen*. Frankfurt/Main, Germany: Suhrkamp.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, UK: Cambridge University Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hove, UK: Erlbaum.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kosslyn, S. M. (1994). *Image and brain*. Cambridge, MA: MIT Press.

Lowe, R. K. (1996). Background knowledge and the construction of a situational representation from a diagram. *European Journal of Psychology of Education*, *11*, 377–397.

Maier, N. R. F. (1931). Reasoning in humans. II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology*, *12*, 181–194.

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*, 1–43.

Miller, L. M. S., & Stine-Morrow, E. A. L. (1998). Aging and the effects of knowledge on on-line reading strategies. *Journal of Gerontology: Psychology Sciences*, *53B*, 223–233.

Montgomery, H. (1988). Mental models and problem solving: Three challenges to a theory of restructuring and insight. *Scandinavian Journal of Psychology*, *29*, 85–94.

Ohlsson, S. (1984a). Restructuring revised – Summary and critique of the Gestalt theory of problem solving. *Scandinavian Journal of Psychology*, *25*, 65–78.

Ohlsson, S. (1984b). Restructuring revised – An information processing theory of restructuring and insight. *Scandinavian Journal of Psychology*, *25*, 117–129.

Ohlsson, S. (1992). Information processing explanations of insight and related phenomena. In M. Keane & K. Gilhooly (Eds.), *Advances in the psychology of thinking* (pp. 1–44). London: Harvester-Wheatsheaf.

Palmer, S. E., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and performance* (Vol. 9, pp. 135–151). Hillsdale, NJ: Erlbaum.

Peirce, C. S. (1931–1935). *Collected writings*, C. Hartshorne, P. Weiss (Vols. 1–6) & A. W. Burks (Vols. 7–8) (Eds.), Cambridge, MA: Harvard University Press.

Peterson, D. (1996). *Forms of representation*. Exeter, UK: Intellect.

Rasch, R. (2003). *42 Denk- und Textaufgaben. Wie Kinder mathematische Aufgaben lösen und diskutieren*. Seelze, Germany: Kallmeyer.

Schnotz, W. (1994). *Aufbau von Wissenstrukturen. Untersuchungen zur Kohärenzbildung beim Wissenserwerb mit Texten*. Weinheim, Germany: Psychologie Verlags Union.

Schnotz, W. (2001). Sign systems, technologies, and the acquisition of knowledge. In J. F. Rouet, J. Levonen, & A. Biardeau (Eds.), *Multimedia learning – Cognitive and instructional issues* (pp. 9–29). Amsterdam: Elsevier.

Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14*, 101–120.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 49–69). Cambridge, UK: Cambridge University Press.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction*, *13*, 141–156.

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review*, *19*, 469–508.

Silveira, J. (1971). *Incubation: The effect of interruption timing and length on problem solution and quality of problem processing*. Unpublished doctoral dissertation. University of Oregon, U.S.A.

Soederberg Miller, L. M. (2001). The effects of real-world knowledge on text processing among older adults. *Aging, Neuropsychology and Cognition*, *8*, 137–148.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). Cambridge, UK: Cambridge University Press.

Sweller, J., van Merrienboer, J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychological Review*, *10*, 251–296.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Wertheimer, M. (1938). *Laws of organization in perceptual forms in a source book for Gestalt Psychology*. London: Routledge & Kegan Paul.

Chapter 2

Instructional considerations in the use of external representations

The distinction between perceptually based depictions and pictures that represent conceptual models

Stella Vosniadou

Existing research on the comprehension of text and pictures

Research on the comprehension of text and pictures has blossomed in recent years. Most of this research investigates the interaction between verbal information and the information contained in pictures in order to understand the conditions under which pictures are more likely to facilitate the comprehension of a given text. Various theories try to explain this text/picture interaction, such as the Dual Coding Theory (Clark & Paivio, 1991; Paivio, 1986), the Conjoint Processing Theory (Mayer, 1997), and the Integrative Model of Text and Picture Comprehension (Schnotz, 2001, 2002). In these theories the emphasis is usually placed on understanding the cognitive mechanisms that underlie the interplay between verbal and pictorial representation.

Mayer and his colleagues (Mayer, 1997, 2003; Mayer & Moreno, 2002, 2003), for instance, argue that students learn better when both words and pictures are present in book-based or computer-based environments, than when only words are presented to them. They also claim that there are certain methods or principles that optimise the effectiveness of such multimedia environments, such as the *coherence effect* – in which students learn more deeply when extraneous material is excluded rather than included – the *spatial contiguity effect* – in which students learn more deeply when printed words are placed near rather than far from corresponding pictures – and the *personalisation effect* – in which students learn more deeply when words are presented in conversational rather than formal style.

Schnotz and his colleagues (Schnotz, 2002; Schnotz & Bannert, 2003; Schnotz & Kürschner, 2008), on the other hand, have looked at the interaction between internal and external representations, the effects of multiple representations, and the relationship between external representations and cognitive processes including the amount of cognitive load. They argue that the inclusion of different pictures in the same text can result in the construction of different mental representations by the reader and in different patterns of performance in subsequently presented tasks. They also explain some of the conditions under which the presentation of external representations may not always be beneficial for the acquisition of new knowledge.

Finally, Lowe and Schnotz (2007) and Ainsworth (1999; Ainsworth & VanLabeke, 2004) have investigated dynamic representations and animations, and have found that it is not always beneficial to have multiple representations, as students, and particularly young students, may find it difficult to translate from one representation to another.

The explanatory framework used to interpret these research findings is usually that of cognitive load theory. In other words, external representations are said to facilitate performance because they reduce the cognitive load of the task or because they make specific cognitive processes easier, faster, and more accessible.

The above-mentioned studies have provided a great deal of information about the ways in which readers comprehend pictures in texts and have generated very useful recommendations for the improvement of curriculum materials and their instructional uses. There are two areas, however, where we believe that more research is needed in order better to understand the interaction between verbal and pictorial information in text comprehension. The first area concerns the distinction between external representations that are *perceptually based depictions*[1] and those that represent *conceptual models*. The existing research has not explicitly addressed students' difficulties in understanding the meaning of these two types of external representation. As will be explained later in detail, conceptual models usually have an analogical relationship to the perceptual situation that they represent and they may differ from that situation substantially in surface similarity, compared with perceptually based depictions.

The second area where more research is needed has to do with the effects of prior knowledge. While many researchers are concerned about how pictorial representations interact with verbal information, there is little discussion of how external representations may interact with domain-specific prior knowledge. The interaction between prior knowledge and the comprehension of pictorial information is particularly relevant in the case of conceptual models, because conceptual models depend more on domain-specific knowledge in order to be understood, compared with perceptually based depictions.

In this chapter we attempt to show that a distinction can be made between perceptually based external representations and those that can be considered conceptual models. Moreover, we argue that conceptual models are more difficult to understand than perceptually based representations, because they demand (a) specific scientific and mathematical domain knowledge, and (b) substantial epistemological sophistication. For this reason, we claim, it is important to distinguish between these two kinds of external representations and to be careful about how each one is used in curricula and instruction.

Two kinds of external representations: perceptually based depictions and conceptual models

Many of the pictures used in textbooks, particularly those in the areas of the physical and biological sciences, are not perceptually based representations, but conceptual models. In order to understand this distinction better, let us look at the two representations appearing in Figures 2.1a and 2.1b. The picture in Figure 2.1a is a perceptually based representation. It shows the sun setting behind the mountains and the moon up in the sky, signalling the approach of night. The picture in Figure 2.1b, however, cannot be considered a perceptually based representation. It is a conceptual model, representing our current understanding of the solar system. Conceptual models are usually theory-based as opposed to being grounded on everyday observations, being the products of scientific investigation. It took our culture hundreds of years of scientific discovery to come up with a representation of the solar system such as the one portrayed in Figure 2.1b, the Copernican revolution being considered one of the most important scientific revolutions in the history of science (see Kuhn, 1957). Even in the present day, we still debate about the exact nature of our solar system, about the definition of a planet, and about whether the planet Pluto is in our solar system or not.

It is not easy to give an exact definition of conceptual models. Nevertheless, we might say that conceptual models have some of the following characteristics, compared with perceptually based depictions: first, they represent entities, situations or phenomena in ways that are different from those that are perceptually experienced, at least from our usual egocentric perspectives and without the help of technological aids. Second, they include theoretical entities that need to be interpreted on the basis of relevant scientific or mathematical knowledge. For example, the representation of the solar system usually shows the orbits of the planets as they revolve around the sun. The movement of the earth and the other planets is not something that we can observe directly and depict visually in a static picture, but also the very idea that the earth is a planet and that planets orbit the sun is a theoretical one. Throughout the ages mankind has gone through different interpretations of the nature, location and movement of the sun, the earth and the other planets. This brings us to our third point, namely that conceptual models are often rather counter-intuitive in the way they relate to perceptual experience. For example, our everyday experience is that the earth, and not the sun, is at the centre of the solar system, and that the earth is stationary and not moving.

As another example, consider the two pictures in Figures 2.2a and 2.2b. The first is a perceptually based representation of ocean water, while the second is a conceptual model that depicts hydrogen bonding between water molecules in liquid water, with the dash lines signifying hydrogen bonds.

It may be argued that the difference between perceptually based depictions and conceptual models is not as large or as clear-cut as we claim, and that we

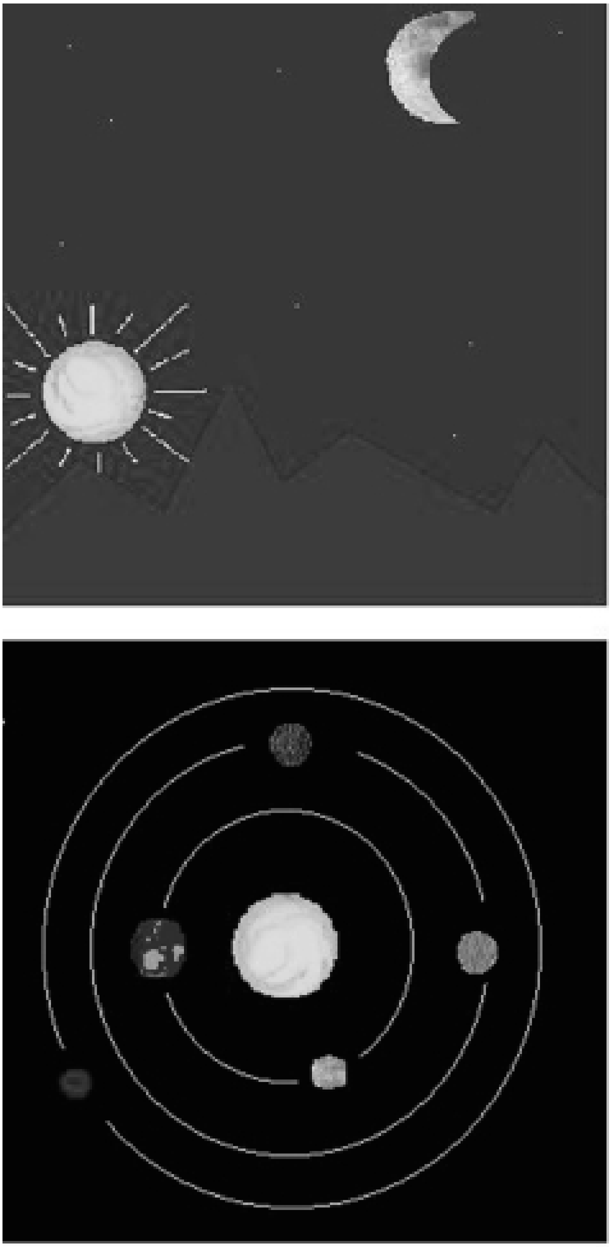


Figure 2.1 Perceptually based representations and conceptual models in astronomy.

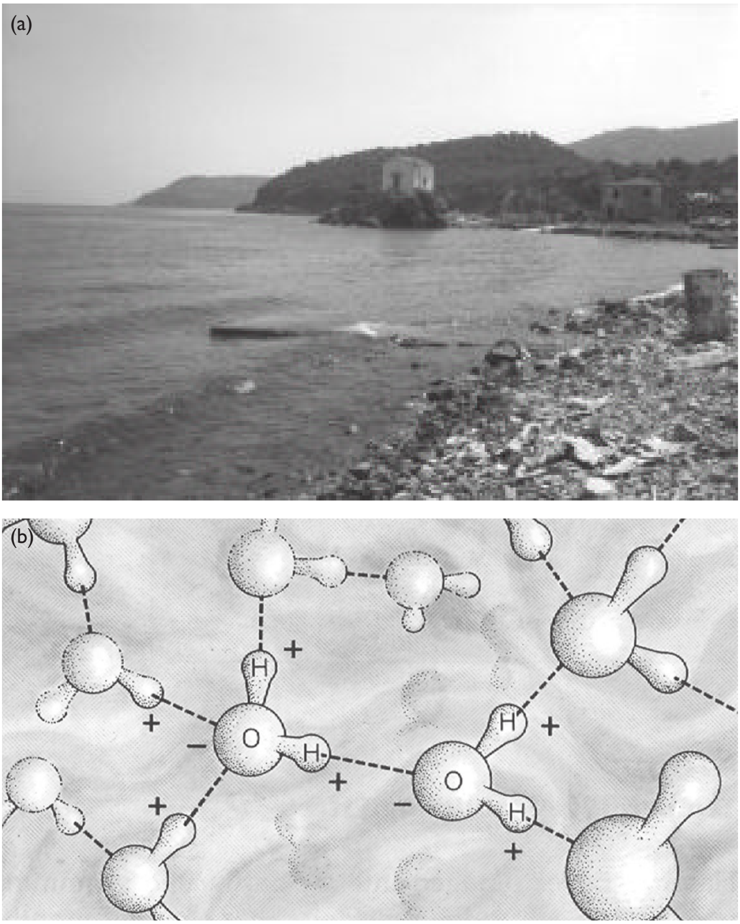


Figure 2.2 Perceptually based representations and conceptual models in chemistry.

are presenting a very simplified picture of a very complex issue. For example, it could be claimed that since we can now take pictures of the earth from satellites, the depiction of a spherical earth in space (if not of the whole solar system) is a perceptually based depiction and not a conceptual model.

We agree that there can be some cases where it may be difficult to distinguish between perceptually based representations and conceptual models. In fact it appears that the two kinds of external representations may not be clearly dichotomous but, rather, may represent a continuum, with some cases being more clearly differentiated than others. But even if we were to consider the picture of the earth from space as 'perceptually based', it still retains many of the characteristics of conceptual models. It is a depiction that represents a very different view from that experienced by an individual living on the earth, and requires, for its understanding, having sophisticated perspective-taking abilities and being able to distinguish between 'appearance' and 'reality'. It also requires the ability to understand that the same entity can have more than one representation, an ability that takes time to develop and can be greatly facilitated through specific instruction (Grosslight et al., 1991; Lehrer et al., 2000). It is for some of these reasons that children often fail to understand that the picture of the spherical earth taken from satellites, refers to the same 'flat' earth on which we live. There is in fact a common misconception among young children that there are two 'earths' – a flat one on which we live and a spherical one which is a 'planet' up in space. This is known as the 'dual earth model' of the earth (Vosniadou & Brewer, 1992).

Conceptual models play an important role in science as they are often used to flesh out the semantics of an axiomatic, syntactic theory, mainly in physics. An abstract, axiomatised theory can usually give rise to a class of models that provide the interpretation of the theory and connect it to the physical world. Giere (1988), following Hesse (1966), argues that conceptual models represent in some way the behaviour and structure of the physical system they represent, i.e., they are structural analogues to the physical system. Thinking of the Bohr atom, for instance, through the model of a system of billiard balls moving in orbits around one ball, with some balls jumping into different orbits at different times, then there are various kinds of analogies that can hold between the model and the real system. According to Morgan and Morrison (1999), there can be different kinds of conceptual models, some visualisable, others mathematical. In all cases, however, models are integral components of theories. They suggest hypotheses, aid in the construction of theories, and are a source of both explanatory and predictive power.

In this chapter we will discuss only the kinds of conceptual models found in external representations used in science and mathematics textbooks. Let us make clear here that we are not talking about an exotic phenomenon that is not worth paying attention to. The pictures we call conceptual models are not rare in science textbooks. On the contrary, they are used very commonly as visual aids to help students understand the theories presented in the text. For example, a random look at the first 13 pages of the Starr and Taggart (1992) college textbook entitled *Biology: The Unity and Discovery of Life*, reveals 22 pictures, all of which represent conceptual models.

Conceptual models are difficult to understand

Science education research has shown that students often fail to understand external representations that have the form of conceptual models. In some cases this happens because the students have misconceptions that may interfere with their interpretation of the conceptual model. Take, for example, the depiction of hydrogen molecules in Figure 2.2b. Many students have persistent misconceptions about molecules. They may believe that molecules are indivisible entities that expand when heated, or that molecules of solids are hard while molecules of gases are soft and weigh less, that molecules are glued together, etc. (see Andersson, 1990; Novick & Nussbaum, 1981; Wiser & Smith, 2008). Many students think that molecules (and/or atoms) are found *inside* matter rather than being the basic constituents of matter. They think that they are embedded in a material substrate rather than being the stuff that things are made out of (Andersson, 1990; Lee et al., 1993; Wiser & Smith, 2008). This 'molecules or atoms *in* matter' model is a very powerful misconception found even in college students taking chemistry courses (Pozo & Crespo, 2005).

In some cases misconceptions may be suggested by textbook illustrations in the form of conceptual models. For example, Wiser and Smith (2008) argue that the common depiction of molecules or atoms as small black dots inside a coloured sphere representing a given substance, can be very suggestive of the 'molecules (or atoms) in matter' misconception. In yet other cases, the specific words used to describe atoms and molecules in the context of a conceptual model, can generate a wrong interpretation. Expressions such as 'Atoms *in* solids vibrate', 'Molecules are less free to move *in* ice than *in* (liquid) water' and 'Bonds are the *glue* between atoms' can reinforce the 'molecules or atoms in matter' misconception. The word 'microscopic' may suggest that atoms can be seen with a microscope and thus generate a wrong sense about their scale, while the word 'particles' may suggest that they are particles of dust.

Difficulty in understanding conceptual models has also been documented in the area of astronomy. Ehrlen (2007) interviewed first-grade elementary children in order to ascertain how they understood the model of the earth as a globe. She found that some of them did not think that the globe had anything to do with the earth on which we live. One of the children called the globe 'a map' and did not seem to regard the earth as a spherical planet. Another called the globe 'a statue' and did not know whether it looked like the earth or not, while a third thought that the countries that appeared on the outside of the globe should actually be inside. The following is an excerpt from this last interview:

Margaret: 1st grader

- I: Does the globe look like the earth?
- M: Yes, but this is inside (points to the surface of the earth)
- I: You are pointing on the outside, on that country there.
- M: It is inside
- I: Yes?
- M: Yes
- I: It is inside. And the people then, where are they?
- M: Inside

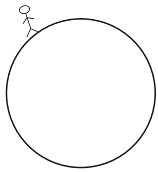


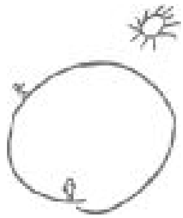
Figure 2.3 External representation used in Vosniadou and Brewer (1994) to investigate children's explanations of the day/night cycle.

Research on the development of students' understanding of the earth and of the day/night cycle has also revealed a number of cases where external conceptual models are misunderstood (Vosniadou & Brewer, 1992, 1994). Vosniadou and Brewer (1994) presented elementary schoolchildren with the drawing shown in Figure 2.3, told them that it shows a person living on the earth, and asked them to 'make it so it is day for that person' and then 'to make it night' for that person.

As can be seen from the examples shown in Figure 2.4, the children interpreted the conceptual model of the round earth in different ways. For example, Tamara (drawing number 1) thought that the drawing was wrong because the person was 'outside the earth'. When she was asked to show where the person should be located, she drew the person inside the earth, at the bottom, and then went on to explain the day/night cycle in terms of the sun being covered by clouds. Allison, on the other hand, accepted the drawing and added the sun to make it day, explaining that the sun 'goes in space' and 'when it gets dark the moon comes back in' (drawing number 3). Timothy also accepted the drawing, but he had a very different model of the day/night than Allison. He thought that the sun went down to the other side of the earth (drawing number 5).

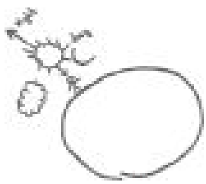
Difficulty in understanding conceptual models can also be found in the case of mathematics. The number line is, for example, an external representation widely used in elementary schools throughout the world to teach young children about natural numbers. However, the number line as a conceptual model can be interpreted differently by young children. Siegler and his colleagues (see Opfer & Siegler, 2007), for instance, have shown that second graders have a different mental representation of the number line than older children and adults. When second graders were asked to place numbers on a number line extending from 0 to 1,000, almost all of them placed the number 150 almost halfway across the number line!

Studies in our lab, where the number line was used to teach sixth-grade children about fractions, showed that children's misconceptions about fractions influenced their interpretation of the number line. For example, some children



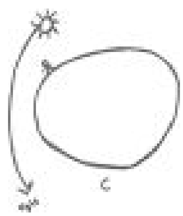
Drawing No. 1

Tamara (No.9, Grade 5)
The sun is occluded by clouds or darkness
E: Now can you make it so it is day for that person?
C: He's outside the earth.
E: Where should he be?
C: In here (see Figure 4, drawing 1)
E: . . . OK now, make it daytime for him.
C: The sun is out here, but it looks like it's in the earth, when it shines . . .
E: OK. What happens at night?
C: The clouds covered it up.
E: Tell me once more how it happens.
C: Cause at 12 o'clock it's dark.



Drawing No. 3

Allison (No.52, Grade 1)
The sun moves out into space.
E: Now make it so it is day for that person.
C: (child makes drawing 3 shown in Figure 4) Right here?
E: Whatever you think. Now make it night.
C: It goes in space.
E: Show me. Tell me how it happens.
C: The sun comes back down. It goes into space and when it gets dark the moon comes back out.

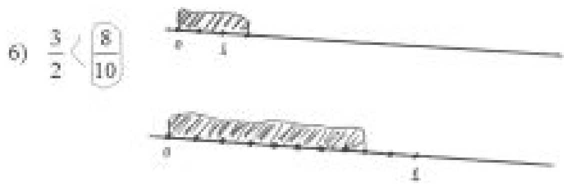


Drawing No. 5

Timothy (No.47, Grade 1)
The sun goes down to the other side of the earth (and the moon goes up)
The child makes the drawings shown in Figure 4.
E: Tell me once more how it happens.
C: When the moon comes up and the sun goes down.
E: Where was the moon before?
C: Under the earth.
E: Show me. Tell me how it happens.
C: What time was it when it goes under the earth?
C: Day

Figure 2.4 Examples of children's interpretations of the external representation given to them in order to explain the day/night cycle.

believe that fractions with large numbers are bigger than fractions with smaller numbers. This belief, which is correct for natural numbers, is transferred to the domain of fractions, causing children to increase the size of the unit when representing fractions consisting of larger numbers on the number line (Pantsidis & Vosniadou, 2006). The child whose response is shown in Figure 2.5, was asked to place the fractions $\frac{3}{2}$ and $\frac{8}{10}$ on the number line. As can be seen, although this child apparently knew that the fraction $\frac{2}{2}$ equals the unit, and that the fraction $\frac{10}{10}$ also equals the unit, he did not keep the size of the unit constant. Instead, he



Use the number line to represent the fractions 3/2 and 8/10.

Figure 2.5 Distortion of external representations representing fractions on the number line.

kept the parts of the unit the same and increased the overall size of the unit on the number line, ending up with a representation where the fraction $\frac{8}{10}$ is bigger than the fraction $\frac{3}{2}$.

The fact that the number line is used in instruction as a representation to help young children visualise natural numbers but that it is also used later to teach older students about rational and real numbers, can be very confusing. Students' early experiences with the number line in the context of discrete natural numbers can seriously inhibit their understanding of the continuous number line in the context of rational and real numbers. The result is that many students come to think of the line as a 'necklace of beads' – as consisting of discrete points laying the one next to the other. According to English (1993), students conceive the number line as a series of 'stepping stones' with an empty space in between, and thus tend to think that there are no numbers between two whole numbers.

Why are conceptual models difficult for students to understand?

There are at least two reasons why pictures in the form of conceptual models may be more difficult to understand than perceptually based depictions. The first concerns the fact that conceptual models are usually related to complex domain-specific theories. Thus, understanding conceptual models requires as a prerequisite to have at least part of this complex, domain-specific knowledge. The second difficulty concerns the gap between students' epistemic beliefs and the epistemic assumptions of conceptual models. Both of these reasons will be discussed in detail below.

Conceptual models are usually related to complex, domain-specific scientific theories. In addition to being rather abstract and complex, these domain-specific theories are difficult for students to understand because they are counter-intuitive, often violating some very entrenched beliefs in the context of naive physics or naive mathematics, which are continuously supported by everyday experience. In contrast, perceptually based depictions usually agree with perceptual experience and do not presuppose scientific or mathematical knowledge.

For example, in the area of chemistry, many students have a naive theory of matter according to which matter is something continuous, something that can be touched and seen, and something that exists in different material kinds, such as solids and liquids (Carey, 1991; Smith, Carey, & Wiser, 1985; Wiser & Smith, 2008). The atomic theory of matter – that matter is made up of atoms far too small to be seen even with a microscope, that there is empty space between atoms, and that each atom has mass and is in constant motion – violates practically all of the beliefs of the initial, perceptually based, naive theory of matter (Andersson, 1990; Novick & Nussbaum, 1981; Wiser & Smith, 2008).

In interpreting conceptual models, children often transfer to these models aspects of their prior knowledge which is based on everyday perception in the context of lay culture. The result of such an implicit transfer is often the creation of a 'synthetic' conception which distorts the conceptual model in ways that make it consistent with the child's original beliefs (see Vosniadou, Vamvakoussi & Skopeliti, 2008, for a detailed discussion of synthetic models). Wiser and Smith (2008) also argue that misrepresentations such as the 'atoms *in* matter' model represent synthetic conceptions, i.e., attempts by the students to synthesise information about the atomic theory of matter received in the context of school science with some of the presuppositions of their naive theory of matter, such as the presupposition that matter is fundamentally continuous.

Studies comparing problem solving in experts and novices show that not only children but also adults who are novices in science have naive theories which are very different from those accepted by current science and which may affect their interpretations of conceptual models. For example, according to Kozma et al. (1996), when chemists see an unknown reddish-brown gas, they may see this object as a model of a 'gaseous system' of indeterminate composition, consisting of one or more substances. They might infer that if the vessel contained more than one substance, they could be continually reacting at certain rates, determined in part by the temperature and pressure of the system and in part by properties of the substances. They could infer that at a stable temperature and pressure, these substances would be at 'equilibrium', reacting at equal and opposing rates such that the adjusted ratio of their partial pressures is a constant. However, when novices see the same object, they may only form an internal representation of 'a gas'. Their representations of the phenomenon may not include the possible existence of more than one substance. They might think that one property of the solitary substance is that it turns colour when heated. Furthermore, these representations do not usually exhibit the dynamic characteristics of the chemists' 'equilibrium system'.

As domain-specific knowledge is acquired, one's internal representations of key concepts in the domain change, increasing the similarity between internal representations and external representations in the form of conceptual models likely to be found in science texts. According to Larkin (1983), significant changes in internal representations with knowledge acquisition can be found in experts versus novices in mechanics. When given problems with 'blocks' and 'pulleys', she argues, experts see conceptual entities like 'forces'. Novices, however, only see the real objects. They construct mental representations that correspond to events or operations in the real world; they envision 'pushing' and 'pulling' the carts and blocks (Larkin, 1983). Similar observations have also been made by Chi and her colleagues (see also Chi, Feltovitch, & Glaser, 1981).

Finally, similar changes happen in the representation of number. Novices in mathematics see numbers as symbols that refer to discrete, physical entities, such as counting numbers. Experts, on the other hand, have formed a different representation of number as an abstract point on the number line which can take the form of different representations (e.g., integers, fractions, decimals) with the properties of density and continuity (Merenluoto & Lehtinen, 2002; Merenluoto & Palonen, 2007; Vamvakoussi & Vosniadou, 2010).

As individuals change their internal representations of key concepts in science and mathematics they become more capable of interpreting correctly external representations in the form of conceptual models. Such conceptual models can, however, easily be misunderstood by novices who lack the necessary background knowledge.

A second, important source of difficulty in understanding conceptual models has to do with students' commitment to a naive, realistic epistemology. There is a great deal of research on students' epistemological development which shows that even secondary school students and sometimes college students and pre-service teachers as well have an underdeveloped epistemological understanding of science (Grosslight et al., 1991; Hoffer & Pintrich, 1997; Smith & Wenk, 2006). They confuse theory and evidence in many ways (Kuhn, Amsel, & O'Loughlin, 1988), they do not understand the role of ideas in guiding the hypothesis-testing process (Carey et al., 1989; Grosslight et al., 1991; Smith et al., 2000), they know very little about the nature of scientific models and their relation to perceptual experience and specifically to the observed characteristics of objects and events (Lehrer et al., 2000; Wiser & Smith, 2008), and they do not know how to engage in model-based reasoning (Duschl, Schweingruber, & Shouse, 2007).

Students need systematic instruction in order to move from an unsophisticated perceptually based epistemology to a more advanced, model-based epistemology. They need to develop from a resemblance-based understanding of models to an understanding of abstract, conceptual models in science that serve an explanatory function and can be used as reasoning tools. In the last years a number of innovative curricula and instructional environments have been designed in order to improve students' epistemological sophistication and their abilities to reason with models (Lehrer et al., 2001; Smith & Wenk, 2006; Smith et al., 1997, 2000; Tytler, Peterson & Prain, 2006). Unfortunately, these experimental attempts have not yet influenced mainstream education in a significant way. Traditional science instruction at all levels continues to consider science as an accumulation of facts and fails to provide students with an understanding of the role of ideas, models and hypothesis-testing in the scientific discovery. At the same time, students' abilities to understand conceptual models in the form of textbook illustrations, and to reason using these models, are grossly overestimated.

Instruction to help students understand conceptual models

Various experimental studies provide good suggestions about how to design learning environments to develop a constructivist epistemology of science even in elementary school students (see, e.g., Brown & Campione, 1994; Lehrer et al., 2000; Smith et al., 2000; White, 1993). Some common elements in all of these are the following: a focus on ideas rather than on facts, on hypothesis-testing and model-based reasoning rather than on memorisation, on argumentation and sharing of ideas rather than on individual learning; and on giving students increasing responsibility for directing their own learning as opposed to having the teacher telling them what to do.

With respect to domain-specific knowledge, we argued that some domain-specific knowledge is a prerequisite to understanding many conceptual models. We must also acknowledge, however, that the use of conceptual models can be an important tool for developing this domain knowledge in the first place. Conceptual models in book-based or computer-based environments can help students understand complex mechanisms and they can help in making invisible entities and processes visible. They can represent phenomena from different perspectives, i.e., they can show how things look in everyday up/down space as opposed to 'deep' space, in real time as opposed to evolutionary time, etc. They can help students understand counter-intuitive concepts by revealing hidden constraints and presuppositions, and by relating theoretical entities to their referents. Such external representations can show, for example, how human movement measures connect to their graphical representations, or they may show an experiment where a liquid changes colour and, at the same time, an animation of an expert's model of it. They can also show how different representations of the same phenomenon can be related to each other (e.g., Vosniadou et al., 1996 and Kozma et al., 1996).

In using conceptual models to develop domain-specific knowledge, researchers, curriculum developers and teachers should be aware of how students' limited background knowledge can influence the way they interpret these conceptual models. Thus, a great deal of attention should be paid to how these models are designed so as not to become the source of misconceptions but, rather, to facilitate knowledge acquisition and text comprehension.

Conceptual models in book-based or computer-based environments are sometimes designed in ways that do not take into consideration students'

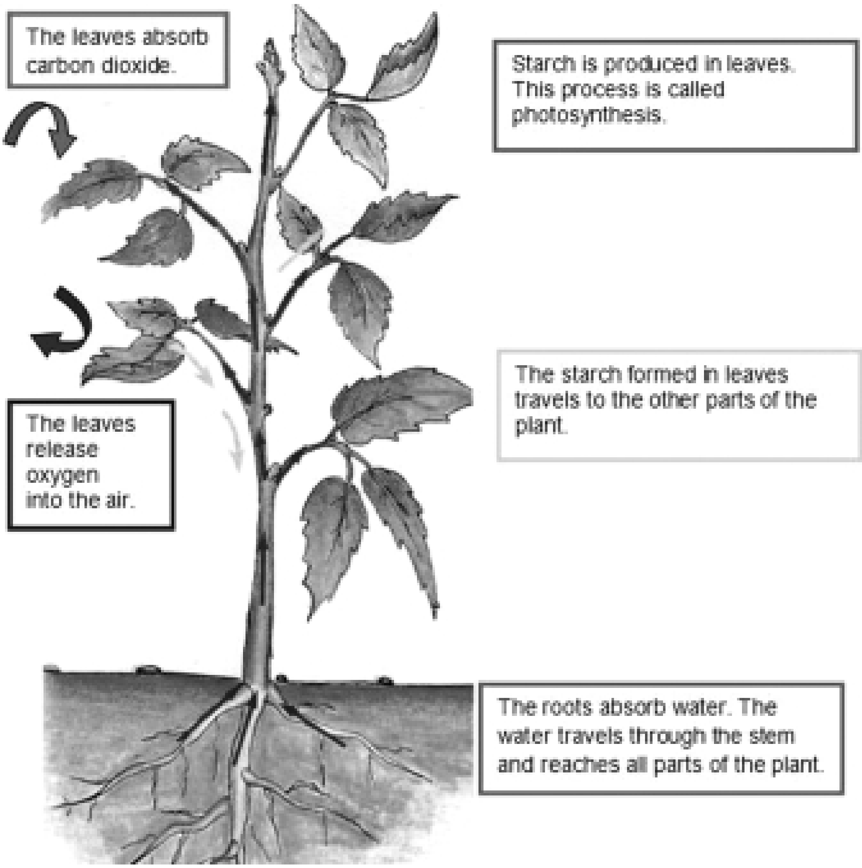


Figure 2.6 Pictorial explanation of photosynthesis.

conflicting prior knowledge, thus becoming the source of misconceptions. For example, the external representation shown in Figure 2.6, found in a Greek science textbook addressed to elementary schoolchildren, is very misleading. The purpose of this picture is to help students understand photosynthesis. However, it does not take into consideration that students often mix up their naive understanding of plant feeding and development – according to which, plants take food from the ground through their roots – with the scientific process of photosynthesis. As a result, such pictures may inadvertently create synthetic conceptions, such as that plants take food both from the ground through their roots and from the atmosphere through their leaves (Hershey, 2004; Kyrkos & Vosniadou, 1997).

In Figure 2.7 we find another conceptual model that is likely to be misunderstood. This picture was used in a science text that explained the

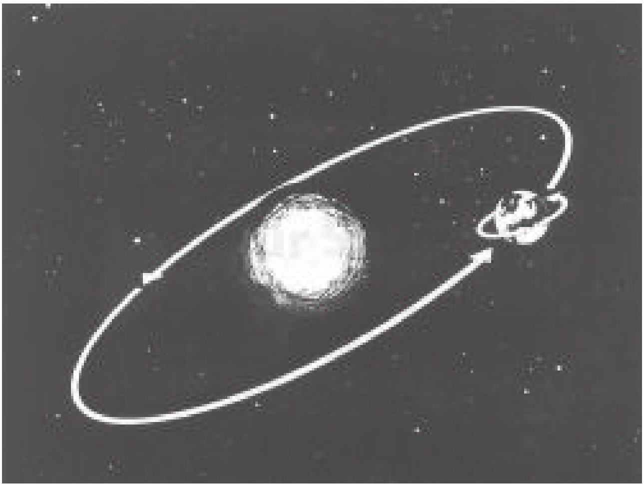


Figure 2.7 The earth's rotation and revolution around the sun.

day/night cycle and shows the earth rotating around its axis and also revolving around the sun, without explaining which of the two movements is responsible for the day/night cycle. The picture has the potential to reinforce a common mistake children make, namely to consider the revolution of the earth around the sun as the explanation of the day/night cycle (Vosniadou & Brewer, 1994). Such conceptual models, which show the earth revolving around the sun in an elliptical orbit, with the earth being closer to the sun some of the time, are also considered the main source of the widespread misconception regarding the change of the seasons, namely that summer happens when the earth comes closer to the sun (Schneps & Sadler, 1988).

Concluding remarks

We have argued that it is important to make a distinction between perceptually based external representations and those that represent conceptual models. Although external representations in the form of conceptual models can help students understand complex and counter-intuitive science and math concepts, they also have the potential to be easily misunderstood. Students often lack the necessary domain knowledge or the epistemic sophistication required to interpret conceptual models correctly. For this reason it is important (a) to develop specific instruction to teach students how to interpret conceptual models, and

(b) to design external representations that are conceptual models carefully, taking into account all that is known about students' prior knowledge, in order to avoid possible comprehension failures.

Acknowledgement

This work is supported by the project ANALOGY: Humans – The Analogy-Making Species, financed by the FP6 NEST Programme of the European Commission (STREP Contr. 029088).

Note

1. Throughout this chapter we will refer to conceptual models and perceptually based *external* representations only.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*, 131–152.

Ainsworth, S., & VanLabeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction, 14*, 241–255.

Andersson, B. (1990). Pupils' conceptions of matter and its transformations (age 12–16). *Studies in Science Education, 18*, 53–85.

Brown, Ann L., & Campione, Joseph C. (1994). Guided discovery in a community of learners. Classroom lessons: Integrating cognitive theory and classroom practice. In McGilly, Kate (Ed), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229–270). Cambridge, MA: The MIT Press, xvi, 317pp.

Carey, S. (1991). Knowledge acquisition. Enrichment or conceptual change? In S. Carey & R. Gelman (Eds.), *The epigenesist of mind* (pp. 257–291). Hillsdale, NJ: Erlbaum.

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). An experiment is when you try it and see if it works: A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education, 11*, 514–529.

Chi, M., Feltovitch, P., & Glaser, R. (1981). Categorization and representation in physics problems by experts and novices. *Cognitive Science, 5*, 121–152.

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

Duschl, R. A., Schweingruber, H. A., & Shouse, A.W., (2007). *Taking science to school: Learning and teaching science in grades K-8.* Washington, DC: The National Academies Press.

Ehrlen, K. (2007). *Conceptions and artefacts: Children's understanding of the earth in the presence of visual representations.* Doctoral dissertation, Stockholm University, Sweden.

English, L. (1993). Reasoning by analogy in constructing mathematical ideas. In G. Bell, B. Wright, N. Leeson, & Geake, J. (Eds.), *Challenges in mathematics education: Constraints on construction* (pp. 213–222). Lismore, Australia: Mathematics Education Group of Australasia.

Giere, R. (1988). *Explaining science: A cognitive approach.* Chicago: University of Chicago Press.

Grosslight, L., Unger, C., Jay, E., & Smith, C. (1991). Understanding models and their use in science. Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching, 28*, 799–822.

Hershey, D. R. (2004). *Avoid misconceptions when teaching about plants.* American Institute of Biological Sciences, from http://www.actionbioscience.org/education/hershey.html.

Hesse, M. (1966). *Models and analogies in science.* Notre Dame, IN: University of Indiana Press.

Hoffer, B., & Pintrich, P. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*, 88–140.

Kozma, R., Russell, J., Johnston, J., Marx., N., & Davis, J. (1996). The use of multiple, linked representations to facilitate science understanding. In S. Vosniadou, E. De Corte, R. Glaser, & H. Mandl (Eds.), *International perspectives on the design of technology-supported learning environments* (pp. 41–60). Mahwah, NJ: Erlbaum.

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills.* Orlando, FL: Academic Press.

Kuhn, T. (1957). *The Copernican revolution: Planetary astronomy in the development of Western thought.* Cambridge, MA: Harvard University Press.

Kyrkos, Ch., & Vosniadou, S. (1997). *Mental models of plant nutrition: A study of conceptual change in childhood.* Poster presented at the Seventh Conference of the European Association for Research on Learning and Instruction, Athens, Greece.

Larkin, J. (1983). The role of problem representation in physics. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 75–98). Hillsdale: NJ: Erlbaum.

Lehrer, R., Carpenter, S., Shauble, L., & Putz, A. (2000). Designing classrooms that support inquiry. In J. Minstrel & E. H. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 80–99). Washington, DC: American Association for the Advancement of Science.

Lowe, R. K., & Schnotz, W. (Eds.). (2007). *Learning with animation: Research and implications for design.* New York: Cambridge University Press.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*, 1–19.

Mayer, R. E. (2003). *Learning and instruction.* Upper Saddle River, NJ: Prentice Hall.

Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction, 12*, 107–119.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*, 43–52.

Merenluoto, K., & Lehtinen, E. (2002). Conceptual change in mathematics: Understanding the real numbers. In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 233–258). Dordrecht, The Netherlands: Kluwer.

Merenluoto, K., & Palonen, T. (2007). When we clashed with the real numbers: Complexity of conceptual change in number concept. In S. Vosniadou, A. Baltas, & X. Vamvakoussi (Eds.), *Reframing the conceptual change approach in learning and instruction* (pp. 247–263). Oxford, UK: Elsevier.

Morgan, M. S., & Morrison, M. (Eds.). (1999). *Models as mediators: Perspectives on natural and social science*. Cambridge, UK: Cambridge University Press.

Novick, S., & Nussbaum, J. (1981). Pupils' understanding of the particulate nature of matter: A cross-age study. *Science Education, 65*, 187–196.

Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology, 55*, 169–195.

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

Pantsidis, Ch., & Vosniadou, S. (2006). *The role of external representations in the ordering of fractions*. Paper presented at the Fifth European Symposium on Conceptual Change 'Bridging the Gap between Mental Models and Situated Cognition? Theoretical and Methodological Considerations'. Stockholm, Sweden.

Pozo, R. M., & Gomez Crespo, M., (2005). The embodied nature of implicit theories: The consistency of ideas about the nature of matter. *Cognition and Instruction, 23*, 351–387.

Schneps, M. H., & Sadler, P. M. (1988). *A private universe*. New York: Annenberg/CPB.

Schnotz, W. (2001). Sign systems, technologies, and the acquisition of knowledge. In J. F. Rouet, J. J. Levonen, & A. Biardeau (Eds.), *Multimedia learning: Cognitive and instructional issues* (pp. 9–29). Amsterdam: Elsevier Science.

Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14*, 101–120.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction, 13*, 141–156.

Schnotz, W., & Kürschner, C. (2008). External and internal representations in the acquisition and use of knowledge: Visualization effects on mental model construction. *Instructional Science, 36*, 175–190.

Smith, C., Carey, S., & Wiser, M. (1985). On differentiation. A case study in the development of the concepts of weight, size and density. *Cognition, 21*, 177–237.

Smith, C., Maclin, D., Grosslight, L., & Davis, H. (1997). Teaching for understanding: A study of students' preinstruction theories of matter and a comparison of the effectiveness of two approaches to teaching about matter and density. *Cognition and Instruction, 15*, 317–394.

Smith, C. L., Maclin, D., Houghton, C., & Hennesey, M. G. (2000). Sixth-grade students' epistemologies of science: The impact of school experiences on epistemological development. *Cognition and Instruction, 18*, 349–422.

Smith, C. L., & Wenk, L. (2006). Relations among three aspects of first-year college students epistemologies of science. *Journal of Research in Science Teaching, 43*, 747–785.

Star, C., & Taggart, R. (1992). *Biology the unity and diversity of life* (6th ed.). Belmont, CA: Wadsworth Publishing Company.

Tytler, R., Peterson, S., & Prain, V. (2006). Picturing evaporation: Learning science literacy through a particle representation. *Teaching Science, the Journal of the Australian Science Teachers Association, 52*(1), 12–17.

Vamvakoussi, X., & Vosniadou, S. (2010). How many decimals are there between two fractions? Aspects of secondary school students' understanding of rational numbers and their notation. *Cognition and Instruction 28*, 181–209.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24*, 535–585.

Vosniadou, S., & Brewer, W. F. (1994). Mental models of the day/night cycle. *Cognitive Science, 18*, 123–183.

Vosniadou, S., De Corte, E., Glaser, R., & Mandl, H. (1996). *International perspectives on the design of technology-based learning environments.* Mahwah, NJ: Erlbaum.

Vosniadou, S., Vamvakoussi, X., & Skopeliti, I. (2008), The framework theory approach to the problem of conceptual change (pp. 3–34). In S. Vosniadou (Ed.), *International handbook of research on conceptual change*. New York: Routledge.

Wiser, M., & Smith, C. (2008). Learning and teaching about matter in grades K-8: When should the atomic-molecular theory be introduced? In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 205–239). New York: Erlbaum.

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

Chapter 3

Critical thinking about biology during web page reading

Tracking students' evaluation of sources and information through eye fixations

Lucia Mason and Nicola Ariasi

Introduction

Our research addresses the role of external representations in reasoning and problem solving by examining the epistemic processing of texts and pictures about a biology topic, which are presented on multiple Internet pages. Epistemic processing is intended as processing that takes into account the source, reliability, and accuracy of information (Hofer & Pintrich, 1997). It is manifest in evaluative processes of the trustworthiness of an informational source and the veracity of its content, especially in the context of web page reading. As such, epistemic processing reflects an important aspect of critical thinking (King & Kitchener, 2002; Kuhn & Weinstock, 2002) activated by a reader who deals with multiple documents on the same topic.

Since the beginning of the new millennium, the World Wide Web has become a tool used by almost anyone – at least in western and many eastern countries – to search information on a huge variety of topics. Although it was not originally developed for educational purposes, the Web is now the biggest source of information for students, who rely on it, rather than on printed material, more and more for academic and school assignments. Undoubtedly, this is an important aspect of the democratisation of our current cultural context. At the same time, however, the infinite body of information made available on this potentially global learning environment by a simple click of the mouse, poses questions of locating, selecting, and evaluating the sources of information, as well as about the information itself, which requires new and complex skills. First of all, information sources must be located. The information must then be read, extracted, integrated, and synthesised into coherent material which can be used to build new knowledge (Kuiper, Volman, & Terwel, 2005). These crucial processes, however, do not cover all those required to use Internet-based material effectively. Navigating the Web to learn more about a topic is not only a question of formulating efficient search queries and applying appropriate reading strategies. It is also a question of evaluating the credibility of websites and the veracity of the information provided. Epistemic processes about the knowledge accessed and learning from the Web must be activated to deal with multiple sources and to follow up only the most reliable of these (Stadtler & Bromme, 2008).

In this chapter we focus on the latter question to examine whether university students take into account the authoritativeness of sources when reading Internet pages on a biology issue and whether their epistemic processing is related to individual characteristics. In particular, epistemic processing was addressed indirectly by the analysis of eye fixations during Internet page reading. Methodologically, an 'online' technique was used that allows objective measurement of the allocation of visual attention while reading, that is, eye fixation measurement. The analysis of visual attention allocated *within* various pages makes it possible to track epistemic evaluation of sources, and their contents.

Information on the Web is presented in a wide variety of representational formats. In this regard, as the multimedia tool par excellence, the Web is a wealth of external representations. Web pages may contain graphs, diagrams, animations, etc. as well as texts and pictures. In this study, we examined visual attention allocated while processing the texts and pictures appearing on various web pages. The research focused on evaluative processes of source reliability, which may lead to concentrating more on certain information and less on other, or to processing the same information differently according to the epistemic authority of the source itself.

Epistemic evaluation of web sources

The value of a search is not reflected in the number of documents retrieved ('the more, the better'), nor the retrieval speed ('the faster, the better'). It does not make sense to find hundreds of documents quickly, although typically this is what happens when students do not have the necessary skills to access and manage information meaningfully. A number of studies have been carried out to examine patterns of information-seeking behaviour (Brand-Gruwel, Wopereis, & Vermetten, 2005; Lin & Tsai, 2007; Tu, Shih, & Tsai, 2008) and descriptive models of it have been proposed (e.g. Brand-Gruwel et al., 2005).

Evaluation of the credibility of a web source is included in these models, as a crucial aspect of information problem solving, although few studies have been carried out on students' judgements of the authoritativeness of an electronic source or the accuracy of its content. In the past, the difficult task of controlling the truth of information was undertaken by editors and publishing companies, while today this task must be carried out by the learners themselves when navigating the Internet to learn more about a topic (Bråten & Strømsø, 2006; Tsai, 2004). Difficulty in recognising the epistemic value of an information source can be expected in younger students, as documented in the literature. Fifth-graders may believe that all they find on the Web is true (Schacter, Chung, & Dorr, 1998), and together with some seventh-graders, may not be aware that the knowledge provided on multiple sites is conflicting because of a one-sided perspective, or may be unable to perceive the plurality of the information read (Mason & Boldrin, 2008). It has also been documented that junior and senior high school students' evaluation is limited when assessing the credibility or accuracy of websites for their varying authoritativeness (Brem, Russell, & Weems, 2001), or they may not take into account the authority of a web source when evaluating a theory on a controversial scientific topic (Clark & Slotta, 2000).

Limited ability to evaluate Internet sources is more surprising and worrying in older students. College students may not necessarily be aware of an explicit epistemology when evaluating online information (Hofer, 2004). In addition, they may rely only on a general search engine (Google) to find out more about an unfamiliar scientific topic and not access any specialised scientific databases. They might not make judgements about the authoritativeness of a source in itself, or might appeal to naive criteria (Mason & Boldrin, 2008). Only when university students are prompted to use evaluation criteria to select web links and assess web pages, do their quality- and credibility-oriented criteria increase (Kammerer, Werner, & Gerjets, 2008).

Epistemic beliefs and web navigation

Beliefs about the nature of knowledge and the process of knowing, namely epistemic beliefs (Hofer & Pintrich, 1997), have been the focus of research which has flourished over the last two decades in both developmental and educational psychology.

From a developmental point of view, the trajectory that characterises the evolution of representations about knowledge and knowing has been indicated (King & Kitchener, 1994; Kuhn, 2000). The shared developmental sequence that can be identified across the proposed models leads individuals to shift, in Kuhn's (2000) terms, from an absolutist to a multiplist to an evaluativist position about knowledge and knowing. According to the absolutist view, knowledge is absolute, certain, non-problematic, right or wrong. It is based on observations from reality or authority, thus it does not need to be justified (knowledge as *fact*). According to the multiplist position, knowledge is ambiguous and idiosyncratic, thus each individual has his or her own views and truths (knowledge as *opinion*). At the evaluativist level, knowledge is conceived as plural and hypothetical but the individual also believes that there are shared norms of inquiry and knowing, thus some positions may be reasonably more supported and sustainable than others (knowledge as *judgement*).

In educational psychology, scholars have identified the various dimensions around which epistemic beliefs articulate. There is substantial agreement among scholars about four epistemic dimensions underlying these beliefs, two regarding the nature of knowledge and two regarding the nature of knowing (Hofer, 2000). The first epistemic belief dimension about the nature of knowledge comprises convictions about the simplicity vs. the complexity of knowledge (from knowledge as a set of discrete and simple facts to knowledge as a complex and interrelated network of concepts). The second epistemic dimension about the nature of knowledge concerns beliefs about the certainty vs. the uncertainty of knowledge (from knowledge as stable and absolute to knowledge as tentative and evolving). The first epistemic dimension about the nature of knowing concerns the relationship between knower and known (from transmitted knowledge that resides outside the self to knowledge that it is rationally constructed by the self). The second epistemic dimension about the nature of knowing concerns the justification of knowledge (from observation or authority as a source, to rules of inquiry and standards for the evaluation of a source).

Beliefs about the nature, justification, and source of knowledge are not extraneous to students' behaviour during Web use. Research has documented that epistemic beliefs are related to decision-making processes during information searching (Withmire, 2003), search outcomes (Tu, Shih, & Tsai, 2008), preference for higher-order metacognitive activities in Internet-based learning environments (Tsai & Chuang, 2005) and discussion and communication activities (Bråten & Strømso, 2006). In all these studies, more advanced beliefs about knowledge and knowing are associated significantly with better performance. In addition, there is evidence that university students spontaneously activate epistemic beliefs while reading web sources to acquire information on a controversial topic, and that this epistemic activation influences Internet-based learning (Mason, Boldrin, & Ariasi, 2010), favouring students who process the material at a higher epistemic level. Furthermore, it has been documented that eighth-graders' epistemic reflections about the justification of online knowledge during a retrospective interview influence their learning of the debated topic under consideration, favouring more epistemically sophisticated students (Mason, Boldrin, & Ariasi, 2008).

Eye tracking and web page reading

In the above-mentioned studies, students' evaluation of Internet informational sources was examined by two methods: one online, thinking aloud, and the other offline, as a retrospective interview. The former is advantageous for providing a rich source of data but, because of its intrusiveness, can alter the process of thinking itself, since cognitive resources are diverted from the execution of the primary task (Veenman, Van Hout-Wolters, & Afflerbach, 2006). Retrospective interviews are not intrusive but cannot reveal what spontaneously comes to mind since students are solicited to evaluate the web sources accessed. In the study reported below, a method that combines the positive aspects of both online and offline measures was used, that is, eye-tracking measurement, to focus on the allocation of visual attention during Internet page reading. Although it has been used extensively in several fields of cognitive psychology, especially to study lexical access and syntactic parsing (Rayner, 1998), this technique has very rarely been used in instructional psychology (Verschaffel, De Corte, & Pauwels, 1992) until recently (Hyönä, Lorch, & Kaakinen, 2002; Jarodzka et al., 2010). The theoretical assumption underlying this method is that saccades (eye movements) and attentional shifts are necessarily linked. If, in simple discrimination tasks, the locus of attention can be decoupled from eye location (Posner, 1980), in complex information processing tasks, such as reading, scene perception, and visual search, the link between the two is necessarily close (Rayner, 1998). Individuals process in the mind the visual information that is currently being fixated by the eye (Just & Carpenter, 1980). Eye movements are, therefore, a strong online indicator of the cognitive processes implicated, reflecting them 'moment-to-moment' (Liversedge & Findlay, 2000). In this regard, very recent research has clearly indicated that eye-tracking methodology provides fine-grained data useful for analysing, or even stimulating, information processing in multimedia learning (van Gog & Scheiter, 2010). Eye fixation data, therefore, provide the means for tracking cognitive processes also while reading Internet pages (Hyönä, Lorch, & Rinck, 2003): gazes on relevant information have higher densities than gazes on irrelevant information.

Although several studies have investigated the allocation of visual attention during web page reading in relation to the utility of electronic resources or web advertising (e.g., Wang & Day, 2007), very little research has been carried out on visual behaviour during web page reading in an educational context. One study examined evaluation processes during online searching, in particular the effectiveness of prompts in evaluation. Data showed that prompts did not affect university students' gaze behaviour, although their verbal utterances expressing criteria oriented to credibility increased (Kammerer et al., 2008). Another study focused on the effects of information problem-solving skills on judgement of a Google search page with 20 results. Findings regarding eye fixations revealed that students with low information problem-solving skills did not consider all results, but quickly selected a top-of-page result (Meeuwen, Brand-Gruwel, & van Gog, 2008).

In the study reported below we were interested in examining whether web pages would be processed differently according to their different levels of authoritativeness. Measurement of eye fixations on different types of information would make it possible to track information processing at an epistemic level, in this case in relation to source credibility. An Internet reader allocates more visual attention to information that is processed as reliable and accurate, while less attention is allocated to information that is evaluated as scarcely reliable or not reliable. We were also interested in examining whether individual differences, such as prior knowledge, domain epistemic beliefs, and argumentative reasoning, can influence epistemic processing. Greater prior knowledge can lead to searching new information and higher awareness of the accuracy of information (Hirsh, 1999). More constructivist beliefs about knowledge in a domain can lead to greater awareness of the complex and changing nature of knowledge (Mason, Gava, & Boldrin, 2010). Higher argumentative reasoning skills can lead to a better detection of fallacies in the arguments provided (Weinstock, Neuman, & Glassner, 2006).

The present study

Research questions and hypotheses

The study was aimed at expanding, theoretically and methodologically, current research on Internet searching and information processing. Theoretically, we related issues of research lines that are generally independent, that is, research on web navigation (Brand-Gruwel et al., 2005; Kuiper et al., 2005; Tsai, 2008) and research on epistemic beliefs (Hofer, 2000; Hofer & Pintrich, 1997). Methodologically, an online technique was used that allows objective measurement of the allocation of visual attention during reading, that is, eye-fixation measurement to track epistemic processing.

During web page reading not all students were expected to be epistemically involved to the same extent in information processing, depending on the source. Based on the literature, we examined the role of prior knowledge, domain epistemic beliefs, and argumentative reasoning, that is, all the individual factors that can potentially play a role in epistemic processing.

Specifically, the following research questions guided this study:

- Does information receive attention *within* the same web source and *across* different sources according to the source authoritativeness?
- Is attention allocation influenced by individual characteristics such as prior knowledge, epistemic beliefs, and argumentative reasoning?

For the first research question, we hypothesised that source reliability would play a role since it ‘suggests’ to the reader to concentrate more on the type of information that is relevant to the particular site, that is, the classic information on the most expert and reliable page and the less familiar information on the least competent and reliable page. Specifically, *within* the most authoritative source, text 1 and picture 1, which provided information about the classic, better known, universal validity of the central dogma of molecular biology, would be fixated with more attention than text 2 and picture 2, which provided less familiar information against the dogma. More attention on text 1 and picture 1 would mean that they are elaborated more as they deserve to be followed up. An opposing pattern of allocation of visual attention should emerge *within* the least credible page. In addition, we also expected that students would allocate attention on the same information differently *across* web pages according to their credibility. Specifically, we expected that both text 1 and picture 1 would receive more attention on the institutional, most reliable page than on the other pages, likewise text 2 and picture 2 on the alternative, less credible source. More relevant information from the unreliable page would be the new, less familiar issues about the biology dogma, which may conflict with a classic perspective on the question.

For the second research question, we hypothesised that individual differences would play a role in the allocation of attention during reading. Higher prior knowledge, more advanced epistemic beliefs and higher argumentative reasoning skills would lead to greater attention on less-known information, especially from authoritative sources, as these personal characteristics would be resources which help readers evaluate less familiar accounts about the examined question.

Method

Participants

Thirty-seven college students, from the faculties of biology and psychology, were involved (F = 19, M = 18). Their mean age was 22.6 (SD = 2.1).

Material

With the simulated aim of collecting information to write a report on the topic, participants were asked to read four 'web stimuli', prepared by taking information from real pages. Each was made up of five types of information: headline, two texts and two pictures (Figure 3.1). One text and picture composed an

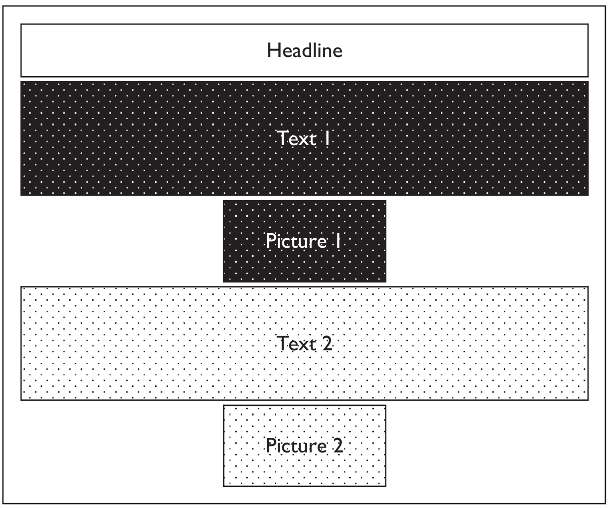


Figure 3.1 The five areas of interest in each web page that correspond to the different types of information provided.

argument for, while the other text and picture made up an argument against the universal validity of the central dogma of molecular biology, first enunciated by Crick (1970). It postulates the unidirectional nature of the transfer of genetic information, that is, from nucleic acids to proteins and not vice versa: DNA can be copied to DNA (DNA replication), DNA information can be copied into mRNA (transcription) and proteins can be synthesised using the information in mRNA as a template (translation). This dogma was considered universally valid until the discovery of retroviruses, such as HIV. In the case of retroviruses, the reverse of normal transcription occurs, as genetic information is transferred from RNA to DNA.

The four pages included information that was exactly the same in terms of content, style and structure regarding both texts and pictures. To make these characteristics as similar as possible across the pages, the information was manipulated slightly and not simply taken from the Internet. Each page provided an argumentation for and an argumentation against the central dogma of molecular biology. For conceptual reasons related to the topic, argumentation 1 (pro) should precede argumentation 2 (con) in all pages. In addition, in all pages the texts were identical in the numbers of sentences and words, and in the size of the images. What differentiated the four pages was only their location on an 'epistemic continuum' from the least to the most authoritative: institutional (the site of the CNR, the National Council for Research in Italy), encyclopaedic (Wikipedia), popular (Science and Knowledge), and alternative ('The genetics of chickens'). The epistemic status of the pages could be inferred from the information at the top of the page, that is, the heading including its URL (e.g. www.igb.cnr.it) and full name (e.g. Istituto di Genetica e Biofisica, Consiglio Nazionale delle Ricerche).

To avoid the possibility of order of appearance interfering with the selection of web sources to be read, four different combinations of results were prepared: in the first outcome, the first source was institutional; in the second outcome, the first was the encyclopaedic page; in the third outcome, the first was the popular site, and in the fourth outcome, the first was the alternative page. One of these four combinations was presented in a random order to each participant, who could spend up to 60 seconds reading. This length of time was chosen for two main reasons, one methodological and the other conceptual. First, we needed to build an environment that was as controlled as possible to ensure the most appropriate measurement of visual behaviour during web page reading. In a pilot study, we asked participants to read each source while their reading times were measured. The average reading time of all web pages was about 60 seconds; therefore this time limit was chosen as the threshold reading time in the present study. Second, eye movement data, like response times for other types of task, can be considered viable indices of the fast and efficient processes involved in epistemic processing, which takes place at an early stage of reading information (Richter, Schroeder, & Wöhrmann, 2009). Of course, we do not maintain that 60 seconds are enough for a deep learning of the concepts presented in the texts and pictures within the pages. Given that no educational studies have examined epistemic processing of online information using quantitative process data, we relied on issues of very recent research in social cognitive psychology and assumed that epistemic processing of information is activated early.

Each type of information within each page represented an 'area of interest' for the analysis of eye fixations (see below). Eye movements were recorded by a 50 Hz Tobii 1750 eye tracker, supported by the software program Clear View 2.7.1.

Measures

Participants' prior knowledge about the biology topic was indicated by the grade obtained in a biology examination.

Epistemic beliefs about the domain of biology were measured by the CAEB (Connotative Aspects of Epistemological Beliefs, Stahl & Bromme, 2007) – a semantic differential scale, made up of 17 pairs of items that measure beliefs about the source, simplicity and certainty of knowledge. The alpha reliability coefficient of this instrument was .72.

Argumentative reasoning skills were measured by the Argumentative Reasoning Task (Neuman, 2003), which introduces six stories to measure skill in identifying argumentative fallacies. The alpha reliability coefficient of this instrument was .72.

On the basis of the medians, scores for each individual difference were dichotomised to obtain two subgroups: one higher and the other lower for prior knowledge, epistemic beliefs about biology and argumentative reasoning skills.

Regarding eye measures, the following indices were analysed for each area of interest within each page: (1) *total fixation time* (TFT in seconds) as the summed duration of all fixations on an area of interest (this is an indicator of the total amount of attention allocated on an area, which reflects the 'duration' of elaboration on it), and (2) *mean fixation time* (MFT in milliseconds) as the ratio between the TFT on an area of interest and the number of fixations on it. It therefore reflects the mean 'duration' of elaboration on the attended area. The five areas of interest, which corresponded to the five types of information provided on each page, varied for size. Therefore, measurement of the MFT was necessary to compare the visual behaviour on the five areas, and to compare (in the first analyses, see below) attention allocation on information located on different page positions (top and bottom), which could be more or less attended.

Procedure

In a departmental laboratory equipped with Tobii 1750, participants were first asked to indicate their gender and age, and administered the CAEB and

Argumentative Reasoning Task via computer. They were then presented with the sequence of the four web pages to be read in a randomised order. After the execution of this task, they were asked to indicate the grade received in the biology exam.

Results

Main findings are presented in relation to each research question.

Attention allocation and source authoritativeness

We first examined attention allocation *within* each source by focusing on the five different types of information that appeared on it. Attention allocation *across* the different sources was then examined by focusing on the four sources varying in authoritativeness.

Attention allocation within each source

An ANOVA for repeated measures with the five types of information (headline, text 1, picture 1, text 2 and picture 2) as independent variables and MFT as the dependent variable was performed separately for each of the four pages. Significant differences emerged for the allocation of visual attention in relation to the type of information within the institutional, more authoritative page, $F(2.589, 93.193) = 5.26$, $p < .05$, $\eta^2 = .13$ (Greenhouse-Geisser), as well as within the encyclopaedic, more well-known page, $F(1.882, 65.609) = 6.65$, $p < .01$, $\eta^2 = .16$ (Greenhouse-Geisser). For both these pages, *post hoc* analyses[1] revealed that text 1 was fixated for a longer mean time than text 2 (Figure 3.2).

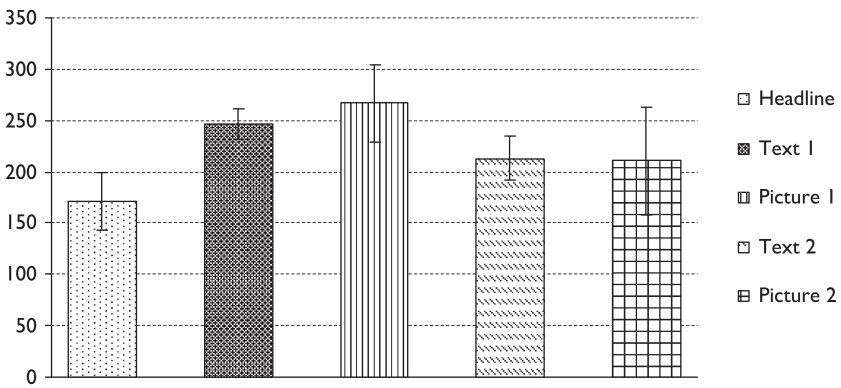


Figure 3.2 Mean fixation time (milliseconds) by type of information on the institutional web page.

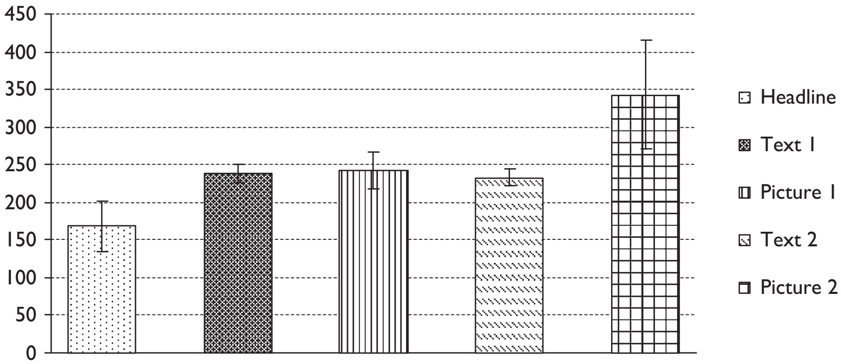


Figure 3.3 Mean fixation time (milliseconds) by type of information on the alternative web page.

Significant differences also emerged within the alternative, least reliable source, $F(1.682, 60.568) = 13.05$, $p < .05$, $\eta^2 = .27$ (Greenhouse-Geisser). *Post hoc* analyses showed that picture 2 was fixated for a longer mean time than picture 1 for that source (Figure 3.3).

Attention allocation across sources

An ANOVA for repeated measures with the four web sources (institutional, encyclopaedic, popular and alternative) as independent variables and TFT and MFT as dependent variables was performed for each of the five types of information. Significant differences emerged for the allocation of visual attention in relation to the source authoritativeness when considering text 2, $F(3, 108 = 3.38$, $p < .05$, $\eta^2 = .09$.

As revealed by *post hoc* analyses, the alternative, least credible page received a longer TFT ($M = 17.43$, $SD = 5.48$) than the institutional, more credible page ($M = 13.78$, $SD = 6.28$). Moreover, visual attention was also allocated differently on picture 1, $F(3, 108) = 2.91$, $p < .05$, $\eta^2 = .08$.

Specifically, the MFT on this picture was longer when the encyclopaedic ($M = 278.90$, $SD = 119.45$) and scientific sources ($M = 267.12$, $SD = 112.50$) were read and shorter when the popular ($M = 237.64$, $SD = 71.05$) and alternative pages ($M = 241.52$, $SD = 72.93$) were read. The opposite pattern of visual attention emerged for picture 2, $F(2.279, 82.045) = 8.23$, $p < .001$, $\eta^2 = .19$ (Greenhouse-Geisser). *Post hoc* analyses revealed that it was fixated for a longer mean time when the alternative ($M = 342.69$, $SD = 216.34$), unreliable page was read, while it was fixated for a shorter mean time during the reading of the institutional ($M = 210.97$, $SD = 157.72$), more reliable source. Significant differences related to the type of

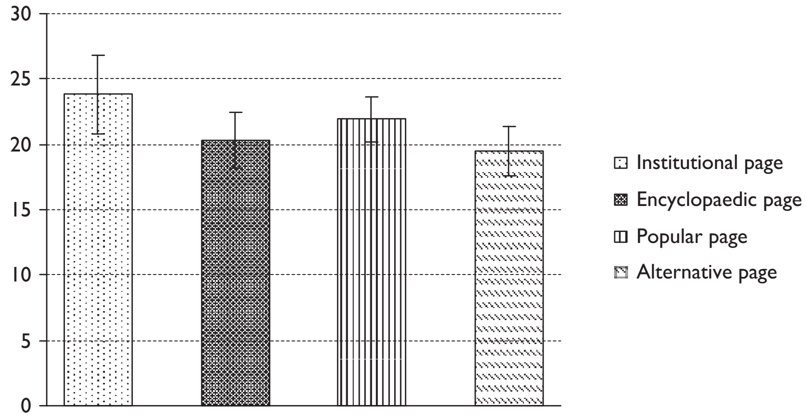


Figure 3.4 Total fixation time (seconds) on argument 1 by web page.

source also emerged for argument 1 (both text and picture) about the universal validity of the central dogma of biology, $F(3, 108) = 3.84$, $p < .05$, $\eta^2 = .10$. *Post hoc* analyses showed that the institutional page was fixated for a longer total time than the alternative page (Figure 3.4).

When argument 2 was considered, significant differences again emerged, $F(3, 108) = 4.17$, $p < .05$, $\eta^2 = .10$, but an opposing pattern of visual attention was identified compared with argument 1 (Figure 3.5). No significant differences emerged for visual attention allocated to page headings.

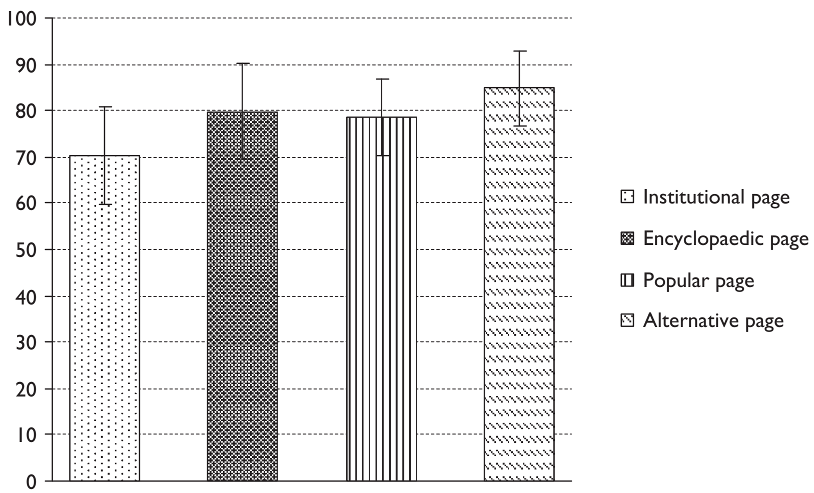


Figure 3.5 Total fixation time (seconds) on argument 2 by web page.

Attention allocation and individual differences

Preliminary analyses revealed a significant difference in prior knowledge between biology ($M = 23.35$, $SD = 4.06$) and psychology students ($M = 20.71$, $SD = 2.95$), $t(35) = 2.28$, $p < .05$. In contrast, no differences emerged regarding their epistemic beliefs about biology and argumentative reasoning skills. When considering the role of the latter characteristics, therefore, we took into account the former as a covariate in the statistical analyses.

PRIOR KNOWLEDGE

An ANOVA revealed that what students already knew about the transfer of genetic information influenced the MFT on all argument 1 through the four pages, $F(1, 35) = 7.01$, $p < .05$, $\eta^2 = .17$.

Lower prior knowledge participants ($M = 270.75$, $SD = 45.70$) allocated more attention to the more solid and traditional information in all pages than higher prior knowledge participants ($M = 233.40$, $SD = 40.10$).

EPISTEMIC BELIEFS

From an ANCOVA (with prior knowledge as covariate), it emerged that students with less sophisticated beliefs about the biological domain ($M = 262.40$, $SE = 8.77$) fixated text 1 for a longer mean time on all pages than students with more advanced representations about biology knowledge ($M = 230.09$, $SE = 8.05$), $F(1, 34) = 7.01$, $p < .05$, $\eta^2 = .17$. It is interesting to note that the latter ($M = 3.87$, $SE = .7$) dedicated more time to picture 2 than the former ($M = 1.71$, $SE = .7$) only within the institutional, reliable source, $F(1, 34) = 4.21$, $p < .05$, $\eta^2 = .11$.

In contrast, readers with less sophisticated beliefs about biological knowledge ($M = 261.97$, $SE = 10.52$) spent on average more time than readers with less sophisticated epistemic beliefs about biology ($M = 231.50$, $SE = 9.66$) reading the encyclopaedic, well-known page, idem $F(1, 34) = 4.33$, $p < .05$, $\eta^2 = .11$.

ARGUMENTATIVE REASONING SKILL

Another ANCOVA revealed that participants with greater skill in detecting informal reasoning fallacies ($M = 45.68$, $SE = 1.84$) allocated attention to the alternative page for longer than students less skilled ($M = 38.09$, $SE = 2.42$) in that respect, $F(1, 34) = 5.730$, $p < .05$, $\eta^2 = .14$.

Discussion and conclusion

The purpose of this study was to extend current research on the epistemic processing of online information by using eye tracking as a non-intrusive methodology that provides data as an indicator of cognitive processing during reading (Rayner, 1998). Eye-fixation measures can therefore track the evaluation of the trustworthiness of electronic resources and their contents. We were particularly interested in examining whether online information would be processed differently in relation to the authoritativeness of its source. According to the literature (Hofer, 2000, 2004), one essential dimension of epistemic thinking concerns beliefs about the source of knowledge.

Our first research question asked whether students allocated visual attention differently *within* each page and *across* web pages on the basis of source authoritativeness. As hypothesised, findings indicated that *within* each page, the attention paid to the various parts varied significantly. Specifically, attention patterns for text 1 arguing for the universal validity of the central dogma, and for text 2 arguing against it, were the same for the scientific, most reliable source, as well as for the well-known encyclopaedic source: text 1 had a longer fixation time than text 2. Another significant attentional variation was identified for pictures within the alternative web page, where picture 1 illustrating the ordinary genetic transmission was fixated for a shorter mean time than picture 2 illustrating the reversed genetic transfer. These outcomes indicate that students look for solid information when reading authoritative pages, and their classic information about the central dogma and its role in the basic biological processes of living beings was therefore elaborated with more attention.

In addition, findings substantially confirmed our hypothesis about opposing patterns of attention allocation *across* pages concerning the classic and the new argumentation about the universal validity of the central dogma in biology. The former was fixated for a longer total time on the institutional, more reliable source, the latter on the alternative, less trustworthy page. In particular, the picture illustrating the reverse genetic transfer received a greater amount of visual attention on the two sources with a lower epistemic status, while the picture showing well-known information about the unidirectionality of genetic transfer was fixated more in the two sources with a higher epistemic status. These findings lead us to infer that if the source in which the information is presented influences visual behaviour while reading, epistemic processing occurred. Information that developed the 'stable' and shared basis of the biological issue under examination was processed with more attention within an official and authoritative source. In contrast, 'alternative' information, which illustrates how the consolidated explicative model cannot account for new conflicting evidence, was processed with more attention within a non-official, non-consolidated source. These data are in line with those regarding spontaneous epistemic activation during web navigation as revealed by thinking aloud. Although at different levels of sophistication, university students were able to take source reliability into account when dealing with multiple, if not conflicting, documents about a debatable topic (Mason & Boldrin, 2008; Mason et al., 2010).

Our second research question asked whether individual differences influenced the allocation of attention while reading the four web pages. Findings substantially support our hypotheses as participants' examined characteristics, prior knowledge, epistemic beliefs about biology, and argumentative reasoning skill differentiated to some extent their visual behaviour in the expected direction. Specifically, in all pages less knowledgeable students needed to allocate more attention to the argument that transmitted the basic information about the transfer of genetic information than students with higher prior knowledge, who did not need to concentrate so much on the basic principles of the topic. The favouring role of prior knowledge also emerged in previous research about Internet searching for information (Kuiper et al., 2005).

As expected, participants with less advanced domain epistemic beliefs, that is, who believed more in biology knowledge as absolute, certain, and stable, fixated more the classic argumentation than participants believing more in biological knowledge as tentative, uncertain, and unstable. On average, lower epistemic sophistication led to spending more time reading the most well-known source, the encyclopaedic. Higher epistemic sophistication led to concentrating more on the picture representing reverse genetic transfer when it was presented on the institutional, authoritative source. Epistemically more sophisticated students, who believed in biological knowledge as uncertain and evolving, were more likely to pay attention to newer information, provided that it comes from an authoritative source. This outcome is in line with previous data about the role of epistemic beliefs in Internet-based learning environments (Bråten & Strømso, 2006), especially in processes of online information evaluation (Mason et al., 2010; Tsai, 2008).

Finally, participants who were more able to identify argumentative fallacies spent more time reading the alternative page than participants low in argumentative reasoning skills. It can be speculated that the former had a cognitive resource available allowing them to pay attention even to a scarcely reliable web page, as they were more able to evaluate the accuracy of information, regardless of the type of source in which it is presented. To some extent this finding calls up previous evidence regarding the positive correlation between epistemic belief and argument generation (Mason & Scirica, 2006) and evaluation (Weinstock et al., 2006).

In sum, this study shows first that external representations provided within a web page as information texts and pictures, receive different amounts of attention. Second, it indicates that the credibility of an information source influences reading: the same materials receive different attention according to the type of page in which they are presented. Third, individual differences play a role in epistemic processing of external representations on the Web.

The study, however, has two main limitations. First, the four web pages were presented in a linear sequence, with no possibility of returning to a previous page. In addition, each page was presented for a fixed reading time to avoid a time variable (which was not investigated), interfering with the task. Future research must overcome linearity in accessing the information sources and take into account the hypertextual nature of the Web. Moreover, reading time can be taken into account as a covariate if it is not fixed for all participants.

Second, the study did not examine participants' learning from the multiple sources. Future studies must reveal the relationship between ocular behaviour during reading and learning through web sources.

The study makes a contribution to scientific research by providing objective evidence that web pages are read differently by college students, who allocate different amounts of attention to various external representations according to type and source. From a methodological point a view, it documents that eye-movement measurement provides indices of processing that can be collected simultaneously during the execution of complex tasks, such as those investigated by instructional psychology. Analyses of eye-fixation patterns can therefore complement other online methodologies used to investigate attention allocation and processing demands during reading, such as thinking aloud, which has the limitation of being intrusive. Eye tracking does not disrupt normal reading and readers are not interrupted by a secondary task. Moreover, this methodology may provide input for instructional design, for example by uncovering experts' cognitive processes when they interact with multiple external representations, such as texts and pictures, both online and in print.

Finally, some educational implications about critical thinking in relation to online external representations can be drawn from our study. First, Internet reading requires epistemic evaluation to be able to differentiate multiple sources of information. This type of evaluation is sustained by individual characteristics, which can be improved dramatically by education. Research on multiple-document literacy has shown that higher-order processes and skills are involved in the construction of an integrated and meaningful representation of an issue, phenomenon or event (Bråten & Strømsø, 2010). These skills include being able to evaluate multiple sources of information (Bråten, 2008; Rouet, 2006), which are even more crucial when students surf the ocean of information available on the Web. Evaluation processes are insufficient if they concern only relevance to the topic of the information accessed. They should also regard source reliability and veracity of the information offered.

The second implication is that the Internet can be a powerful tool in helping novice students develop standards for evaluating online texts and other representations. Through the frequent experience of comparing perspectives concerning the same issue, the Internet provides the opportunity to reflect on the credibility of knowledge sources and the accuracy of information (Tsai, 2008). The Internet requires students to be able to evaluate sources without being overwhelmed, and, at the same time, it develops or refines their evaluative processes by requiring them to reflect on the nature, source and justification of knowledge, that is, to practise critical thinking.

The third implication is that instructional interventions need to be implemented to teach students to be epistemically active when navigating the Web, and equip them with important tools for becoming lifelong learners who use validated knowledge to make informed decisions. The new literacy skills for the Internet era imply critical thinking to transform web surfing into web navigation, where ongoing outcomes are monitored and judged with the aim of building new knowledge intentionally. In the use of the Internet as an infinite source of information for scientific learning, students' skills in formulating reasoned judgements about the credibility of websites and veracity of arguments can make a very great difference.

Note

1. All *post hoc* analyses mentioned in this session have been carried out with Bonferroni correction.

References

Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novices: Analysis of a complex cognitive skill. *Computers in Human Behavior, 21*, 487–508.

Bråten, I. (2008). Personal epistemology, understanding of multiple texts, and learning within Internet technologies. In M. S. Khine (Ed.), *Knowing, knowledge and beliefs: Epistemological studies across diverse cultures* (pp. 377–404). New York: Springer.

Bråten, I., & Strømsø, H. I. (2006). Epistemological beliefs, interest, and gender as predictors of Internet-based learning activities. *Computers in Human Behavior, 22*, 1027–1042.

Bråten, I., & Strømsø, H. I. (2010). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes, 47*, 1–31.

Brem, S. K., Russell, J., & Weems, L. (2001). Science on the Web: Student evaluations of scientific arguments. *Discourse Processes, 32*, 191–213.

Clark, D. B., & Slotta, J. D. (2000). Evaluating media-enhancement and source authority on the Internet: The Knowledge Integration Environment. *International Journal of Science Education, 22*, 859–871.

Crick, F. (1970). Central dogma of molecular biology. *Nature, 227*, 561–563.

Hirsh, S. (1999). Children's relevance criteria and information seeking on electronic resources. *Journal of the American Society for Information Science, 50*, 1265–1283.

Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology, 25*, 378–405.

Hofer, B. K. (2004). Epistemological understanding as a metacognitive process: Thinking aloud during online searching. *Educational Psychologist, 39*, 43–55.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*, 88–140.

Hyönä, J., Lorch, R. F., Jr,. & Kaakinen, J. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns, *Journal of Educational Psychology, 94*, 44–55.

Hyönä, J., Lorch, R.F., Jr., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313–334). Amsterdam: Elsevier Science.

Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction. 20*, 146–154. Just, M. A., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.

Kammerer, Y., Werner, B., & Gerjets, P. (2008, July). *What evaluation processes are applied during Web search? The impact of prompts to evaluate.* Paper presented at the XXIX International Congress of Psychology. Berlin, Germany.

King, P. A., & Kitchener, K. S. (1994). *Developing reflective judgement.* San Francisco, CA: Jossey-Bass.

King, P. A., & Kitchener, K. S. (2002). The reflective judgment model: Twenty years of research on epistemic cognition. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 37–61). Mahwah, NJ: Lawrence Erlbaum Associates.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science, 9*, 178–181.

Kuhn, D., & Weinstock, M. (2002). What is epistemological thinking and why does it matter? In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology. The psychology of beliefs about knowledge and knowing* (pp. 121–144). Mahwah, NJ: Erlbaum.

Kuiper, E., Volman, M., & Terwel, J. (2005). The web as an information resource in K-12 education: Strategies for supporting students in searching and processing information. *Review of Educational Research, 75*, 285–313.

Lin, C. C., & Tsai, C. C. (2007). A 'navigation flow map' method of representing students' searching behaviors and strategies on the Web. *Cyber Psychology and Behavior, 10*, 689–695.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4*, 6–14.

Mason, L., & Boldrin, A. (2008). Epistemic metacognition in the context of information searching on the Web. In M. S. Khine (Ed.), *Knowing, knowledge and beliefs: Epistemological studies across diverse cultures* (pp. 377–404). New York: Springer.

Mason, L., Boldrin, A., & Ariasi, N. (2010). Epistemic metacognition in context: Evaluating and learning online information. *Metacognition and Learning, 5*, 67–90.

Mason, L., Gava, M., & Boldrin, A. (2008). On warm conceptual change: The interplay of text, epistemological beliefs, and topic interest. *Journal of Educational Psychology, 100*, 291–309.

Mason, L., & Scirica, F. (2006). Prediction of students' argumentation skills about controversial topic by epistemological understanding. *Learning and Instruction, 16*, 492–509.

Meeuwen, L., Brand-Gruwel, S., & van Gog, T. (2008, July). *Surfing the ocean of information: Evaluation of information on the Internet.* Paper presented at the XXIX International Congress of Psychology. Berlin, Germany.

Neuman, Y. (2003). Go ahead, prove that God does not exist! On high school students' ability to deal with fallacious arguments. *Learning and Instruction, 13*, 367–380.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology, 96*, 538–558.

Rouet, J. F. (2006). *The skills of document use: From text comprehension to web-based learning*. Mahwah, NJ: Erlbaum.

Schacter, J., Chung, G., & Dorr, A. (1998). Children's Internet searching on complex problems: Performance and process analyses. *Journal of the American Society for Information Science, 49*, 840–849.

Stadtler, M., & Bromme, R. (2008). Effects of the metacognitive computer-tool met.a.ware on the web search of layperson. *Computers in Human Behavior, 24*, 716–737.

Stahl, E., & Bromme, R. (2007). An instrument for measuring connotative aspects of epistemological beliefs. *Learning and Instruction, 17*, 773–785.

Tu, Y. W., Shih, M., & Tsai, C. C (2008). Eighth graders' web searching strategies and outcomes: The role of task types, web experiences and epistemological beliefs. *Computers and Education, 51*, 1142–1153.

Tsai, C. C. (2004). Beyond cognitive and metacognitive tools: The use of Internet as an 'epistemological' tool for instruction. *British Journal of Educational Technology, 35*, 525–536.

Tsai, C. C. (2008). The use of Internet-based instruction for the development of epistemological beliefs: A case study in Taiwan. In M. S. Khine (Ed.), *Knowing, knowledge and beliefs*: *Epistemological studies across diverse cultures* (pp. 273–285). New York: Springer.

Tsai, C. C., & Chuang, S. C. (2005). The correlation between epistemological beliefs and preferences toward Internet-based learning environments. *British Journal of Educational Technology, 36*, 97–100.

van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20*, 95–99.

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning, 1*, 3–14.

Verschaffel, L., De Corte, E., & Pauwels, A. (1992). Solving compare problems: An eye movement test of Lewis and Mayer's consistency hypothesis. *Journal of Educational Psychology, 84*, 85–94.

Wang, J. C., & Day, R. F. (2007). The effects of attention inertia on advertisements on the WWW. *Computers in Human Behavior, 23*, 1390–1407.

Weinstock, M., Neuman, Y., & Glassner, A. (2006). Identification of informal Reasoning fallacies as a function of epistemological level, grade level, and cognitive ability. *Journal of Educational Psychology, 98*, 327–341.

Withmire, E. (2003). Epistemological beliefs and information-seeking behaviour of undergraduates. *Library & Information Science Research, 25*, 127–142.

Chapter 4

Representational flexibility in linear-function problems

A choice/no-choice study

Ana Acevedo Nistal, Wim Van Dooren, Geraldine Clarebout, Jan Elen, and Lieven Verschaffel

Theoretical background

Representational fluency and flexibility in mathematical problem solving

In the literature on mathematics education, it is often claimed that the ability to use external representations facilitates mathematical problem solving (Dienes, 1960; Even, 1998; Yerushalmy, 2006). The literature also mentions several skills that students need to possess in order to benefit from using external representations. These skills can be categorised in two groups.

The first group of skills can be called *representational fluency*. It involves the ability to interpret or construct representations (Ainsworth, Bibby, & Wood, 1998), as well as the ability to translate and switch between representations (on demand) accurately and quickly (Even, 1998). In sum, representational fluency refers to the efficiency (in terms of accuracy and speed) with which students can interpret, construct, translate, and switch between external representations.

The second group of skills involves making *appropriate* representational choices in a given problem-solving or learning situation. The way in which 'appropriateness' is conceptualised varies across studies. In some studies (e.g., Larkin & Simon, 1987; Schnotz & Bannert, 2003), a representational choice is considered appropriate if it is in line with the demands of the task at hand. In some studies from the strategy choice literature (e.g., Verschaffel et al., 2007), a choice is considered appropriate if it matches not only task demands, but also the characteristics of the subject that has to make the choice, and sometimes even the context in which the choice takes place. The main difference between these two conceptualisations is that the first one is (purely) based on a rational evaluation of the to-be-solved task, which results in a series of task-to-representation(s) matches which are expected to benefit the resolution of the task at hand, whereas the second conceptualisation also takes into account the individual's capacity to use the different representations and the context in which the choice takes place. What counts as an appropriate or flexible choice varies greatly across studies (for a review, see Warner, 2005), but for the purpose of this study we will focus on the two conceptualisations described above.

In this chapter, students' ability to make appropriate representational choices will be referred to as *representational adaptivity or flexibility*[1], and this skill will be the focus of this study. The first group of skills described above (and more specifically: representational efficiency) will be considered insofar as it relates to representational adaptivity (i.e. choosing the most appropriate representation may mean taking into account one's efficiency in using particular representations).

Representational choices in mathematical problem solving

Several studies (e.g., Dienes, 1960; Even, 1998; Kaput, 1992; Yerushalmy, 2006) have focused on the possible beneficial effects of creating, interpreting and combining multiple external representations when solving mathematical problems. However, research on students' ability to make appropriate representational choices is not as common. As Cox and Brna (1995) explain, 'to date, there has been much folk wisdom and speculation but little empirical work on the issue of representation selection' (p. 10). We consider this a topic of utmost importance, since merely confronting students with multiple representations does not ensure that problem solving will benefit. As Uesaka and Manalo (2006) showed, a problem solver confronted with a multitude of representations is often unable to decide which representation(s) to choose, and for what purposes. Interestingly, this inability to choose is not only present in students, but also in pre-service mathematics teachers (Even, 1990). Even found that, for example, in a problem that could easily be solved by means of a graph, 80 per cent of the prospective teachers used a formula and subsequently failed to reach the correct solution due to their inappropriate representational choice. Similar results were found by Mousoulides and Gagatsis (2004), who tested 95 pre-service mathematics teachers' understanding of functions. Their participants displayed a strong tendency to use an algebraic approach to the resolution of problems, even in cases where a geometric approach would have been easier and more efficient.

The purpose of our study was to observe and evaluate students' representational choices when solving linear-function problems. We focused on functions because this concept is considered by many (e.g., Eisenberg, 1992) as a key notion which supports the development of mathematical learning. Moreover, functions are a typical subdomain of mathematics where different representations (mainly tables, graphs, and formulae) can be used (Even, 1998). But our study also had a methodological purpose: we wanted to test whether a method that is used very frequently to study flexibility of strategy choices, the choice/no-choice method of Siegler and Lemaire (1997), could also shed some light on the complex processes that underlie representational flexibility.

Design of the study

Participants

Sixty-one Spanish secondary school students participated in the study. The 33 students from 4° ESO were aged 14–15, whereas the 28 students from 1° de Bachillerato were aged 15–16. For the sake of simplicity, the two age groups will be referred to as 10th-graders (14–15-year-olds) and 11th-graders (15–16-year-olds). Approximately half of the students were female.

Both age groups had prior knowledge about linear functions. In Spain, the topic is introduced in 9th grade, where students learn how to represent linear functions by means of formulae, tables, and graphs. In 10th grade, students are introduced to concepts such as slope, intercept, and intersection. Other types of functions such as quadratic, radical, and exponential are mentioned briefly at the end of 10th grade, but they are not studied in detail yet. In 11th grade, students get a quick review on linear functions at the beginning of the year, and the focus then moves to quadratic, exponential, rational, radical, trigonometrical and logarithmic functions.

Task

The problems used differed from each other in terms of two variables:

Variable 1: *Problem type*. One-third of the problems were related to the concept of slope (see Figures 4.1, 4.2, and 4.3), one-third to the concept of y-intercept (see Figure 4.4), and one-third to the concept of intersection between two functions (see Figure 4.5).

Variable 2: *Contextualisation*. Half of the problems were contextualised, the other half were decontextualised. Contextualised problems presented students with a story that ended with a question concerning the data provided in the problem (see Figures 4.1, 4.3, and 4.5). Decontextualised problems asked directly about a concrete mathematical concept (see Figures 4.2 and 4.4).

Method

The method used, namely the choice/no-choice method (Siegler & Lemaire, 1997), was borrowed from the strategy choice literature. This method has been applied in studies dealing with strategy choices in mathematical topics such as multiplication (Siegler & Lemaire, 1997), numerosity estimation (Luwel, Lemaire, & Verschaffel, 2005; Luwel et al., 2003), and addition and subtraction (Torbeyns, Verschaffel, & Ghesquière, 2004), but to our knowledge it has never been used before in research on representational choice.

The test administration was computer-based. Participants solved a series of problems under different conditions: one choice condition (C-condition) where they could choose either a table, a graph, or a formula to solve each problem,

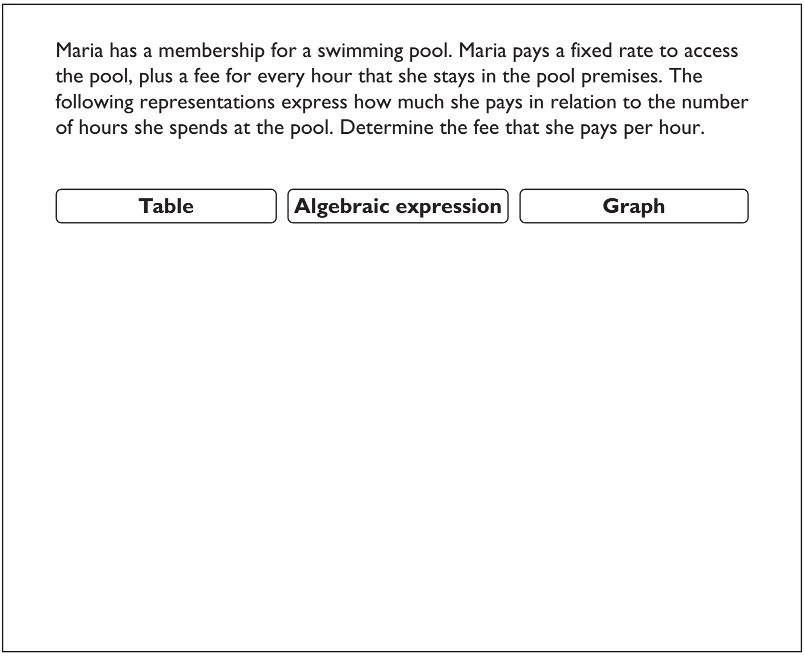


Figure 4.1 Contextualised slope problem in the C-condition.

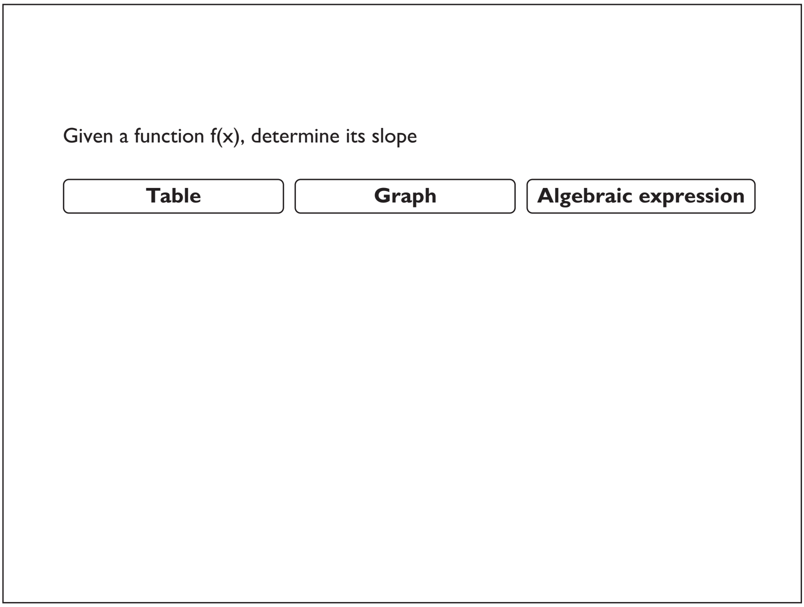


Figure 4.2 Decontextualised slope problem in the C-condition.

In a summer day, a meteorological station located near Madrid measures air temperature every hour for a whole morning. The following table expresses the registered temperature (in °C) in relation to the time elapsed (in hours) since the temperature started being measured. Determine by how many degrees the temperature increases per hour.

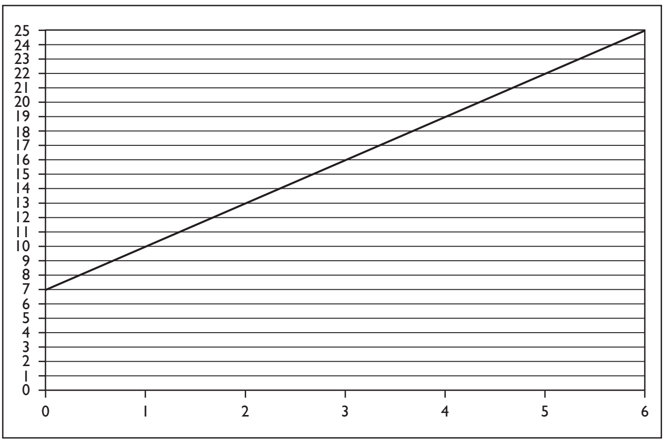
Hours	°C
0	7
1	9
2	11
3	13
4	15
5	17
6	19

The temperature increases [ ] °C per hour

Next

Figure 4.3 Contextualised slope problem from the NC-table condition.

Determine the intercept of f(x)



The intercept of f(x) is [ ]

Next

Figure 4.4 Decontextualised intercept problem from the NC-graph condition.

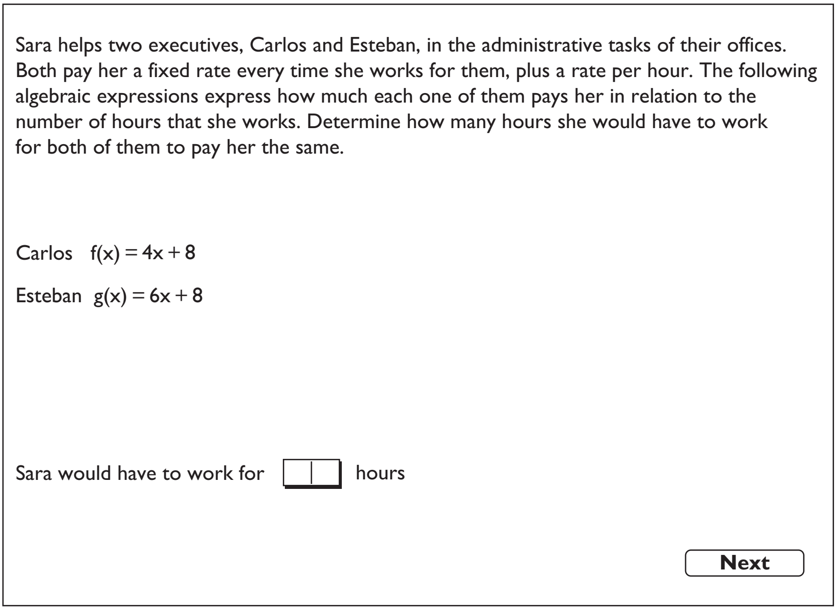


Figure 4.5 Contextualised intersection problem from the NC-formula condition.

and three no-choice conditions (NC-conditions), namely a NC-table condition, a NC-graph condition, and a NC-formula condition, where students were forced to use only one of these three representations to solve all the problems. The problems in the C- and NC-conditions were parallel.

In the C-condition, students received the problem wording together with three clickable buttons (‘table’, ‘algebraic expression’, and ‘graph’, see Figures 4.1 and 4.2), which appeared in random order on the screen. To ensure that students really used the selected representation, data were only provided in the representation and not in the problem wording itself.

In the NC-conditions, the problem wording was immediately followed by a predetermined representation: a table, a graph, or a formula, depending on which NC-condition the problem belonged to (see Figures 4.3, 4.4, and 4.5).

Procedure

As suggested by Siegler and Lemaire (1997), the students were exposed to the C-condition first, to avoid carry-over and recency effects from the NC- to the C-condition with regard to representational choice. The NC-test was administered four days later. Within each condition, half of the students received the contextualised problems first, and the others received the decontextualised problems first. Slope, intercept and intersection problems appeared in random order.

Students were instructed to solve the problems as accurately and as quickly as possible. For each student, the program (Macromedia Director Shockwave Studio 8.5, educational version) logged the following information:

- Accuracy: The answers to all the problems in the C- and NC-conditions were logged and coded as correct (1) or incorrect (0).
- Speed: The program logged the elapsed time from the moment a student clicked on a button to access a problem to the moment he/she made his/her representational choice. The total time spent on a problem was also logged.

Analysis

From the NC-conditions, we analysed the efficiency (in terms of accuracy and speed) with which the students used the different representations, as well as the influence of grade, problem type, contextualisation, and type of representation on this efficiency.

From the C-condition, we analysed the frequency with which each representation was chosen. The impact of grade, problem type, and contextualisation on these choices was also analysed. Students' representational choices were examined in detail to determine if choices were made in a flexible way. Two different conceptualisations and operationalisations of flexibility were adopted: a task-based flexibility and a task × student-based flexibility.

Task-based flexibility

According to researchers such as Cox (1996) and Larkin and Simon (1987), students' ability to successfully interact with external representations depends highly on their ability to identify the representation that matches the demands of the task at hand. As Gilmore and Green (1984) argue in their match–mismatch conjecture, performance is likely to be benefitted if the representation chosen to solve a particular problem matches the requirements of the problem at hand. Hereunder we briefly describe what such matches/mismatches between tasks and representations could look like for the tasks and representations used in our study. A rational task analysis is used for this purpose.

With regard to the variable problem type, a match could be argued between the formula and slope problems. In our study, all formulae were provided in slope/intercept form ($f(x) = mx + b$), which meant that the slope of the functions could easily be read off from the provided formulae. Solving slope problems with tables and graphs requires at least some calculations, and thus slope problems solved with tables and graphs can be expected to require more time and to be more error-prone than problems solved using a formula.

With regard to intercept problems, two possible matches were considered: on the one hand, one could argue that tables and graphs are a match to intercept problems, since the term 'intercept' translates into Spanish as 'ordenada en el origen'. The fact that the word 'origin' is mentioned in the term itself could lead students to the (0, *y*) value in the table or graph, thus obtaining a correct answer more straightforwardly than with a formula. It is important to note that all tables and graphs in our study explicitly included the intercept point.

On the other hand, since all formulae were provided in slope/intercept form, students might find it easier to read off the intercept from the formulae. If this was the case, then formulae could be considered a match to intercept problems.

Tables and graphs could be argued to be the best matches to intersection problems, since in all problems the intersection point was given both in the tables and graphs provided. Using a formula to find the intersection between two functions requires either extensive algebraic work or a laborious trial-and-error approach.

Contextualisation was expected to act as a mediating variable which could influence students' representational choices. Based on the results of Koedinger and Nathan (2004), contextualised problems were expected to prompt students to use more informal representations such as tables and graphs, whereas decontextualised problems were expected to elicit the use of more formal/symbolic representations such as formulae.

As has been pointed out, the representations provided to students to solve the problems in this study had specific characteristics: all formulae were in slope/intercept form, and all tables and graphs explicitly displayed the intercept or intersection points needed to solve the problems. These aspects were controlled for reasons of internal validity. The implication that follows is that the matches described above might only be applicable to these specific problems, and thus restrict the external validity of the results.

Task × student-based flexibility

While the first conceptualisation and operationalisation of flexibility was based on a rational evaluation of the characteristics of and relationships between the different problems and representations, the second conceptualisation also takes into account students' ability to interact with the different representations. This conceptualisation is based on Siegler and Lemaire's (1997) model of strategy choice, as well as on Verschaffel et al.'s (2007) conceptualisation of adaptive strategy choice as 'the conscious or unconscious selection and use of the most appropriate solution strategy on a given mathematical problem, for a given individual, in a given sociocultural context' (p. 19). As opposed to the first conceptualisation and operationalisation of flexibility, which is purely task-based, flexibility here is understood as a student's ability to make a representational choice which is in line not only with task characteristics, but also with student characteristics. More specifically, an individual student is said to make a flexible representational choice if he/she chooses in the C-condition the representation that, *according to the student's NC-data*, is the most likely to yield the most efficient answer for that particular student in that particular problem.

Results

Representational efficiency

Our approach to the study of representational efficiency involved analysing the accuracy and speed data yielded from the NC-conditions. This decision was based on Siegler and Lemaire's (1997) claim that the data in the NC-conditions yield unbiased estimates of students' fluency with the different representations, since all students are forced to solve all problems (or parallel problems) with all the available representations.

Accuracy data

A repeated-measures logistic regression analysis was carried out to determine which variables had an effect on students' performance in the NC-conditions. The dependent variable was NC-accuracy. The independent variables were grade, representation, problem type, contextualisation, and the interactions between them.

Grade did not have a main effect on accuracy – both grades performed at a comparable level in the NC-conditions (10th-graders: 59 per cent correct; 11th-graders: 64 per cent correct). However, the remaining three independent variables did have a main effect: representation, $\chi^2(2, N = 55) = 17.84$, $p < .01$, problem type, $\chi^2(2, N = 54) = 7.24$, $p = .03$, and contextualisation, $\chi^2(1, N = 54) = 24.73$, $p < .01$.

For the representation variable, pair-wise comparisons[2] showed that formulae yielded the lowest accuracy of all, whereas tables and graphs yielded higher (and similar) accuracies (formula: 51 per cent; table: 64 per cent; graph 68 per cent).

For problem type, the lowest accuracy corresponded to slope problems, whereas accuracy in intercept and intersection problems was significantly higher (slope: 55 per cent; intercept: 65 per cent; intersection: 64 per cent).

There was also a significant interaction between problem type and grade, $\chi^2(2, N = 54) = 6.65$, $p = .04$. Both grades performed similarly in slope and intercept problems, but 11th-graders were significantly better at solving intersection problems (see Table 4.1).

Table 4.1 Accuracy data from the NC-conditions by problem type and grade.

	Slope	*Intercept*	*Intersection*
10th grade	.57	.63	.56
11th grade	.53	.67	.71

With regard to contextualisation, contextualised problems yielded significantly higher accuracy (70 per cent correct) than decontextualised problems (52 per cent correct).

A significant interaction between problem type and representation was found, $\chi^2(4, N=54)=12.42$, $p=.01$. Pair-wise comparisons showed that, for intercept problems, graphs yielded higher accuracy than tables and formulae. In slope and intersection problems, graphs and tables yielded higher accuracies than formulae (see Table 4.2).

A significant interaction between representation and contextualisation was also found, $\chi^2(2, N=54)=27.42$, $p<.01$. In contextualised problems, tables and graphs yielded higher accuracies than formulae. In decontextualised problems, the three representations yielded comparable accuracies (see Table 4.3).

Speed data

A repeated-measures linear regression analysis was carried out to determine which variables affected the speed with which students solved the problems in the NC-conditions. The dependent variable, NC-speed, was operationalised as the time elapsed from the moment a student started a new problem to the

Table 4.2 Accuracy data from the NC-conditions by problem type and representation.

	Slope	*Intercept*	*Intersection*
Table	.66	.60	.67
Graph	.56	.75	.71
Formula	.42	.59	.52

Table 4.3 Accuracy data from the NC-conditions by representation and contextualisation.

	Table	*Graph*	*Formula*
Contextualised	.79	.76	.50
Decontextualised	.47	.58	.51

moment that the student proceeded to the next one. The independent variables were grade, problem type, contextualisation, representation, NC-accuracy and the interactions between them.

Two variables did not have a main effect on NC-speed: grade (it took both grades approximately the same time to solve the problems in the NC-conditions – 10th-graders: 35s, 11th-graders: 31s) and NC-accuracy (correct answers were given in approximately the same time as incorrect answers – 31s and 35s respectively).

The remaining variables had a main effect on students' NC-speed: problem type, $\chi^2(2, N=54)=57.89$, $p<.01$, contextualisation, $\chi^2(1, N=54)=15.42$, $p<.01$, and representation, $\chi^2(2, N=54)=20.47$, $p<.01$.

With regard to problem type, intersection problems took significantly longer to solve (on average 48s, $SD=51$s) than intercept and slope problems (intercept: 25s, $SD=20$s; slope: 27s, $SD=22$s).

Decontextualised problems (28s, $SD=37$s) were solved significantly faster than contextualised problems (39s, $SD=33$s). This difference in speed is probably due to the fact that contextualised problems had longer wordings and thus took longer to read.

With regard to representation, the problems where students were forced to use a graph were solved significantly faster (27s, $SD=30$s) than the problems solved with tables (34s, $SD=33$s) or formulae (39s, $SD=42$s).

As explained above, grade did not have a main effect on NC-speed. However, the interaction between representation and grade was significant, $\chi^2(2, N=54)=6.88$, $p=.03$ (see Table 4.4).

For 10th-graders, the fastest representation was the graph, whereas tables and formulae were significantly slower. A considerable improvement in the speed of use of the table was observed from 10th to 11th grade. The 11th-graders were significantly faster at using tables than the 10th-graders, to the extent that the table becomes (together with the graph) one of the (two) fastest representations in 11th grade.

A significant interaction between problem type and representation was also observed, $\chi^2(4, N=54)=37.32$, $p<.01$ (see Table 4.5).

For slope problems, all representations yielded equal NC-speeds. For intercept problems, the graph was the fastest representation. For intersection problems, the graph and the table were the fastest representations.

Table 4.4 Mean speed (and SD) in seconds in the problems from the NC-conditions by representation and grade.

	Table	*Graph*	*Formula*
10th grade	40 (40)	28 (33)	38 (35)
11th grade	27 (20)	25 (25)	41 (49)

Table 4.5 Mean speed (and SD) in seconds in the problems from the NC-conditions by problem type and representation.

	Slope	*Intercept*	*Intersection*
Table	28 (18)	27 (20)	46 (48)
Graph	27 (27)	21 (14)	32 (41)
Formula	26 (21)	27 (24)	65 (59)

Representational flexibility

Frequency of representational choice

A *t*-test was conducted to find out whether students' overall representational choices differed from random. The results showed that tables (chosen in 32 per cent of the cases) were chosen at random level, $t(329) = 0.35$, $p = .72$, formulae (40 per cent) were chosen above random level, $t(329) = 2.47$, $p = .01$, and graphs (28 per cent) were chosen below random level, $t(329) = 2.34$, $p = .02$.

T-tests, furthermore, showed that the formula was chosen significantly more often than the table, $t(330) = 2.03$, $p = .02$, and the graph, $t(330) = 3.40$, $p < .01$, while the table and graph did not differ significantly from each other.

However, when the choice data were analysed per grade, only 10th-graders chose formulae (45 per cent) significantly more often than graphs (23 per cent), $t(346) = 4.41$, $p < .01$, and tables (32 per cent), $t(346) = 2.44$, $p < .01$.

The representational choices of 11th-graders were distributed equally among the three representations (table: 33 per cent; graph: 32 per cent; formula: 34 per cent).

Task-based flexibility

Students' representational choices were analysed per problem type in order to determine whether they were in line with the matches between problem tasks and representations suggested by our rational task analysis. The following significant differences in representational choice were identified:

In slope problems, formulae (44 per cent) and tables (38 per cent) were used significantly more often than graphs (18 per cent), $t(218) = 4.23$, $p < .01$ and $t(218) = 3.37$, $p < .01$, for formulae and tables respectively. Our rational task analysis suggested that slope problems might be solved more easily by means of formulae. Thus, the fact that many students chose the formula to solve slope problems could therefore be considered a flexible choice if one understands flexibility from the purely task-based perspective that was explained above.

In intercept problems, formulae (45 per cent) were chosen significantly more often than graphs (29 per cent), $t(218) = 2.40$, $p < .01$, and tables (26 per cent), $t(218) = 2.86$, $p < .01$, with no differences between the latter two. In our rational task analysis, one of the possibilities considered was that formulae could be a good match to intercept problems given the fact that all formulae were provided in slope/intercept form. The fact that the formula was the representation chosen the most to solve intercept problems by both grades could thus be considered a flexible choice from such a purely task-based perspective.

In intersection problems, there were no significant differences in the frequencies with which the three representations were chosen (table: 33 per cent; graph: 35 per cent; formula: 32 per cent). Our rational task analysis indicated that tables and graphs would be the best match for intersection problems, but students' choices did not reflect this fact.

In contextualised problems, the table (41 per cent) was the representation chosen the most, significantly more often than the formula (31 per cent), $t(328) = 1.84$, $p = .03$, and the graph (28 per cent), $t(328) = 2.33$, $p = .01$, with no differences between the latter two. In decontextualised problems, the formula (49 per cent) was chosen the most, significantly more often than the table (24 per cent), $t(328) = 4.83$, $p < .01$, and the graph (27 per cent), $t(328) = 4.30$, $p < .01$, with no differences between the latter two. These choices were in line with our rational task analysis, which means that students' choices in contextualised and decontextualised problems could be considered flexible from the purely task-based perspective elaborated above. Our results were very similar to those obtained by Koedinger and Nathan (2004), who found that students preferred to use more informal representations (such as tables and graphs) to solve contextualised problems and more formal representations (such as formulae) to solve decontextualised problems.

To sum up, when a purely task-based conceptualisation/operationalisation of flexibility was used, students' representational choices turned out to be mostly flexible. However, as we will see in the next section, a task × student conceptualisation of flexibility shows a quite different picture.

Task × student-based flexibility

Based on Siegler and Lemaire (1997), students' representational choices were compared with the performance that these representations yielded in the NC-conditions. As a first step, for each problem and on a group level, we computed the difference between the accuracy of a particular representation (e.g., the table) and the averaged accuracy of the other two representations (the formula and the graph). These differences were then correlated with the frequency of representational choice of the first representation (the table) for the same problems. In both grades, there was a significant negative correlation between these two measures (10th grade: $r = -.49$, $p = .03$; 11th grade: $r = -.055$, $p = .01$).

According to Siegler and Lemaire (1997), these results of a group-wise analysis should be interpreted as a sign of students' lack of flexibility, since most students failed to select the representations which, according to the NC-data, would have been the most effective for the problems in question.

As a second step, an individual score was calculated for each student based on Siegler and Lemaire's (1997) adaptivity score. This score is calculated combining individual data from the C- and NC-conditions. From the C-condition, the frequency of choice of each representation for a particular student (i.e. the percentage of times that the student used each representation in the C-condition) is computed (F_{C_table}, F_{C_graph}, and F_{C_form}). From the NC-condition, the accuracy yielded by the different representations is also computed (A_{NC_table}, A_{NC_graph}, $A_{NC\_formula}$). Siegler and Lemaire (1997) argue that, since the student is forced to use all representations to solve all problems, the NC-conditions yield unbiased estimates about students' fluency with the different representations.

Then, a simulated accuracy in the C-condition (A_{C_sim}) is calculated using the NC-accuracy of each representation weighed by its frequency of choice. The resulting formula is as follows:

$$A_{C\_sim} = (F_{C\_table} \times A_{NC\_table}) + (F_{C\_graph} \times A_{NC\_graph}) + (F_{C\_form} \times A_{NC\_form})$$

A_{C_sim} is the simulated accuracy that a particular student would have obtained if he/she had chosen each representation with the same frequency as he/she actually chose them in the C-condition, but applying them to problems randomly (i.e., without taking into account his/her own performance with each particular representation in the different problems in the NC-condition). The adaptivity score is calculated by subtracting A_{C_sim} from the student's actual performance in the C-condition (A_{C_actual}).

If $A_{C\_sim} \approx A_{C\_actual}$, then the flexibility score is close to 0, meaning that the student does not perform better in the C-condition than if he/she had made random representational choices (but at the same base rates for each representation). If $A_{C\_actual} > A_{C\_sim}$, then the flexibility score is positive, which means that the student made flexible representational choices and performed better than if he/she had chosen between the representations randomly. If $A_{C\_actual} < A_{C\_sim}$, then the flexibility score is negative, and this means that the student's representational choices were even less effective than if he/she had chosen randomly. In other words, the student's choices were not only inflexible – they were counterproductive.

There were no highly flexible or highly inflexible students in either of the grades. In 10th grade, the highest flexibility score was 0.33 and the lowest −0.38. In 11th grade, the highest score was 0.16 and the lowest −0.33.

In 10th grade, 8 out of 27 students obtained a positive score. These students performed 15 per cent higher in the C-condition than the students who obtained a negative flexibility score: students with a positive score on average got 3.8 out of 6 problems (63 per cent) correct; students with a negative score on average got 2.9 out of 6 problems (48 per cent) correct, but this difference was not significant, $t(23) = 1.41$, $p = .08$. The 11th-graders with a positive flexibility score (6 out of 25) displayed an average performance in the C-condition that was 16 per cent higher than the performance of students with a negative score. Those with a positive score on average got 4.0 out of 6 problems (66 per cent) correct; students with a negative score on average got 3.0 out of 6 problems (50 per cent) correct, and this difference in performance was significant, $t(20) = 1.7$, $p = .05$.

The mean flexibility score per grade was very close to 0 in both grades (10th grade $M = -0.08$, $SD = 0.18$; 11th grade $M = -0.09$, $SD = 0.14$). There was no significant improvement in flexibility from 10th to 11th grade, $t(50) = 0.21$, $p = .41$. We interpreted these results as a general lack of representational flexibility.

This general lack of representational flexibility also becomes apparent when the accuracy data from the C- and NC-conditions are compared. If the two grades had been adaptive, then they should have performed better in the C- than in the NC-conditions, since their adaptivity would have allowed them to choose the representations most likely to yield a correct answer for the problems at hand. However, in both grades, average performance in the NC-conditions was better than that in the C-condition (10th-graders, $t(26) = 1.76$, $p = 0.04$; 11th-graders, $t(24) = 2.47$, $p = 0.01$), meaning that, in general terms, students performed worse when they were given the choice to select a representation than when they were forced to solve each problem with a predetermined representation. This was probably due to their inability to make flexible representational choices.

Discussion

The main aim of this study was to examine students' ability to make flexible representational choices while solving linear-function problems. In addition, the potential of Siegler and Lemaire's (1997) choice/no-choice method to study representational flexibility was assessed.

Representational efficiency

Contrary to our expectations, grade did not have an effect on students' ability to interact with the representations. Since 11th-graders had been exposed to the concept of linear functions for a longer time, we expected them to be more accurate and faster at interacting with the different representations. However, this was not the case.

Certain representations were shown to facilitate the resolution of certain types of problems. Slope problems were solved more efficiently with graphs and tables, while the three representations yielded comparable speeds in this type of problem. Intercept problems were solved more accurately and also faster with graphs. Intersection problems were solved more accurately and faster with graphs and tables. With regard to contextualisation, tables, graphs and formulae yielded comparable accuracies in decontextualised problems, but contextualised problems were solved more accurately with graphs and tables.

Representational flexibility

In the C-condition, an interesting shift in frequency of representational choice was found from 10th to 11th grade. The 10th-graders displayed an overall preference towards using formulae. However, such preference was no longer present in 11th grade, where all representations were chosen in a similar number of occasions. Such a shift in representational preferences could have been due to instruction, since the books used in 10th grade (and presumably also instruction) focus on the use of formulae to introduce the concept of linear function, whereas in 11th grade the use of tables and graphs is also encouraged.

To what extent were the representational choices made by students flexible? Using a purely task-based approach to flexibility, the conclusion was that students' representational choices were mostly flexible in the sense that they were in line with the matches predicted by our rational task analysis. However, the accuracy data from the NC-conditions showed that students often selected representations that were neither the most effective nor the fastest for the to-be-solved problem. The fact that there was a negative correlation between the number of times that each representation was chosen to solve a particular problem in the C-condition and the performance that that representation yielded for the corresponding parallel problem in the NC-condition already hinted that students were not as flexible in their choices as the task-based measure of flexibility suggested. An understanding of flexibility based on a mere match between tasks and representations does not seem to capture the quintessence of flexibility, since it disregards certain variables (e.g., students' fluency with the different representations) that should also be taken into account when determining whether a student is able to make flexible choices.

Siegler and Lemaire's (1997) conceptualisation of adaptivity partly addresses this issue by bringing into the picture a subject variable, that is, students' representational efficiency, obtained from the data from the NC-conditions. The task × student-based conceptualisation allowed us to discover that the reason why most of the students performed better in a NC than in a C setting was because they were unable to make flexible representational choices. The few students who did make flexible representational choices displayed a somewhat higher overall performance in the C-condition than their inflexible counterparts.

Implications for research

One of the methods used to assess students' representational flexibility was Siegler and Lemaire's (1997) choice/no-choice method, a method which, to our knowledge, had never been applied to the field of mathematical representations. Contrary to the traditional, purely task-based understanding of flexibility, this method allowed us to obtain a more fine-grained view of students' flexibility by taking into account not only task, but also subject variables.

However, using the choice/no-choice method to study representational flexibility also had its flaws. One of them is related to the wide variety of flexibility profiles which are encapsulated under the 0 flexibility score. A student who obtains an overall performance of 0 both in the C- and in the NC-conditions always gets a 0 adaptivity score. A student who obtains an overall performance of 1 in both the C- and NC-conditions also gets a 0 adaptivity score, but the adaptivity profiles of these two students are very different: in the first case, no representational choice can be considered flexible, since none of the representations is bound to yield a correct answer. In the case of the second student, choosing any of the representations could be considered a flexible choice, since all the representations are bound to yield a correct answer.

The task × student flexibility score inspired from Siegler and Lemaire (1997) only took into account the accuracy data yielded by the NC-condition. Further insight into students' representational flexibility could be obtained by computing a flexibility score that combines both accuracy and speed data extracted from the NC-conditions. By doing so, flexibility could be conceptualised as a student's ability to select the representation that leads him/her *the fastest* to the correct answer for the problem at hand.

As explained previously, the representations that students used to solve the problems had very specific characteristics: all formulae were in slope/intercept form, and in all tables and graphs the intercept or intersection points were explicitly displayed. This was done to get higher internal validity for our study. The disadvantage, however, is that the observed trends may only apply to the specific set of problems and representations that we used. As such, this is not problematic, since our intention was mainly to show the theoretical and methodological potential and implications of various conceptions of flexibility in the use of representations, as well as testing the potential of the choice/no-choice method in this field. Nevertheless, in order to generalise the findings to the broader field of function problem solving and to derive instructional implications from it, it seems worthwhile to do further work using different problem types, or providing students with representations that are not always in the ideal format to solve the proposed problems.

Future research in the field of representational choice might also benefit from combining the data obtained from choice/no-choice studies with one-to-one interviews with students after the tests, where they are prompted to explain their choices. Encouraging explicit reflection on representational choices by means of interviews might help students to verbalise their criteria, thus providing a different perspective on the variables that regulate students' representational choices.

Implications for instruction

Both the 10th- and 11th-graders who participated in this study had been exposed to regular instruction about the topic of linear functions, but they had never received instruction specifically aimed at improving their flexibility. Judging by the lack of flexibility displayed by the majority of the students, the ability to choose flexibly does not seem to be a skill that develops spontaneously in the traditional mathematics class.

Our results show that representational flexibility is a very complex and nuanced concept: what is considered flexible for a particular problem and a particular student might be considered inflexible when a different problem and/or student are involved. Even when a simplistic, task-based-only approach to flexibility is used, it is still challenging to state which representational choices are flexible and which ones are not, since it is extremely difficult (if not impossible) to identify clear-cut links between problem types and representations which are guaranteed to favour performance. Due to the great number of variables that play a role in determining what can be considered a flexible representational choice, representational flexibility is a skill that is very difficult to teach directly and explicitly in the traditional mathematics classroom. Considerable research effort needs to be invested to design powerful learning environments which explicitly address representational flexibility if we want our students to be equipped to make appropriate representational choices.

Notes

1. As it can be noted, we use the terms 'representational flexibility' and 'representational adaptivity' as synonyms. We are aware of the fact that some authors do not consider these terms as synonyms (see Acevedo Nistal et al. (2009) for a discussion on this issue), but for the sake of simplicity we have decided to use them as such in this article. For an explanation regarding the different conceptualisations of these terms, we refer to Verschaffel, Luwel, Torbeyns, and Van Dooren (2009).
2. For purposes of readability, statistical details were only included in text for the main effects and interaction effects. For the numerous contrast analyses (e.g., the pairwise comparisons between grades, problem types, etc.), we only reported whether differences were statistically significant ($p < .05$) or not, without mentioning each time χ^2 scores, degrees of freedom and exact p-values.

References

Acevedo Nistal, A., Van Dooren, W., Clarebout, G., Elen, J., & Verschaffel, L. (2009). Conceptualising, investigating, and stimulating representational flexibility in mathematical problem solving and learning: A critical review. *ZDM – The International Journal on Mathematics Education, 41*, 627–636.

Cox, R. (1996). *Analytical reasoning with multiple external representations.* Unpublished PhD thesis, Department of Artificial Intelligence, University of Edinburgh.

Cox, R., & Brna, P. (1995). Supporting the use of external representations in problem solving: The need for flexible learning environments. *Journal of Artificial Intelligence in Education, 6*, 239–302.

Dienes, Z. P. (1960). *Building up mathematics.* London: Hutchinson Educational.

Eisenberg, T. (1992). On the development of a sense for functions. In E. Dubinsky & G. Harel (Eds.), *The concept of function: Aspects of epistemology and pedagogy* (pp. 153–174). Washington, DC: Mathematical Association of America.

Even, R. (1990). Subject matter knowledge for teaching and the case of functions. *Educational Studies in Mathematics, 21*, 521–544.

Even, R. (1998). Factors involved in linking representations of functions. *Journal of Mathematical Behaviour, 17*, 105–121.

Gilmore, D. J., & Green, T. R. G. (1984). Comprehension and recall of miniature programs. *International Journal of Man-Machine Studies, 21*(1), 31–48.

Kaput, J. (1992). Technology and mathematics education. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 515–556). New York: Macmillan.

Koedinger, K. R., & Nathan, J. N. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *The Journal of the Learning Sciences, 13*, 129–164.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65–100.

Luwel, K., Lemaire, P., & Verschaffel, L. (2005). Children's strategies in numerosity judgment. *Cognitive Development, 20*, 448–471.

Luwel, K., Verschaffel, L., Onghena, P., & De Corte, E. (2003). Strategic aspects of numerosity judgment: The effect of task characteristics. *Experimental Psychology, 50*, 63–75.

Mousoulides, N., & Gagatsis, A. (2004). Algebraic and geometric approach in function problem solving. In M. Johnsen Høines & A. Berit Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 385–392). Bergen, Norway: PME.

Siegler, R. S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General, 126*, 71–92.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction, 13*, 141–156.

Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2004). Strategic aspects of simple addition and subtraction: The influence of mathematical ability. *Learning and Instruction, 14*, 177–195.

Uesaka, Y., & Manalo, E. (2006). Active comparison as a means of promoting the development of abstract conditional knowledge and appropriate choice of diagrams in math word problem solving. In D. Barker-Plummer, R. Cox, & N. Swoboda (Eds.), *Diagrammatic representation and inference: 4th international conference, Diagrams 2006*, Stanford, CA, USA, June 28–30, 2006, Proceedings (pp. 181–195).

Verschaffel, L., Luwel, K., Torbeyns, J., & Van Dooren, W. (2009). Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education. *European Journal of Psychology of Education, 24*, 335–359.

Verschaffel, L., Torbeyns, J., De Smedt, B., Luwel, K., & Van Dooren, W. (2007). Strategy flexibility in children with low achievement in mathematics. *Educational & Child Psychology, 24*(2), 16–27.

Warner, L. B. (2005). *Behaviors that indicate mathematical flexible thought*. Unpublished doctoral dissertation, Rutgers, The State University of New Jersey, New Brunswick.

Yerushalmy, M. (2006). Slower algebra students meet faster tools: Solving algebra word problems with graphing software. *Journal for Research in Mathematics Education, 37*, 356–387.

Chapter 5

Representations and proof

The case of the Isis problem

Brian Greer, Dirk De Bock, and Wim Van Dooren

Introduction

The Isis problem, so named because of its link with the Isis cult, as recorded by the Greek author Plutarch, is very simple:

> What rectangles with integral sides (in some unit) have area and perimeter numerically equal? Prove the result.

It is relatively easy to find out that there are precisely two such rectangles, 4×4 and 3×6. (For convenience, they will be referred to hereafter as the 'Isis rectangles'.) The interest of the problem lies in the variety of representations that can be used to support forms of argument that are the bases of proofs, and also in the fact that the problem is accessible to students of almost all ages. The focus of this chapter is on the domain of mathematical proving, and, more specifically, on the crucial roles that (various) mathematical representations play in proving.

In a first section we set out a variety of approaches leading to proofs, showing thereby how proofs can rely on substantially different mathematical representations, each having distinctive affordances. In this section, we will also argue how representational flexibility (Acevedo-Nistal et al., 2009) can provide deeper insight into the problem.

In a second section, we report on an experimental study in which student teachers were asked to construct proofs themselves and to evaluate and comment upon a variety of given proofs. We consider how the results resonate with major theoretical themes in the psychology of mathematical thinking. Again, the issue of mathematical representations will come forward as a major theme, as student teachers seem to attribute differential status to the representational systems in which a mathematical proof is developed.

Finally, we summarise the main theoretical and pedagogical implications of our analysis and experimental findings.

Variety of approaches to a proof for the Isis problem

This section can be considered as a 'rational analysis' of the task showing the variety of proofs relying on different mathematical representations. As shown below, each proof uses a particular type of argumentation that seems intrinsically linked to the representational system in which it occurs. Given that different representational systems emphasise (and de-emphasise) different aspects of the situation represented (Lesh & Doerr, 2003), switching between different representations and linking insights from these representations may provide a unique, multifaceted understanding of the Isis problem that could not occur when only focusing on one proof or staying within one representation. We will first consider an empirical approach (grounded in a numerical/tabular representation), then several approaches within the algebraic representational system, and finally a geometrical approach.

Empirical

An approach that even young children could use is simply to check a large number of rectangles. Such exploration is likely to progress from generation of cases unsystematically to more systematic generation and recording of cases, for example, by keeping one dimension constant and increasing the other systematically. Representations may also progress from drawings of rectangles with dimensions marked (perhaps using squared paper), to use the formulae for area and perimeter, and thence to recording the results in a tabular representation. From systematically generated results, particularly when represented in a table, patterns and consequent hypotheses are likely to emerge, and, in the long run, the basis of a proof.

One possible path to a proof is the following. First, a two-dimensional table is constructed in which, for each combination of the dimensions of the rectangle, is recorded A – P where A is the area (numerically) and P the perimeter (numerically) (see Figure 5.1). The zeroes in this table correspond to the Isis rectangles. (Note that there are three, since in this representation 3×6 and 6×3 are not the same.) To complete a proof, it is necessary to show that infinite extension of the table will not result in any more zeroes. One general argument to establish that fact is to consider what happens when an $x \times y$ rectangle grows to an $(x + 1) \times y$ rectangle (the reader is invited to construct a diagram to represent this change). When this happens, the area increases by y and the perimeter by 2, so A – P increases by $y - 2$, which explains the arithmetic progressions in each row and column of the table. From this result, it is straightforward to construct a proof that there are no more Isis rectangles.

	length						
width	1	2	3	4	5	6	7
1	–3	–4	–5	–6	–7	–8	–9
2	–4	–4	–4	–4	–4	–4	–4
3	–5	–4	–3	–2	–1	0	1
4	–6	–4	–2	0	2	4	6
5	–7	–4	–1	2	5	8	11
6	–8	–4	0	4	8	12	16
7	–9	–4	1	6	11	16	21

Figure 5.1 Tabular representation of the numerical value of area–perimeter for different rectangles.

Algebraic

Anyone who has studied algebra is very likely to rely on a symbolic representation and respond to the problem by writing down an equation using the formulae for calculating the area and perimeter of a rectangle:

$$xy = 2x + 2y$$

(or its equivalent using different symbols). The problem then is to solve this as a Diophantine equation (i.e., an equation whose solution set is restricted to whole numbers). From any such equation, an infinite number of equivalent equations can be generated. When staying within the algebraic representational system, the 'trick' then is to find such equations that are useful in leading to a solution. Efforts by many people over many years have led to numerous ways of rewriting the 'core equation' that are amenable to arguments that are the bases for proofs. Here are some of them:

Completing the rectangle (by analogy with completing the square)

A standard move is to convert to an equation with zero on the right-hand side, thus:

$$xy - 2x - 2y = 0$$

By adding 4 to each side, the left-hand side can be factorised:

$$xy - 2x - 2y + 4 = 4$$
$$(x - 2)(y - 2) = 4$$

Since $x-2$ and $y-2$ are integers whose product is 4, there are only a few possibilities to consider and the proof follows quickly (here and elsewhere, details are left to the reader).

Writing one variable as a function of the other

Another standard step for an equation with two variables is to rewrite (if possible) in the form $y=f(x)$), i.e. y as a function of x. In this case, it becomes:

$$y=2x/(x-2) \text{ or } y=2+4/(x-2)$$

Whence:

(a) $x-2$ is a factor of $2x$, or, even more clearly, of 4, again leading to a small number of cases that can be checked one by one (proof by exhaustion);
(b) The equation can be recognised as that of a hyperbola. In this case, it may be convenient to switch to a graphical representation (Figure 5.2). When the graph is constructed, the question then becomes one of the finding which points on the hyperbola have coordinates that are natural numbers.

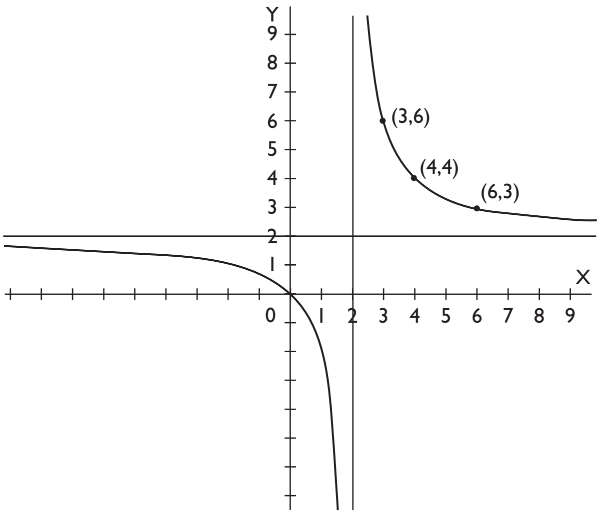


Figure 5.2 Hyperbola $y = 2x/(x - 2)$ showing the points on it whose coordinates are positive integers.

Unit fractions

By dividing each term by $2xy$, the equation becomes:

$$\frac{1}{2}=\frac{1}{x}+\frac{1}{y}$$

Now it can be argued that either $1/x$ and $1/y$ are both $1/4$ or one is less than $1/4$ and the other greater than $1/4$. Whichever is greater than $1/4$ must be $1/3$, $1/2$ or $1/1$ and, again, it is simple to check these three possibilities.

Harmonic mean

The equation can be rewritten as:

$$\frac{2xy}{x+y}=4$$

The expression on the left-hand side may be recognised as the formula for the harmonic mean of x and y. (The harmonic mean for two numbers arises in such situations as the following: if you travel from A to B at x miles per hour, and back from B to A at y miles per hour, the average speed for the whole journey is the harmonic mean of x and y.) If the harmonic (or any other) mean of two numbers is 4, it immediately follows that they are both 4 or one is greater than 4 and one less than 4. Hence to a proof, as above.

A 'silly' equation

Some rewritings of the equation might seem unpromising, such as:

$$yx+xy=4x+4y$$

Yet this equation implies that y and x cannot both be greater than 4, since if this were the case, yx would be greater than $4x$ and xy greater than $4y$, so the left-hand side would be greater.

Considerations of equality/inequality

If x and y are equal, then it is easy to see that $x=y=4$ is the only solution. Otherwise, assume x is less than y (a standard move in mathematical arguments justified by the symmetry of the equation, hence interchangeability of the symbols, and often introduced with the phrase 'without loss of generality'). Then $2x+2y<4y$, so if

$$xy=2x+2y, xy<4y, \text{ whence } x<4 \text{ (and so on).}$$

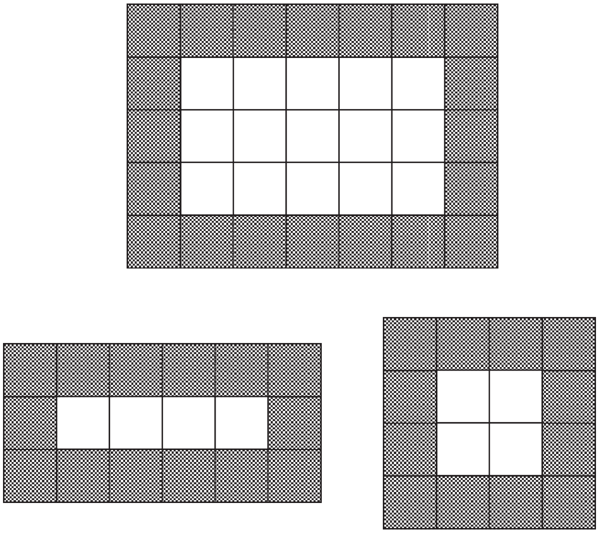


Figure 5.3 For any rectangle made up of unit squares, area of tiles around the perimeter = P – 4. For the Isis rectangles, the interior area = 4.

Geometrical

A radically different approach is based on the representation of the rectangle as a set of unit tiles (Figure 5.3). The number of tiles around the outside, as shaded in the figure, can be seen to be closely related to the length of the perimeter (P). Indeed, it is easy to see that the number of shaded tiles is P – 4 (the 4 coming from the 'overlaps' at the corners of the rectangle). But the number of shaded tiles is also equal to their area, which is part of the total area (A) of the rectangle. For A and P to be equal, therefore, the area of the interior, unshaded, tiles must be 4 to compensate for the 'missing' 4. Hence the interior must either be a 1×4 or 2×2 rectangle, leading to the 3×6 and 4×4 Isis rectangles, respectively (Figure 5.3).

Thus, while this representation of the rectangle as a set of tiles may first invite trial and error approaches (constructing several rectangles and checking whether the required property holds), it may develop from a series of empirical observations into a general argument which is the basis of a proof. Note also that the algebraic representation of the condition that the interior has area 4 is $(x - 2)(y - 2) = 4$, which was arrived at earlier through algebraic transformations.

An experimental study: student teachers' initial attempts to solve the problem and their ideas about proof

A study was carried out with three groups of students/future mathematics teachers to see, first, how they would approach the problem, working individually, and second, how they would react to and evaluate a variety of proofs. The groups were as follows:

- Group 1: future middle school teachers in a class at a West Coast US urban university ($N = 9$).
- Group 2: future lower high school teachers in two classes in the Flemish part of Belgium ($N = 23$).
- Group 3: future upper high school teachers in two classes in the Flemish part of Belgium ($N = 16$) who already had a degree in mathematics or were expected shortly to complete it.

All participants were given a task consisting of two parts and on each part they could work for one hour. First, they were invited to solve the Isis problem, including proving that there are only two Isis rectangles, and to look for more than one proof. Having finished that first part, the participants were invited to study five given proofs, to rank them in 'order of quality' (not defined by the researchers) from best to worst, and to comment on them. The five proofs were selected among the variety of proofs that are discussed in the previous sections, taking care that the various types of argument occurring in the different representational systems would all be included. The proofs, provided in counterbalanced orders, were: (1) the 'Table Proof' (see Figure 5.1), including an argument why extending the table will not produce any more zeroes, (2) the 'Graph Proof' in which the original equation $xy = 2x + 2y$ was rewritten as $y = 2 + 4/(x - 2)$, the corresponding hyperbola sketched (see Figure 5.2) and it was argued why (3, 6), (4, 4), and (6, 3) are the only three points with positive integer coordinates on that hyperbola, (3) the 'Unit Fractions Proof' starting from the equivalent form $1/x + 1/y = 1/2$, showing that either $x = y = 4$ or one of them is less than 4, then checking the limited number of possible cases, (4) the 'Factorisation Proof' starting from the $(x - 2)(y - 2) = 4$ equivalent form and, once more, checking a limited number of possible cases, and (5) the 'Tiles Proof', the geometrical proof explained above (Figure 5.3). In each case, the essence of the proof was given, with some logical steps 'telescoped', rather than a precise line by line argument.

Generation of proofs

Table 5.1 records the number of complete and partial proofs generated. A proof was judged partial when the essence of the proof was discernible but an essential element to complete the argument was missing. Just identifying the two Isis rectangles was not considered as a proof.

The predominant approach in *Group 1* was generation of a large number of examples, often sketching rectangles on squared paper. In several cases, this exploration shifted to more systematic recording of results in tables, using the

Table 5.1 Types of proofs found by the three groups (C = complete, P = partial).

	Group 1		*Group 2*		*Group 3*	
	C	*P*	*C*	*P*	*C*	*P*
Graph			1		3	
Factorisation					3	
Tiles					4	
Divisibility[a]			1	1	4	2
Exhaustion[b]			2	2		3
Other		1			5	
Total	0	1	4	3	19	5

[a]This refers to arguments based on $2x$ being divisible by $x - 2$, the natural number x being one dimension of the rectangle.
[b]This refers to arguments considering all the possibilities for $2x/(x - 2)$ being a positive integer.

formulae for area and perimeter. In a number of cases, regularities were noted in these tables. The closest to a proof came from a student who wrote as follows:

> There is going to be a limited amount of times where the area and the perimeter are equal. When you have $b \times h$ for the [area] the number is growing at a much larger rate than perimeter, which is growing by two. In order for the area and the perimeter to be equal you must wait for the numbers to cross.

The predominant approach in *Group 2* was to write the equation $xy = 2x + 2y$ (or equivalent) and work on manipulation of that equation. Three complete and three partial proofs emerged from finding the equivalent equation $y = 2x/(x - 2)$ and either arguing from this that $x - 2$ must divide $2x$, or by a method of exhaustive searching finding all natural number values of $2x/(x - 2)$. One proof was found by arguing from the graph of the hyperbola.

Not surprisingly, most of the complete proofs were found by students in *Group 3*. Like Group 2, their predominant approach was through manipulation of the starting equation, with more success. Moreover, there was also greater variation in the proofs found, as shown in Table 5.1. Individual differences were very marked, with 12 of the 19 proofs stemming from just three individuals.

One student in particular produced five different proofs, all coherently argued (Greer, De Bock, & Van Dooren, 2009). His fourth proof was particularly original and interesting. It started from the quadratic equation $X^2 - cX + 2c = 0$. If the roots of this equation (the values of X which satisfy the equation) are x and y, then $x + y = c$ and $xy = 2c$, so $xy = 2x + 2y$. The problem is now changed into that of finding for what values of c the given equation has roots that are natural numbers. The full details are not important; rather, the point is, as discussed later, that this creative act nicely illustrates the heuristic of 'think of a related problem' (Polya, 1957).

Evaluation of proofs

Table 5.2 lists the mean ratings for students' rankings of the five given proofs. The most striking aspect of these data is the marked difference between Group 1 and the other two groups. The students in Group 1 indicated they had much less fluency in algebraic manipulations, and many of them found some of the proofs hard to follow, or commented that steps were missing. Students in Groups 2 and 3 found the proofs easy to follow, but some also found them lacking in rigour (a reasonable comment).

Whereas average rankings (as well as typical comments) showed the preference of most students in Groups 2 and 3 for the algebraic proof by factorisation, this generally got low rankings from the students in Group 1. Several of them commented that it was not clear where the 4 came from (an interesting point, indeed). For example, one wrote 'why adding 4?' followed by 'Oh, to factorise . . . to easily factorise. You need to practice a lot to make this format. I couldn't see this.' The students in the other two groups appeared to follow the factorisation proof easily enough, and liked the fact that it led directly to the solution. For example, one wrote: 'In my view, the factorisation proof is the best because it is the least intuitive. Every step is mathematical' and another 'The proof by factorisation . . . is very tight, mathematically correct and easy to follow. Moreover, no tricks are used; the problem becomes easier by adding 4 on both sides.' Another student wrote: '+ 4 was a good trick! From then on, it is child's play.' A possibly relevant factor is the strong emphasis on factorisation in the Flemish curriculum (Vlaams Verbond van het Katholiek Secundair Onderwijs, 2002), and the general appreciation of future (secondary) mathematics teachers for algebraic approaches to mathematical problems as opposed to (systematic) numerical approaches that are considered as 'trial and error' (Van Dooren, Verschaffel, & Onghena, 2002).

The other case where Group 1 gave markedly lower evaluations was the Tiles Proof, and several remarked that the proof was not clear. Indeed, this proof

Table 5.2 Mean ranks for proofs, scored from 5 (best) to 1 (worst).

	Group 1	*Group 2*	*Group 3*	*Total*
Factorisation	3.4	4.0	4.4	4.0
Unit fractions	2.9	3.4	3.1	3.2
Tiles	2.2	3.5	3.0	3.1
Table	4.1	1.8	2.3	2.4
Graph	2.3	2.4	2.1	2.3

gave rise to the most varied and interesting responses. Reactions, as judged by comments from the Flemish students, were ambivalent, reflecting its perceived simplicity and elegance, yet somehow finding it not 'properly mathematical' because of the lack of algebraic expressions in the argumentation. For example: 'Proof with tiles: This is a better proof because it is clear and from the beginning till the end, it is neatly reasoned. Nevertheless, I miss some equations and it is a rather intuitive proof.' Others wrote: 'A much lower level and it lacks the real great mathematical thinking' and 'very clear and simple but not enough mathematics to be a real proof'. Others found its intuitive nature and clarity and the lack of involvement of algebraic expressions appealing: 'Even a ten years old child can understand it. You don't need any mathematical knowledge!'

Many of the Flemish students bracketed the Unit Fractions Proof with the Factorisation Proof as most convincing merely because of the representational system they rely on, i.e. algebraic expressions that are manipulated. Thus, 'you really prove it with factorisation and unit fractions', and 'both proofs are proofs in which algebraic manipulations are used, which I prefer because, in my view, it leads to the deepest insight'. Others valued the Unit Fractions Proof less. The proof is an example of 'proof by exhaustion', that is to say, showing that a finite number of possibilities exist (in this case three) and checking these one by one; in all groups, some of the students appeared to confuse this logic with a trial and error approach that can be used to find the solution of an algebraic equation without the need to manipulate the algebraic expression as such.

By contrast with the other groups, Group 1 ranked the Table Proof highest. The probable reason for this is that it was most accessible to them, in particular because they had, in class, been carrying out somewhat similar empirical investigations. The student who found the germ of a proof based on looking at rate of increase of area and perimeter wrote: 'I love this proof! I used [a] similar proof like this and it was easy for me to keep track visually.' On the other hand, many of the students in this group were convinced by studying the patterns in the table, but did not find clear the explanation of why there could be no further solutions. Among the Flemish students, the average ranking for the Table Proof was low and, in some cases, it was even rejected as a proof.

One reason for the low ranking of the Graph Proof was probably again the representational system it relied on. One response was: 'Proof with the graph: I find it very bad because no real mathematical arguments are used. It is just read off the drawing.' The reasoning needed to establish the limited number of points with natural number coordinates on the graph was also very sketchily presented in the material given to the students. It should, nevertheless, be noted that one subgroup of Group 2 (with six students) rated the Graph Proof relatively highly, which can be attributed to the fact that they were accustomed in class to using graphing software, and were allowed to, and did, use it while responding to the two parts of the study. The proofs collectively also elicited several emotional and aesthetic reactions, which not only related to the soundness of the arguments given in the proofs as such, but also to the representations relied on. Examples are: ‘Checking by trial and error which number can and which cannot [work] and I do not find this pleasant. It is indeed a proof, but it doesn’t look like it’, ‘The factorisation is very simple, clear and beautiful’, ‘The proof with the tiles comes over as a little bit playful’, ‘The proof with the tiles, I found by far the most creative and funny, but it does not fit very well with my perception of a mathematical proof’. One reaction pointed to the difference between proofs that convince logically and proofs that illuminate: ‘The proof with the tiles is the most visual one: you are not only convinced about the truth of the judgment, you also get the feeling that you “see” why it is so.’

Because of the limitations of the study, not too much weight should be put on the results presented here. However, it seems reasonable to claim that they establish the potential of the problem as a means of probing both students’ thinking in relation to a novel problem, and their conceptions of proof in various kinds of representations.

Links with theories of problem solving and proof

Here we again go through the empirical, algebraic and geometric approaches to the Isis problem, pointing out connections with important theoretical elements that are at the core of this book – the role of representation, interplay of routine and adaptive expertise, the search through a problem space, the importance of personal experience, the use of heuristics in the style of Polya.

An empirical approach to the Isis problem, as in many other cases, is good for generating a conjecture – virtually all the students tested found the two Isis rectangles. Mathematicians have always worked in this way and, with computers available, it is becoming even more common, and even changing the conception of proof, with a resurgence of ‘experimental mathematics’ (Hanna, 2007, pp. 6–7). The more difficult part is moving from the conjecture, via systematisation of the cases considered, to an argument that is the basis for a proof (Warner, Schorr, & Davis, 2009). It is also a subtle part of pedagogy to convince children that a conjecture that they believe to be true, and that has been verified for a great many cases, still requires a proof that applies to all cases. In the case of the Isis problem, one starting point for moving from conjecture to proof might be by the tabular representation suggested in Figure 5.1, a move that requires shifting from a simple noticing of the patterns in the table to a systematic analysis of what underlies those patterns, e.g. through a general argument about what happens when one dimension increases by 1 (see above).

The various approaches through the algebraic representational system remind us of a useful distinction made by Hatano (2003, p. xi), who defined routine expertise as ‘Simply being able to complete school exercises quickly and accurately without understanding’, whereas adaptive expertise means ‘the ability to apply meaningfully learned procedures flexibly and creatively'. The interplay between these two forms of expertise can be seen very clearly in how students approached the problem algebraically. The most obvious and prevalent form of routine expertise was to write the relationship between (numerical values of) area and perimeter as an algebraic equation. Almost all the students in all three groups wrote the equation at some point, though, as mentioned above, most in Group 1 were unable to develop the equation in useful ways. Having done this, the problem has been redefined as finding natural number values of x and y that satisfy $xy = 2x + 2y$. Given that the algebraic representation is constructed, solving the problem can be done without reference to rectangles, areas, and perimeters, exemplifying the power of symbolic algebra in general – particularly when the method for finding a solution of the equation set-up is known (routine knowledge):

> when we solve an algebraic equation which models a problem, we detach ourselves from the meaning of the symbols and their referents. Intermediate steps are usually not regarded as having meaning with reference to the situation, they are rather mechanical (and thus efficient) manipulations towards the solution.
>
> (Arcavi, 1994, p. 30)

From the starting equation a route can be traced through the space of equivalent equations generated by applying allowable transformations. How this works out for individuals will depend a lot on what is routine expertise for them. Davis (1985, pp. 89–95) presents a long and winding journey through many representations recorded by a mathematics teacher searching through the space of equivalent equations for the basis of a proof.

As was pointed out above, one routine step led to the equation:

$$y = \frac{2x}{x - 2}$$

Some may recognise this as the equation of a hyperbola (more familiar in the form $(x - 2)(y - 2) = 4$ and, if they have graphing facilities available, recur to the graphical representation (e.g. on a calculator). Others, as with several of the students in Groups 2 and 3, will stay within the algebraic representation, and recognise that for $2x/(x - 2)$ to be an integer (equal to y) $x - 2$ must divide exactly into $2x$, a tight constraint on x that can be used to find the possible values. A further simplifying step is to transform to:

$$y = 2 + \frac{4}{x - 2}$$

so that $x - 2$ must be a factor of 4, whence the proof follows very easily.

Other pieces that might be more or less routine expertise include the range of algebraic manipulations of the initial equation allowed, knowledge about inequalities, about the harmonic mean, and knowledge about unit fractions. As pointed out by Acevedo-Nistal et al. (2009), representational flexibility is heavily dependent on subject characteristics, in particular prior conceptual and procedural knowledge about representations, and domain-specific knowledge.

Using symbolic algebra is, thus, an almost routine, automatic response, given a certain level of experience with it, one that may or may not lead, and with more or less struggle, to a successful proof. However, as discussed by Arcavi (1994), there are cases where algebra may be invoked that are rather intractable, and a strategic shift to a very different kind of representation may be appropriate. With the Isis problem, a shift to a geometrical representation changes the nature of the problem and the form of the solution. When Brian Greer first told Bob Davis about what we now refer to as the Tiles Proof (which is certainly not original but which Brian Greer discovered independently) Davis was very excited because, he suggested, this was a line of argument available to the Egyptians, who had not developed symbolic algebra (Davis, 1993). Referring to a number of different solutions of the Isis problem that had been published in the *Journal of Mathematical Behavior*, Davis commented:

> [Greer] thought about the concrete square tiles . . . and thought directly in terms of these tiles. Clearly, the ancient Egyptians *could* have done that, and one feels it quite likely that they did. All of the rest of us moved immediately away from concrete representations of the problem, and dealt instead with abstract representations. This is characteristic of late-twentieth-century analysis. (p. 6).

In the various approaches to the problem, a number of examples of heuristic thinking in the sense of Polya (1957, 1962) have been observed. Davis (1988, p. 338) asked the question:

> How does . . . anyone . . . think of rewriting:
> $xy = 2(x + y)$
> in the form $(x - 2)(y - 2) = 4$?

Based on introspection, we can suggest two reasons. The form of the equation $xy - 2x - 2y = 0$ strongly cues the 4 that is missing (as suggested by Gestalt theorists, for example). So that answers the question raised by Davis (and by several students in the study). The 4 is needed to ‘complete the rectangle’.

How about the heuristic of considering special cases? A simple application leads to the result that the only Isis rectangle that is square is the 4×4 case.

More generally, holding one dimension constant while varying the other systematically forms the basis of an approach through considering rate of change, as was discussed earlier.

A heuristic that goes to the heart of mathematics is that phrased by Polya (1957, p. 110) as the deceptively simple question 'Have you seen it before?' For the Isis problem, this heuristic more specifically may be 'Where have I seen xy and $x+y$ related?' which may call to mind the quadratic equation $x^2 + ax + b$ in which the sum of the roots is $-a$ and the product of the roots is b (see the discussion above of one student who found a proof in this way). Another question that can be asked is 'Where else do the pairs of numbers (4, 4) and (3, 6) occur?' which leads to a very interesting problem devised by Lewis Carroll (discussed by Polya, 1962, p. 41) where precisely these pairs of numbers are involved – as two pairs of natural numbers with harmonic mean 4.

These considerations lead more generally to the topic of making connections in mathematics, which is taken up below within the discussion of educational implications.

Theoretical and pedagogical implications

In this chapter we have shown how the Isis problem embodies a web of deep mathematical ideas within a context that does not rely on complex technical mathematical knowledge. As such, it has much to offer both for the teaching of mathematics and as a conceptually rich example for the application of theories of problem solving, notably for probing the roles of representations.

Theoretical implications

Schnotz et al. (this volume) suggest ways in which perspectives of Gestalt psychology and information-processing psychology may be synthesised. Our analysis instantiates, and extends, this synthesis. The focus, from Gestalt psychology, on the centrality of the initial representation of the problem, is illustrated by the differences among the proofs, most markedly the Tiles Proof and the various proofs emanating from the equation $xy = 2x + 2y$. Development of an algebraic proof may be characterised, in part, as a search through a problem space of equivalent equations. However, there is more involved than finding a goal state within that problem space since there is no well-defined goal state. Rather, the solver has to recognise a state within the problem space as being tractable as the basis of a proof. As we have seen, both from the task analysis and from the work of a few exceptional students, there are many such states, and many forms of argument leading from them to proofs. Thus the algebraic proofs combine elements of searching a problem space with the need for insight. There is a further important element to this synthesis, however, namely that insight does not occur in a vacuum. We have documented, on the one hand, the roles of routine expertise and domain knowledge and, on the other, of flexible expertise through the application of heuristics in the tradition of Polya (1957, 1962), notably 'think of a related problem'. The analysis also illustrates the possible inhibiting effects of certain representations, notably in this case the algebraic, through processes akin to the Gestalt notion of functional fixedness (Schnotz et al., this volume) and Whitehead's (1929) notion of inert knowledge, as discussed below.

Schnotz et al. (this volume) also emphasise the contrast between a descriptive representation (based on symbols within a conventional symbol system that have no intrinsic relationship with the mathematical objects represented) and a depictive representation (which embodies some structural commonalities between the representation and the represented entity). The algebraic solutions for the Isis problem exemplify the former type, and the geometric solution the latter. There is also a link here with a distinction made in the literature on proofs between 'proofs that prove' (typically through a sequence of operations on a descriptive representation) and 'proofs that explain' (more usually associated with depictive representations).

Pedagogical implications

From the perspective of teaching mathematics, a virtue of the Isis problem we can point to is its interesting extensions. For example:

- What triangles with integer sides (in some unit) have the PANE property (Perimeter and Area Numerically Equal)?
- More generally, what plane figures have the PANE property?
- Ramping up a dimension, what cuboids (rectangular box-shaped solids) with integral dimensions have the property that the volume and surface area are numerically equal?

The quest for extensions provides an additional rationale for finding various alternative proofs in different representational systems: which proofs are adaptable to the new problem, and why are the other proofs not? The reader may like to consider which proofs for the Isis problem can be adapted for the three-dimensional analogue at the end of the above list.

A more radical extension starts from the observation that the Isis problem, and the extensions mentioned above, are of interest mathematically, at least as a kind of intellectual puzzle, but, as far as we know, do not have important applications – for a fundamental reason that, in itself, is extremely important. The opening statement of the problem carefully includes the words 'in some unit', 'numerically', and 'integral', thereby revealing the arbitrariness of the measurements and the sense in which an area and a length cannot be equal. The former can be seen in the general theorem that, for *any* plane figure for which area and perimeter are defined and finite, one and only one suitable change of the unit of measurement will give that figure the PANE property (and a similar theorem applies for three-dimensional figures in relation to volume and surface area). The problem throws light on fundamental questions about dimensionality. If the problem is extended to general issues about relationships across dimensions, then we enter a very important realm, with many biological and engineering applications (e.g. Galilei, 1638; Haldane, 1928). However, this discussion lies beyond the scope of this chapter.

As we have demonstrated through task analysis and through examining students' responses to the problem, the Isis problem is notable for the variety of different approaches to it. Relatively recently, we were alerted (by Abraham Arcavi) to a paper by Usiskin (1968, p. 388) called 'Six non-trivial equivalent problems' in which he makes exactly this point, commenting that: 'Equivalent problems demonstrate perhaps the most important quality of mathematics, the ability of one theoretical concept to be used as a model for many different ideas' (for an eloquent exposition of this position, see Poincaré, 1908) and arguing that suitable examples can be found to introduce understanding of this fundamental principle early in a child's mathematics education. In our experience, not enough emphasis is put in mathematics education on showing connections – in Poincaré's words 'mathematics is the art of giving the same name to different things' (and see Greer & Harel, 1998, p. 16).

For several of the proofs presented, it seems reasonable to assume that the students, in some sense, 'had' the necessary knowledge. That so few of them generated a proof within an hour (admittedly a short time to investigate an unfamiliar problem) suggests that their knowledge was 'inert' (Whitehead, 1929). Another explanation is that representing the problem by means of an algebraic equation cued familiar techniques, but not the different styles of argument needed to solve Diophantine equations. The students involved in the study probably did not very often (if ever) solve such equations. As an example of inert knowledge, Brian Greer remembers a discussion with a class of university students, some of whom had taken quite advanced mathematics classes. Having been presented with the equation $1/x + 1/y = 1/2$ they were at a loss how to proceed to find natural number solutions for x and y, and did not seem to find it natural to argue that either $1/x$ and $1/y$ are $1/4$ or one is greater than $1/4$ and the other less. It was as if the algebraic representation straitjacketed their thinking to known algebraic manipulations, so that they were inhibited of thinking of $1/x$ and $1/y$ as (positive) numbers whose sum is $1/2$ (so that both must equal $1/4$ or one must be more and one less, leading to a proof by exhaustion as discussed above). In other words, this could be characterised as an extreme case of the phenomenon Arcavi (1994) described as 'lack of symbol sense'.

Acknowledgement

This research was partially supported by grant GOA 2006/01 'Developing adaptive expertise in mathematics education' from the Research Fund K. U. Leuven, Belgium.

References

Acevedo-Nistal, A., Van Dooren, W., Clarebout, G., Elen, J., & Verschaffel, L. (2009). Conceptualising, investigating, and stimulating representational flexibility in mathematical problem solving and learning: A critical review. *ZDM – The International Journal on Mathematics Education, 41*, 627–636.

Arcavi, A. (1994). Symbol sense: Informal sense-making in formal mathematics. *For the Learning of Mathematics, 14*(3), 24–35.

Davis, R. B. (1985). The role of representations in problem solving: Case studies. *Journal of Mathematical Behavior, 4*, 85–97.

Davis, R. B. (1988). Contributions from Marion Walker and Murray Klamkin. *Journal of Mathematical Behavior, 7*, 337–338.

Davis, R. B. (1993). The Isis problem and the question of representations. *Journal of Mathematical Behavior, 12*, 3–8.

Galilei, G. (1638). *Discorsi e dimostrazioni matematiche intorna a due nuove scienze: 1954; Dialogues concerning two new sciences.* New York: Dover.

Greer, B., De Bock, D., & Van Dooren, W. (2009). The Isis problem as an experimental probe and teaching resource. *Journal of Mathematical Behavior, 28*, 237–246.

Greer, B., & Harel, G. (1998). The role of isomorphisms in mathematical cognition. *Journal of Mathematical Behavior, 17*, 5–24.

Haldane, J. B. S. (1928). *On being the right size in possible worlds.* New York: Books for Libraries Press (www.cmmp.ucl.ac.uk/~jlf/haldane_essay.html).

Hanna, G. (2007). The ongoing value of proof. In P. Boero (Ed.), *Theorems in school: From history, epistemology and cognition to classroom practice* (pp. 3–18). Rotterdam, The Netherlands: Sense Publishers.

Hatano, G. (2003). Foreword. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. xi–xiii). Mahwah, NJ: Erlbaum.

Lesh, R., & Doerr, H. M. (Eds.). (2003). *Beyond constructivism. Models and modeling perspectives on mathematical problem solving, learning and teaching.* Mahwah, NJ: Erlbaum.

Poincaré, H. (1908). The future of mathematics [English version translated from] *Revue Générale des Science Pures et Appliquées, 19th Year, No. 23*, Paris. Available at: http://home.att.net/~o.caimi/Poincare.pdf.

Polya, G. (1957). *How to solve it: A new aspect of mathematical method* (2nd ed.). Princeton, NJ: Princeton University Press.

Polya, G. (1962). *Mathematical discovery: Vol. 1.* New York: Wiley.

Schnotz, W., Baadte, C., Müller, A, & Rasch, R. (this volume). Creative thinking with depictive and descriptive representations.

Usiskin, Z. (1968). Six nontrivial equivalent problems. *Mathematics Teacher, 61*, 388–390.

Van Dooren, W., Verschaffel, L., & Onghena, P. (2002). The impact of pre-service teachers' content knowledge on their appreciation of pupils' strategies for solving arithmetic and algebra word problems. *Journal for Research in Mathematics Education, 33*, 319–351.

Vlaams Verbond van het Katholiek Secundair Onderwijs. (2002). *Leerplan Wiskunde voor de Tweede Graad van het Algemeen Secundair Onderwijs* [Mathematics syllabus

for Grades 9 and 10 of General Secondary Education]. Brussels, Belgium: Author.
Warner, L. B., Schorr, R. Y., & Davis, G. (2009). Flexible use of symbolic tools for problem solving, generalisation, and explanation. *ZDM – The International Journal on Mathematics Education, 41*, 663–679.
Whitehead, A. N. (1929). *The aims of education*. New York: Macmillan.

Chapter 6

Secondary school students' availability and activation of diagrammatic strategies for learning from texts

Michael Schneider, Catrin Rode, and Elsbeth Stern

Introduction

Visual-spatial knowledge representations such as matrices, Venn diagrams, hierarchical trees, and path diagrams are valuable tools for illustrating the gist of texts as well as complex quantitative data patterns. They make abstract concepts more concrete by mapping them onto spatial layouts with familiar interpretational conventions, and they clarify and highlight aspects of the problem that might otherwise be obscured by the text (Novick, 2001). Thus, using appropriate diagrams to represent information is an important tool which enables deeper understanding and facilitates problem solving in academic as well as non-academic domains.

Stern, Aprea, and Ebner (2003), for example, provided empirical evidence for the benefits of using spatial representations. They showed that graphical representation of the content of texts aids reasoning and transfer of the contents to new situations. Bauer and Johnson-Laird (1993) found that subjects responded faster and drew more valid conclusions in a deductive reasoning task when the premises were presented diagrammatically.

In order to provide students with proficiency in the use of all kinds of visual-spatial representations, diagrammatic literacy must go beyond simple diagram reading and construction skills. Students need to understand how to use diagrammatic representations as tools for thinking. Consequently, mathematics and science educational research in particular (Greeno & Hall, 1997; Lewis, 1989) has emphasised the need for students to learn to use such representations and to develop an understanding of the strengths and weaknesses of various representations for different purposes. Students should learn to actively create and adaptively use diagrams as tools for problem solving, as opposed to just reading provided diagrams in preconceived ways.

Despite the usefulness of diagrams, in Germany, where our study was conducted, as well as in many other countries, school instruction on diagram use has remained limited (e.g., Hardy et al. 2005; Mevarech & Kramarski, 1997).

Textbooks, for instance, are often supplemented with highly specific illustrations which support the understanding of a given problem but do not necessarily teach students how to generate diagrams to reconstruct or represent novel information. Moreover, diagrams used in media inside and outside of school often serve merely illustrative purposes. In many German classrooms, systematic instruction on graphical representational techniques hardly occurs at all, and, if it does, it is usually in restricted contexts (e.g., Cartesian graphs representing linear functions in mathematics classes). Typically, when solving a problem at school, students either are instructed to draw a pre-specified type of diagram or are given a diagram along with the problem. Thus, there is hardly any curriculum-based opportunity to learn how to translate a novel problem into a representation that captures the deep structure of the task at hand (Hardy et al., 2005).

In spite of the large body of literature indicating that graphs and diagrams can support and facilitate problem solving, there is considerably less research on how people actually use diagrams. Novick, Hurley, and Francis (1999) studied college students' ability to select the appropriate type of diagram from a set of alternatives and demonstrated at least some schematic knowledge about the conditions of applicability for particular spatial structures. However, other studies suggest that this knowledge is not used spontaneously. Schoolchildren's limitations in using graphical representations as thinking tools become apparent when they are presented with arithmetic word problems. Although graphical representations help considerably to highlight relevant aspects of the described situation, college students show very limited competencies in depicting appropriately the relevant information of word problems (Lewis, 1989). Moreover, Stern and Staub (2000) showed that even elementary school math teachers of relatively high-achieving classrooms failed completely in using appropriate spatial representations for word problems.

In and out of school, students encounter a broad variety of different diagrammatic representations such as graphs, circles, tables, matrices, tree diagrams, path diagrams, and so on. But to what extent do the learning conditions typically encountered by secondary school students enable them to create and use diagrammatic representations when faced with a new and complex problem which requires a diagrammatic representation? Tversky, Kugelmass, and Winter (1991) showed that even without systematic instruction, preschool children begin to use linear orderings for representing non-spatial information. This shows that even young children benefit from instructions on using space for non-spatial purposes. However, since using space to represent content is hardly ever practiced outside of limited context in math and science classrooms, students have few opportunities for strengthening this competence. Thus we expect serious deficits in diagrammatic competency to occur. These potential deficits could be attributed to at least two major causes: lack of availability or lack of activation.

Availability of diagrammatic strategies

The first possible cause of deficient diagram use is lack of availability. Students may lack the competencies for constructing diagrams because they do not know how to use space for non-spatial constructs. Since preschool students already use linear orderings spontaneously, a lack of availability can hardly be expected for one-dimensional representations. However, even college students show deficiencies in the use of representational forms that require the integration of two or more dimensions, such as clusters, crosstabs, matrices, complex path diagrams or tree diagrams (Novick & Hmelo, 1994). As such diagrams are used only rarely in school, students have to develop this knowledge by themselves on the basis of their background knowledge. Linear graphs are part of the curriculum in Grade Eight mathematics in Germany, where our study took place. This instruction on coordinate systems might help students of this age also to better understand other multidimensional diagrams used on other content domains. In this case, there should be improvements in the availability of strategies making use of two-dimensional diagrams from Grade Seven to Grade Nine.

Activation of diagrammatic strategies

The second possible cause for students not using diagrams is lack of activation. Students receive systematic instruction on verbal strategies in representing the gist of a text, such as underlining core information in a written text or excerpting keywords. Therefore, diagrammatic representation may be available in principle, but may not be activated when required. Students may rely on familiar verbal strategies even if it would be more advisable to represent the core information by a diagram. The persistence of strategies that are highly familiar but inefficient for solving the particular problem has been demonstrated repeatedly in various content areas and may be part of a larger, generally adaptive tendency to maintain cognitive variability (Siegler & Stern, 1998). Therefore we expect that even older secondary school students rely on verbal strategies to represent the gist of a text that could also be depicted in a diagram. Given the strong familiarity of verbal strategies, spontaneous use of spatial strategies can rarely be expected.

A lack of explicit knowledge about the potentials of diagrams might further contribute to this problem: learners may lack the knowledge that visuospatial representations can sometimes be more efficient than words. Due to this they do not consider using diagrammatic strategies, even if this method is highlighted in the instruction.

Students' competencies and deficits can be inferred by varying the degree to which the diagrammatic strategy is pointed out to them in the instructions for solving the problems. When students use the diagrammatic strategy, even in the absence of any instruction about it, they demonstrate a good availability and activation of this strategy. When students do not use the diagrammatic strategy, even after an intervention makes it available to them, only a low level of activation can account for the finding. On the other hand, when students who are asked explicitly to use diagrams still do not use them, then they are demonstrating a lack of availability.

Experiment 1: spontaneous and evoked diagram use

Experiment 1 consisted of two phases, a reading phase and a test phase. In the reading phase, the students were presented with six short texts. Students were asked to take down notes about the contents of the texts and were told that these notes would help them in the subsequent test phase. The instructions were given in German (as our participants were German). We were careful to choose a wording that did not suggest the exclusive use of text or the exclusive use of diagrams. In the test phase, students were allowed to use their notes, but not the original texts, to answer a number of questions about the texts. Each text described the relations between a set of items. These relations can be represented with keywords or, alternatively, with a diagram (i.e., hierarchical tree diagram, word clusters, map, or one-dimensional array).

The participants were randomised into two groups: the *control group* was asked only to take notes about the texts, without any guidelines about the nature of these notes. The *treatment group*, however, received a brief explanation of two useful strategies before the test: using keywords and using diagrams for summarising texts.

We hypothesised that students at both age levels would use the diagrammatic strategy only rarely in the control group but more frequently in the treatment group, where the availability of the strategy was increased due to the instruction.

Method

Participants

A total of 131 secondary school students participated in the experiment. The sample consisted of 60 seventh-graders (28 girls; one gender unknown) with a mean age of 12.7 years ($SD = 0.6$) and 71 ninth-graders (34 girls) with a mean age of 14.9 years ($SD = 0.6$). All students were recruited from the highest track (*Gymnasium*) of the German secondary school system, which is attended by one-third of children in that age group. Deficiencies in diagrammatic competencies found in this group of students can be assumed to be even larger in samples from the two lower educational tracks in Germany. Participants were compensated with 15 DM (approximately US$7.00). Students were recruited in their classrooms, but the study did not take place during class and participation by the students was optional.

Procedure

Groups of six to ten students were tested either in large seminar rooms in our institute or in a separate room in their school building. Only students of the same age group and the same instructional conditions were tested together. After a short introduction and an explanation of the procedure by the experimenter, participants were told that they would have to memorise information presented in written texts. They were further informed that they would be asked questions about the texts and would be allowed to consult their written notes when answering the questions. Next they received instructions according to the instructional condition to which they were assigned.

After this, booklets containing the texts and booklets for making the notes were distributed. Students worked through the task booklet at their own pace. When a student completed the task, the two booklets were collected so that none of the students would have extra time with the text. After a break, the booklets containing the notes were redistributed. Participants also received a booklet containing the questions on the texts. Students were allowed to take as much time as they wished to answer the questions. The entire procedure took approximately one hour.

Material

The stimulus material consisted of a booklet with six short written texts (see texts 1–6 listed in the Appendix), one on each page. Each text described relationships between seven or eight items. The six texts varied in their content and in the relational structure of the items described. The texts were constructed to allow for the representation of structural relationships between the items in the texts by using a spatial layout or by noting the keywords in a non-spatial way. For instance, the relationships in text 3 could be represented in a hierarchical tree diagram. This text described the structure of fictitious beings:

> On the faraway planet of Urx, living beings are called pings. There are two kinds of pings: spotted pings and striped pings. There are also two kinds of spotted pings: laughing pings and crying pings. Among the striped pings, there are the noisy ones and the quiet ones. Tip is a crying ping.

Each of the texts described non-spatial relationships that could optimally be represented in one type of diagram. The only exceptions were text 6, which described spatial relations, that is, the positions of seven buildings relative to each other, and text 2, which was entirely episodic, thus offering scant opportunity for spatial representation. These two texts served as manipulation checks, as text 6 was expected to evoke a high number of diagrams and text 2 a very low number of diagrams, assuming that the students adapted their strategy use to what they read. The remaining four texts required the use of space for representing non-spatial information. Two texts required the integration of more than one dimension: text 1 was about similarities and differences in the breeding behaviour of different types of frogs, suggesting a cluster representation. As indicated above, text 3 described a fictitious species which suggested a tree diagram as the most appropriate representation. Text 4, which was about preferences for different items of food, and text 5, which focused on temporal relations, both required one-dimensional arrays.

Students were provided with a second booklet for taking notes, with one page designated for each text. A third booklet contained questions about the six texts (see Appendix). Three questions were asked for each text. Two questions, the memory questions, referred to relationships explicitly stated in the text. The third question, the inference question, referred to a relationship that could only be inferred from the information given in the text, but was not explicitly stated there. The order of all texts was randomised for each student.

Experimental manipulation

Each student was randomly assigned to one of two instructional conditions. Participants in the free-choice condition were informed about the general task and procedure. They were encouraged to take notes in an efficient manner, but no suggestions were given as to what an efficient manner might be. This condition allowed us to observe the degree to which these students spontaneously used a diagrammatic strategy to summarise the texts. Participants in the keyword-or-diagram condition received the same information, but were also told about two different methods of note taking: the keyword strategy and the diagram strategy. In the *keyword strategy* the goal was to summarise the text into a list of keywords. In the *diagram strategy* the goal was to summarise the relations described in the text with a diagram. Two examples of diagrammatic representations were presented, although these examples were different from the diagrams most appropriate for summarising the texts presented in the experiment.

At the end of the instructions, the participants were asked to choose either the keyword strategy or the diagram strategy to summarise the texts presented subsequently. The difference between the two conditions shows the degree to which students are impeded by a lack of availability of diagrammatic strategies.

Coding

Two independent raters determined whether a correct spatial representation, a correct keyword representation, or neither of the two had been used by each participant for each of the six problems. If a representation had a spatial structure that matched the structure of the content of the text, and if most of the items (i.e., all or all but one) were placed correctly into the structure, the notes were coded as diagram use. We chose this strict coding rule which excludes incorrect use of diagrams, because in some conditions of Experiments 1 and 2 the students were asked to use diagrams. Students not being able to use diagrams for summarising a text sometimes responded to this instruction by drawing random representations, which resemble actual diagrams but do not spatially organise the content of a text in a useful way. The current coding system is useful, because it excludes such random behaviour as being coded as diagram use. Written keywords or a written copy of the text was coded as keyword use. The very rare diverging ratings were unified in a discussion.

Results

Figure 6.1 reveals an illustration of the frequency of correct diagrammatic representations for the six texts. As expected, students generated very few diagrams for the episodic text 2 and many more diagrams for the description of a spatial arrangement in text 6. So the manipulation check yielded positive results. A comparison between the different tasks shows that tree diagrams (text 3), simple linear arrays (texts 4 and 5), and maps (text 6) are used by about 60 per cent of students under the keyword-or-diagram condition. In contrast, cluster representations (text 1) were rarely used.

Frequency of diagram use was determined for each person by computing the percentage of texts that were summarised by means of a correct diagram. A score of 0 would mean that a student generated no diagram at all. A score of 100 would indicate that each of the six texts was summarised in a diagram. The means and standard deviations for the two grade-levels and the two experimental conditions are displayed in Table 6.1. An ANOVA with these two factors revealed a significant main effect for grade-level, $F(1,127) = 9.4$, $p = .003$, $\eta^2 = .069$, as well as for condition, $F(1,127) = 35.3$, $p < .001$, $\eta^2 = .127$. There was no interaction effect, $p = .687$, $\eta^2 = .001$.

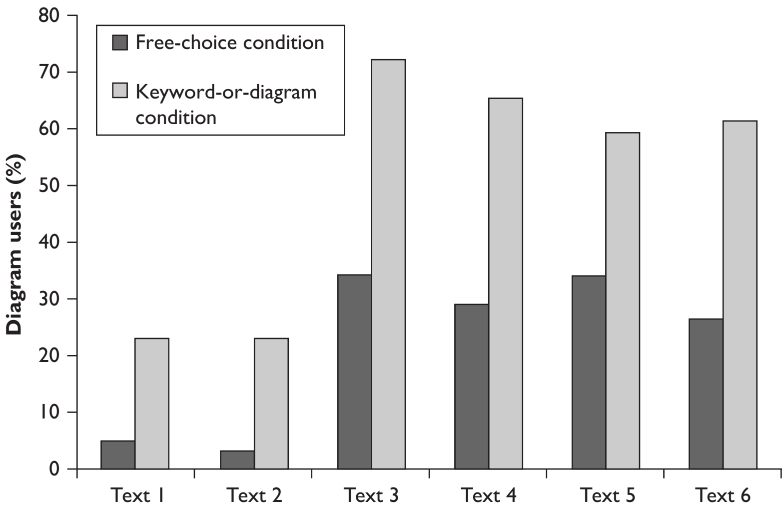


Figure 6.1 Percentage of students who correctly used a diagram to represent one of the six texts presented in Experiment 1.

Table 6.1 Frequency of correct diagram use by each person in Experiment 1 reported in percentage of problems.

	Free-choice condition		*Keywords-or-diagram condition*		*Total*	
	M	*SD*	*M*	*SD*	*M*	*SD*
Grade 7	15	23	42	30	28	30
Grade 9	28	27	58	27	45	30
Total	22	25	50	28	37	30

The more frequent diagram use in the keyword-or-diagram condition suggests that there is a potential for the use of spatial diagrams as representational tools but that this knowledge is not used spontaneously unless it is activated by appropriate instructions.

The interpretation of our results is based on the assumption that using diagrams is as effective as, or even more effective than, alternative strategies for summarising the presented texts. In order to test these hypothesised positive effects of diagrammatic strategy use, we computed for each student the mean solution rate for the two memory questions and the mean solution rate for the inference question.

For the entire sample, the partial correlations between the diagrammatic score on the one hand and the memory scores and inference score on the other hand were computed while controlling for grade-level. The number of diagrams used and the solution rate for the memory tasks correlated with $r = .25$, $p = .004$. The number of diagrams used and the solution rate for the inference task correlated with $r = .57$, $p < .001$. The rather high correlation between the diagrammatic score and the inference score is compatible with the claim that visual-spatial representations are more helpful than keywords for drawing inferences.

One could object, however, that the correlations are caused by general cognitive competencies. More competent students may be better at answering inference questions and they may be more inclined to produce diagrams, even if the diagrams are not used as reasoning tools. This objection, however, can be toned down if the instruction, which led to an increased use of diagrams, can be shown to also have a positive effect on the solution rates of the memory tasks and reasoning task presented at the end of the experiment. Means and standard deviations of these scores are depicted in Table 6.2. As expected, the experimental groups differ highly significantly, both in their solution rate for the memory tasks, $F(1,129) = 14.3$, $p < .001$, $\eta^2 = .100$, and in their solution rate for the inference task, $F(1,129) = 12.7$, $p = .001$, $\eta^2 = .090$.

These results show that diagrams are more helpful than notes for answering memory and inference questions about the texts used in this study. We did not record the individual students' time on task, so we cannot say whether

Table 6.2 Mean solution rates in percentage for the memory questions and the inference question in the two treatment groups in Experiment 1.

	Free-choice condition		*Keywords-or-diagram condition*		*Total*	
	M	*SD*	*M*	*SD*	*M*	*SD*
Memory questions	86	15	94	9	90	13
Inference question	72	18	83	19	78	19

the mechanism by which diagram use facilitated learning was that diagram use led to longer learning times or, alternatively, that it changed the nature of students' reasoning about a text. However, overall, students not using diagrams to summarise our texts did not choose the most effective strategy.

Discussion

The participants of Experiment 1 rarely used diagrams spontaneously. However, the intervention, which pointed out the usefulness of keywords and diagrams, significantly increased the use of diagrams and improved students' memorising and inferences. These results indicate that secondary school students do not have diagrammatic representational strategies available. When the strategies are made available to them by means of a short instructional intervention, students are also able to use them. The intervention had explained the use of keywords as well as of diagrams for summarising texts. Thus, students did not necessarily have to use diagrams since the diagrammatic strategy was only one of the two options presented. The fact that students did use this option also shows that the diagrammatic strategy can be activated and carried out correctly by students once they have acquired it.

However, we do not know from Experiment 1 whether students who did not use diagrams spontaneously lacked the ability to construct diagrams or whether they did not activate diagrammatic strategies because they lacked the necessary understanding of the potential of diagrammatic representations that would enable them to abandon the familiar keyword strategy. From research on strategy change (Siegler, 2007) we know that substituting a familiar strategy for a new one is a prolonged process, even if the new strategy is far more efficient than the old one. Using the keyword strategy was less efficient than using the diagram strategy, but since the keyword strategy still enabled students to answer the questions, they may have chosen it based on familiarity. In school, students typically are not provided with explicit arguments and evidence for the benefits of using diagrams. Therefore, a reluctance to activate the new diagrammatic strategy may persist, even when the students do know the strategy. To find out the extent to which there is a lack of activation of the diagram strategy, we investigated spatial strategy use in a second experiment under the condition that no alternative strategy was allowed. We also used some new texts in Experiment 2 in order to investigate students' competencies and difficulties with depicting two dimensions.

Experiment 2: the activation of diagrammatic strategies

In order to find out whether students' performance in using diagrammatic representations could be boosted further by instruction on how to use this representational strategy, participants were tested under two conditions: the keyword-or-diagram condition or the diagram-only condition. In the keyword-or-diagram condition (the same as in Experiment 1), students were free to choose the representational form they preferred. In the diagram-only condition, students had to use diagrams to represent the problem. Four of the problems from Experiment 1 were used. In addition, four new, complex texts were presented which described two-dimensional relations. Appropriate spatial representations for these new texts were path diagrams and matrices. We hypothesised that participants would use diagrams more frequently in the diagram-only condition than in the keyword-and-diagram condition. This would indicate that our participants have the diagrammatic strategy available (i.e., they know it) but only activate it when explicitly asked to do so.

Method

Participants

A total of 81 volunteers were tested. The 43 seventh-graders (22 girls) and 38 ninth-graders (19 girls) were recruited from the same track of the German educational system as in Experiment 1. All seventh-graders were 13 years old. The ninth-graders had a mean age of 15.2 years ($SD = 0.6$). The students were compensated with 15 DM (approximately US$15).

Materials

Participants in Experiment 2 were presented with eight texts which included four of those used in Experiment 1 (texts 3–6 in the Appendix). The four new problems in Experiment 2 (texts 7–10 in the Appendix) required the use of common diagrams allowing the representation of more than one dimension, such as path diagrams or matrices. Two of the new problems were based on the Cartesian product, that is, they required each element of one set to be combined with each element of another set. One of these problems (text 10, adapted from Novick & Hmelo, 1994) dealt with possible combinations of sweaters and trousers, while the other one, text 8, was about the roads connecting three towns on an island. In two problems only some of the elements of two sets had to be connected: text 9 (adapted from Schwartz, 1993) was about fictitious animals in the jungle that eat, or are eaten by, other animals, while text 7 described a group of children with various hobbies, some common and some different.

Procedure and design

Students in Experiment 2 received basically the same instructions as in Experiment 1. At the beginning of the experiment, the keyword strategy and the diagram strategy of note taking were explained to all students. A major question to be addressed in the second study was whether one can boost students' diagrammatic performance further. Therefore, all students participating in this study solved the task under two conditions: for four of the problems they were free to choose either the keyword or the diagram method; for the remaining four problems they were specifically instructed to use the diagram method. The order of the two conditions was counterbalanced, and across the two conditions, the problems were also counterbalanced.

As in Experiment 1, participants worked through the booklets individually. After completion of the representation task and a short break, they were allowed to consult their notes while they answered the questions about the problems. Diagram use and keyword use in the participants' notes were categorised as in Experiment 1.

Results and discussion

Figure 6.2 shows the proportion of each experimental group who correctly used diagrams for summarising each text. The frequency of correct diagrams varied between texts. Students had more difficulties when more demanding – specifically two-dimensional – problem representations were required. For instance, text 6, which required a map-like representation, and text 4, which required a linear array, yielded more correct diagrams than text 7 or text 10, which both required a two-dimensional representation. Overall, the students demonstrated a good understanding of diagrams under the diagram-only condition.

We computed, for each person, the number of texts summarised by means of a diagram in the keywords-or-diagram condition and the diagram-only condition, respectively. These scores were expressed as percentages of the number of texts presented under the respective condition. For example, when a person generated diagrams for three of the four texts presented in the diagram-only condition the person's score for this condition would be 75 per cent.

Table 6.3 shows the mean numbers of texts the seventh-graders and ninth-graders represented in diagrams under both instructional conditions. A repeated-measure analysis of variance with the within-subjects factor instructional condition and the between-subjects factor grade-level revealed a significant multivariate main effect for the instructional condition, $F(1,79) = 33.1$,

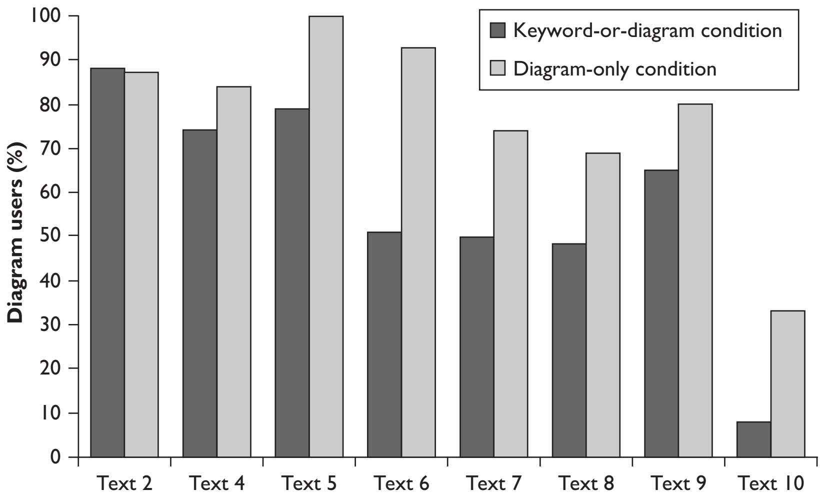


Figure 6.2 Percentage of students who correctly used a diagram to represent one of the eight texts presented in Experiment 2.

Table 6.3 Frequency of correct diagram use by each person in Experiment 2 reported in percentage of problems.

	Keywords-or-diagram condition		*Diagram condition*		*Total*	
	M	*SD*	*M*	*SD*	*M*	*SD*
Grade 7	55	25	78	25	65	19
Grade 9	63	30	80	20	71	20
Total	58	28	78	23	68	20

$p < .001$, partial $\eta^2 = .282$, but no effect of grade-level, $F(1,79) = 1.7$, $p = .195$, partial $\eta^2 = .021$, and no interaction of the two factors, $F(1,79) = 0.4$, $p = .546$, partial $\eta^2 = .005$.

Students produced more correct diagrams when they were instructed to produce a diagrammatic representation than when they were free to choose between diagrams and keywords. Thus, this experiment indicates that even when students have broad knowledge of a variety of representational forms, they often do not use it spontaneously, but only when they are instructed to do so.

As in Experiment 1, we tested whether diagram use led to better answers on the memory tasks or the elaboration tasks. Table 6.4 shows the mean solution rates for the two conditions. The experimental manipulation had no effect on the solution rates for the memory tasks or the elaboration tasks (all *p*s > .4,

Table 6.4 Mean solution rates in percentage for the memory questions and the inference question in the two treatment groups in Experiment 2.

	Keywords-or-diagram condition		*Diagram condition*		*Total*	
	M	*SD*	*M*	*SD*	*M*	*SD*
Memory questions	95	9	96	8	96	6
Inference question	73	23	76	23	74	14

all partial $\eta^2 < .009$). This was found in two repeated-measures ANOVAs, each of which had the within-subjects factor experimental condition and the between-subjects factor grade-level. However, the total number of diagrams used in the experiment correlated significantly with the memory score ($r = .23$, $p = .037$) and with the inference score ($r = .42$, $p < .001$) after we controlled both variables for grade-level.

General discussion

Our results indicate that using diagrams pays off. The more diagrams students used, the better their recall of information or their inference of new information. In addition, an experimentally induced increase in the frequency of diagram use also led to increases in the recall and inference of information in Experiment 1. Despite these advantages of diagrams, they were rarely used spontaneously. In the free-choice condition in Experiment 1, where students received no instructions about how to summarise the texts, diagrams were only generated in 22 per cent of all possible cases.

This conforms to our impression, expressed in the introduction, that school instruction does not help students to use the potentials that diagrams have to their fullest extent. The spontaneous use of diagrams increased highly significantly with the grade-level. This shows that students learn something about diagram use in middle school. However, our participants came from the highest educational track in Germany. Even in this relatively high-achieving group, spontaneous diagram use was still as low as 28 per cent in Grade Nine, indicating that school instruction needs to be improved.

The strong effects of our short and simple interventions show that such improvements would not be difficult to achieve. In the diagram-only condition of Experiment 2, students were able to use diagrams correctly on an average of 78 per cent of all problems. So middle-school students are well able to transform abstract relations described in a text into a visuospatial representation. Moreover, as demonstrated in both experiments, students do not use diagrams in a mindless way, but adapt the diagrams they choose to the nature of the relations to be visualised. We used a total of ten different texts that differed strongly in the nature of the described relations. Students recognised this and flexibly created the appropriate types of diagrams.

When diagram use has positive effects on memory and reasoning, and when students are able to use diagrams appropriately, then why do they usually fail to do so? The results of our two experiments demonstrate that both students' availability and students' activation of diagrammatic strategies are deficient. In Experiment 1, students were free to choose between the two strategies in both experimental conditions. However, diagram use as a strategy for summarising texts was explained and, thus, made more available only to one group of students. This intervention, too, had a highly significant effect on diagram use. The findings from Experiment 2 show that a lack of activation of diagrammatic strategies is part of the problem. In both interventions, the diagrammatic strategy was explained and, thus, made available to all students. However, students preferred to use keywords instead of diagrams, when they had the choice.

As a consequence, school instruction needs to focus on the improvement of both factors. The use of diagrams should not only be discussed in the context of function graphs in mathematics instruction but in other subjects and content areas as well. Students need to practise the use of diagrams, so that they have this strategy available for when it might be useful. In addition, school instruction should highlight the specific advantages and disadvantages of texts, pictures, formulae and various types of diagrams. This knowledge can help students to choose adaptively among alternative external knowledge representations for solving given problems, instead of just resorting to a familiar default strategy (Kramarski & Ritkof, 2002). An important topic for further research are the exact cognitive mechanisms by which diagram use leads to learning gains, because these mechanisms were not in the focus of our study.

In a discussion of the potential of diagrams, Larkin and Simon (1987) came to the conclusion that they 'are (sometimes) worth ten thousand words'. The word 'sometimes' is important, because diagrams are no panacea. Each type of external knowledge representation has its specific advantages and limitations (De Bock et al., 2003; Friel, Curcio, & Bright, 2001). The better students understand this, the more adaptively they can employ the different representational forms to their fullest extent.

Acknowledgements

We thank Theresa Treasure for her help with the manuscript preparation. The study was funded by the Max Planck Institute for Human Development, Berlin, Germany.

Appendix: texts and test questions presented in the two experiments

The materials were presented to the participants in German. Here, we show an English translation.

Problem 1

Text

There are many different kinds of frogs, some of which differ greatly in their eating and nesting habits. Bullfrogs, wood frogs and stone frogs have quite similar eating and nesting habits. Green frogs and leopard frogs also are very similar in their eating and nesting habits. Leopard frogs and wood frogs, however, differ greatly as to their eating and nesting habits. The eating and nesting habits of brown frogs are similar to those of red frogs, but completely different from those of the other frogs.

Most appropriate diagram: clusters

Questions

(1). Do bullfrogs and stone frogs have similar eating and nesting habits? (2). Do brown frogs and red frogs have similar eating and nesting habits? (3). Do green frogs and bullfrogs have similar eating and nesting habits?

Problem 2

Text

Susie and Frank live in Barrington. They are 14 years old, and the school they go to is Linton College. They enjoy reading, and Frank also likes to play basketball. Susie's mother is a photographer. Frank's father is an optician. In summer, they both go to Scotland.

Most appropriate diagram: no diagram

Questions

(1). How old are Susie and Frank? (2). What is Susie's mother doing for a living? (3). Where do they go to in the summer?

Problem 3

Text

On the faraway planet of Urx, living beings are called pings. There are two kinds of pings: spotted pings and striped pings. There are also two kinds of spotted pings: laughing pings and crying pings. Among the striped pings, there are the noisy ones and the quiet ones. Tip is a crying ping.

Most appropriate diagram: hierarchical tree

Questions

(1). Are laughing pings spotted or striped? (2). Are the quiet pings spotted or striped? (3). Is Tip spotted or striped?

Problem 4

Text

Julian likes noodles best, salad not at all, and French fries a little bit. He likes apples more than salad, but French fries more than apples. He likes rice second best. He likes potatoes less than French fries, but more than apples. He likes Cornflakes and crispies less than apples but more than salad.

Most appropriate diagram: linear ordering

Questions

(1). Does Julian prefer potatoes or French fries? (2). What does he like second best? (3). Does he prefer potatoes or crispies?

Problem 5

Text

Mary wants to go for a swim today. Before she does, however, she buys a book. After her swim she goes first to the hairdresser's, then shopping. After that, she has lunch. But before having lunch, she calls her girlfriend, whom she will meet after lunch. Then she goes home.

Most appropriate diagram: linear ordering

Questions

(1). What does Mary do after lunch? (2). What does she do first after her swim? (3). What does she do between going to the hairdresser's and calling her friend?

Problem 6

Text

Louisville is a small town. Facing the church, there is a flower shop. On the right side of the flower shop, there is a supermarket. Opposite the supermarket, beside the church, there is a school. On the right side of the school, there is a drugstore. Beside the drugstore, there is a hairdresser. On the right side of the supermarket, there is a playground.

Most appropriate diagram: map

Questions

(1). What is on the right side of the drugstore? (2). What is on the right side of the flower shop? (3). What is opposite the playground?

Problem 7

Text

The teacher of a fourth-grade class asks her students what hobbies they have. All the children in her class have different names, but some children have several hobbies. Anne says that she likes swimming; Monica likes to collect shells, and Susan and Hannah like to read. Susan also likes to collect shells, and Fanny says that she, too, likes to collect shells. Alicia says that she likes to read; and Gerald likes swimming. Alicia and Susan also like swimming.

Most appropriate diagram: two-dimensional representation

Questions

(1). Which activity does Hanna like? (2). Does Fanny like swimming? (3). Which child has the largest number of activities?

Problem 8

Text

On the island of Mobumbi, there are only three towns. One town is called Adi, one is called Bedi and one is called Cedi. All the roads from Adi to Cedi run through Bedi. There are only four roads from Adi to Bedi and only three roads from Bedi to Cedi.

Most appropriate diagram: Cartesian product in a two-dimensional representation

Questions

(1). How many roads are there from Adi to Bedi? (2). How many roads are there from Bedi to Cedi? (3). How many roads are there from Adi to Cedi?

Problem 9

Text

In nature, there are very complicated food chains. In the jungle of Muzumbi, for instance, Dasings eat Tindals and Sandis. Tindals eat Pondos. Godas eat Dasings. Faltings eat Rondas. Pondos are eaten by Sandis and by Rondas.

Most appropriate diagram: two-dimensional representation

Questions

(1). What do Dasings eat? (2). What do Faltings eat? (3). Which animal is eaten by the largest number of other animals?

Problem 10

Text

Susie always gets a birthday parcel from her grandmother. This year, she asked her grandmother for some clothes. She supposes that her grandmother will buy her blue or green or red or yellow trousers. In addition, she will probably get a sweater. The sweater will also be blue or green or red or yellow. Susie hopes that she will get either a red sweater and green trousers or yellow trousers and a blue sweater or a yellow sweater and yellow trousers, for those are her favourite combinations.

Most appropriate diagram: Cartesian product in a two-dimensional representation

Questions

(1). What does Susie get from her grandmother? (2). What does Susie think goes well with green trousers? (3). How many possible combinations of trousers and sweaters are there?

References

Bauer, M. I., & Johnson-Laird, P. N. (1993). How diagrams can improve reasoning. *Psychological Science, 4*, 372–378.

De Bock, D., Verschaffel, L., Janssens, D., van Dooren, W., & Claes, K. (2003). Do realistic contexts and graphic representations always have beneficial impact on students' performance? Negative evidence from a study on modelling non-linear geometry problems. *Learning and Instruction, 13*, 441–463.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*, 124–158.

Greeno, J. G., & Hall, R. P. (1997). Practicing representation: Learning with and about representational forms. *Phi Delta Kappan, 78*, 361–367.

Hardy, I., Schneider, M., Jonen, A., Stern, E., & Möller, K. (2005). Fostering diagrammatic reasoning in science education. *Swiss Journal of Psychology, 64*, 207–217.

Kramarski, B., & Ritkof, R. (2002). The effects of metacognition and email interactions on learning graphing. *Journal of Computer Assisted Learning, 18*, 33–43.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65–100.

Lewis, A. B. (1989). Training students to represent arithmetic word problems. *Journal of Educational Psychology, 81*, 521–531.

Mevarech, Z. R., & Kramarski, B. (1997). From verbal descriptions to graphic representations: Stability and change in students' alternative conceptions. *Educational Studies in Mathematics, 32*, 229–263.

Novick, L. R. (2001). Spatial diagrams: Key instruments in the toolbox for thought. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 40, pp. 1242–1256). San Diego, CA: Academic Press.

Novick, L. R., & Hmelo, C. E. (1994). Transferring symbolic representations across nonisomorphic problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1296–1321.

Novick, L. R., Hurley, S. M., & Francis, M. (1999). Evidence for abstract, schematic knowledge of three spatial diagram representations. *Memory & Cognition, 27*, 288–308.

Schwartz, D. L. (1993). The construction and analogical transfer of symbolic visualizations. *Journal of Research in Science Teaching, 30*, 1309–1325.

Siegler, R. S. (2007). Cognitive variability. *Developmental Science, 10*, 104–109.

Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General, 127*, 377–397.

Stern, E., Aprea, C., & Ebner, H. G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction, 13*, 191–203.

Stern, E., & Staub, F. (2000). Mathematik lernen und verstehen: Anforderungen an die Gestaltung des Mathematikunterrichts. [Learning and understanding mathematics: requirements for the organisation of mathematics education]. In E. Inckermann, J. Kahlert, & A. Speck-Hamdan (Eds.), *Sich Lernen leisten: Grundschule vor den Herausforderungen der Wissenschaft* [Affording oneself to learn: primary school for the challenges of science] (pp. 90–100). Köln: Luchterhand Verlag.

Tversky, B., Kugelmass S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology, 23*, 515–557.

Part II

Development of representational tools and evaluation of their effects on student learning

Chapter 7

Conceptual change in learning electricity

Using virtual and concrete external representations simultaneously

Tomi Jaakkola, Sami Nurmi, and Erno Lehtinen

Introduction

Previous research has established that teaching students to understand the functioning of electric circuits on a qualitative level is a difficult pedagogical challenge (e.g., Lee & Law, 2001; McDermott & Shaffer, 1992; Reiner et al., 2000). First, the central concepts, such as voltage, current and resistance, are very abstract by nature and refer to processes that are dynamic and often intangible in natural situations. Consequently, it is not easy to provide the students with accurate information about electric circuits in an easily comprehensible format. Second, students have many misconceptions about electric circuits which seem to be exceptionally tenacious and resistant to teaching efforts (e.g., Lee & Law, 2001; McDermott & Shaffer, 1992). The problem is that the students are not well aware of the limitations of their initial model, which is often coherent enough for them to feel that they have arrived at a consistent (albeit incorrect) and satisfactory explanation (Vosniadou, 2002).

According to some of the recent studies, the use of computer simulations seems to promote students' conceptual understanding of electric circuits more effectively than the use of real circuits (e.g., Finkelstein et al., 2005; Jaakkola & Nurmi, 2008; Zacharia, 2007). Finkelstein et al. (2005), for instance, examined the effects of substituting a computer simulation for real circuits on learning the basics of DC circuits in a university physics course. They found that the students using the simulation outperformed the students using the real circuits, both on a conceptual knowledge test and in the coordinated tasks of assembling a real circuit and describing how it works.

One explanation why computer simulations seem to promote conceptual understanding more effectively than real circuits, seems to be that the development of a theoretical understanding of electric circuits through practical manipulation with real circuits can be problematic; in many cases students can only see what is happening on the surface level of the circuit, while being unable to grasp the underlying processes and mechanisms that are important for theoretical understanding (e.g., current flow) (Finkelstein et al., 2005; Hennessy, Deany, & Ruthven, 2006). In comparison with real circuits, a simulation can make the functioning of electric circuits more transparent; it can model circuits on various levels of abstraction (e.g., a circuit in schematic format vs. the mimicking of real bulbs and wires) and visualise processes that are invisible in natural systems[1] (Finkelstein et al., 2005; Goldstone & Son, 2005; Hennessy et al., 2006). This visualisation allows the students to become better aware of the limitations of their initial reasoning (the output of the simulation may be in conflict with their expectations) and discover the properties of the scientific model embedded in the simulation (e.g., the electric circuit is a closed system in which all components interact; Ohm's law; total resistance in parallel and series circuits) (e.g., de Jong, 2006; Lehtinen & Rui, 1996). Another distinctive feature of computer simulations is that the embedded model(s) often highlights the elements that are important for theoretical understanding (e.g., interdependence between current, voltage, and resistance) and excludes (or hides) the elements that are irrelevant or potentially misleading (e.g., poor connections, worn batteries, tangled wires, colour of wires, or even broken wires or bulbs) (e.g., Finkelstein et al., 2005; Goldstone & Son, 2005).

Our own findings suggest that, at least in the elementary school context, computer simulations and laboratory activities should be considered as complementary (rather than competing) instructional tools that, in combination, can provide appropriate conditions for conceptual change and deeper understanding of electrical circuits (Jaakkola & Nurmi, 2008; Jaakkola, Nurmi, & Veermans, in press; see also Ronen & Eliahu, 2000). In our first study (Jaakkola & Nurmi, 2008) fourth- and fifth-grade students solved circuit assignments in three different learning conditions – a computer simulation (using only a simulation), a hands-on laboratory exercise (using only real circuits) and a simulation–laboratory combination (using the simulation and the real circuits in parallel). In each condition the students had 90 minutes to practice with the circuits during the intervention. The results showed that the development of conceptual knowledge was the most notable in the combination condition. Students in the simulation condition also made clear progress during the intervention, but their conceptual understanding of electric circuits did not reach the desired level in the post-test. The progress was the most modest in the laboratory condition where the students' conceptual understanding remained at an elementary level, even after the intervention. In a more recent study (Jaakkola et al., in press) we investigated the role of implicit and explicit instruction (see the section on method, page 137) on students' conceptual learning outcomes when they used the simulation either on its own (simulation environment) or in parallel with real circuits (simulation–laboratory hybrid environment) to learn the basic principles behind the functioning of electric circuits. The main finding was that although the explicit instruction was able to improve students' conceptual understanding of electrical circuits considerably in the simulation environment, their understanding did not reach the level of the students who used the simulation and the real circuits in parallel, even after the amount of time spent on constructing

Table 7.1 Gain in subject knowledge (pre-test post-test change) in standard deviation units, proportion of correct models in the post-test and learning efficiency in each learning condition.

	Simulation implicit	*Simulation explicit*	*Hybrid implicit*	*Hybrid explicit*
Gain in subject knowledge (d[a])	.22	.78	1.51	1.24
Proportion of correct models (%)	17	43	58	67
Learning Efficiency[b]	0.572 (1.840)	1.700 (2.101)	3.554 (2.39)	2.232 (1.597)

[a]d is Cohen's (1988, p. 48) standardised mean difference effect size for one sample paired observations.
[b]Learning efficiency (gain in subject knowledge per learning time; cf. Rasch & Schnotz, 2009) controls the effect of time on learning outcomes. In order to cover the same content in all conditions, the students were allowed to spend an unlimited amount of time to construct and study the circuits during the intervention.

and studying the circuits during the intervention was controlled (see Table 7.1 for main results; more details are provided in Jaakkola et al., in press).

The aim of this chapter is to investigate from video data the issues that could explain why combining and linking the use of virtual external representations (a computer simulation) with concrete external representations (laboratory activities) seems to promote students' conceptual understanding so effectively. The analysis focuses on the three theoretical issues outlined below.

(1) In our study, the students use the simulation and the real circuits side by side, in parallel.[2] That is, they construct each circuit first with the simulation and then reconstruct the same circuit using the real equipment immediately afterwards.[3] This means that the students always have two different representations of electrical circuits available. Although the simulation and the real circuits are superficially dissimilar, the underlying principles are the same in both. Several studies (e.g., Gentner, Loewenstein, & Thompson, 2003; Gick & Holyoak, 1983; Kurtz, Miao, & Gentner, 2001; see also Ainsworth, 2006) have found that analogical encoding – comparing two instances of a to-be-learned principle – is a powerful means of promoting learning, even for novices. Gentner and her colleagues (2003), for instance, examined schema abstraction and transfer among novices learning negotiation strategies. All participants studied two short case examples dealing with negotiation strategies. Half of the participants were encouraged to compare the cases, and half were encouraged to study them one at a time. The outcome was that drawing comparisons led to greater understanding of the schema and greater transfer than did reading the cases separately. The authors explain that drawing comparisons between two complementary cases can help students focus on the common principles shared by the cases, and thus result in a more abstract understanding of the phenomena. This finding

suggests that analogical encoding can be effective even quite early in learning, when learners lack knowledge of an appropriate base domain. Accordingly, we assume that the simultaneous use of virtual and concrete representations can improve students' understanding of the principles behind the functioning of electrical circuits.

(2) In line with Piagetian theory many authors have proposed that cognitive conflict between prior knowledge and the requirements of new tasks is a fundamental driving force in learning scientific concepts (e.g., Chinn & Brewer, 1993; Strike & Posner, 1982). However, recent research has shown that a cognitive conflict introduced by pedagogical arrangements is often insufficient to promote conceptual change and conceptual learning in general (Limón, 2001). Merenluoto and Lehtinen (2004) have shown that there are inter-individual and situational differences in students' sensitivity to unfamiliar aspects of new learning tasks, and in their cognitive, metacognitive and motivational strategies to cope with experienced cognitive conflict. We assume that the simultaneous use of virtual and concrete representations can increase students' sensitivity to novel findings, and the richer external support afforded by the hybrid environment can result in the use of more adequate strategies to deal with the conflict.

(3) Self-explanations may be a factor that helps students' conceptual understanding. In their study on problem solving in physics, Chi and VanLehn (1991) found that good solvers provided more self-explanations during the problem-solving process. They defined self-explanations as comments that pertained to the content of physics. Self-explanations are generated in the context of learning something new. We assume that the opportunity to move between two external representations in the hybrid environment triggers students' self-explanations. On the other hand, the incidence of self-explanation can also be increased by explicit guidance to look for explanations.

Based on the above theoretical analysis we present two general research questions:

- Are there differences between the students participating in the four experimental conditions in terms of experienced cognitive conflicts and presented self-explanations?
- Is there evidence in the hybrid environment to support the use of reasoning and learning by analogical encoding?

Method

Some parts of this section are provided in condensed format. More details about the method can be found in Jaakkola et al. (in press).

Design

The students constructed various circuits in four different learning conditions between pre-test and post-test:

- In the simulation implicit condition (SI) the students used a simulation to construct electric circuits and they received implicit instruction. Implicit instruction means that the students were provided only with procedural guidance, i.e. they were told what kind of circuit to construct, how to construct it, and what kind of electrical measurements to conduct, but they were not told on what aspects of the circuits they should focus.
- In the simulation explicit condition (SE) the students used the simulation and received explicit instruction. Explicit instruction means that the students constructed exactly the same circuits as the students receiving implicit instruction, but they were given more support and structure for their inquiry process, i.e. when they constructed the circuits and conducted electrical measurements they were guided to focus on the circuit elements that are important for a theoretical understanding (e.g., current flow, changes in voltage across the bulbs in various circuits) and asked to explain their findings (e.g., under which condition there is a current flow).
- In the hybrid implicit condition (HI) the students used the simulation and the real circuits in parallel and received implicit instruction. Parallel use means that the students constructed exactly the same circuits as the students in the simulation conditions (SI, SE), but they constructed each circuit twice in a row: first using the simulation and then, immediately after succeeding with the simulation, reconstructing that (same) circuit with the real equipment (circuits) that was placed next to the computer (see Figure 7.1).[4] This means that they had continuously two different representations of electrical circuits available.
- In the hybrid explicit condition (HE) the students used the simulation and the real circuits in parallel and received explicit instruction.

Participants

The participants were 50 fifth- and sixth-grade students (11–12 years old; 31 girls and 19 boys) from three different classrooms of one urban Finnish elementary school. They had no previous formal education in electricity. Student allocation into the four conditions was based on matching; sets of four students were matched on pre-test scores, and from each set one student was allocated randomly to one of the four learning conditions.[5] This was to ensure that the conditions would have the nearest to equal spread of subject knowledge at the baseline. After the students were matched into the conditions, pairs were formed randomly within each condition (each pair worked in the same condition). Working in pairs is a natural procedure in science classrooms in Finland

Figure 7.1 Screenshot example of students working in the hybrid environment.

and previous studies have shown that working in pairs can be especially effective when the work involves computers or requires complex problem-solving processes. The students were taught by the same teacher in all conditions, in order to control for a possible teacher effect.

Materials

Simulation

The simulation used was the 'Electricity Exploration Tool' (EET; 2003; Figure 7.2). The representation level of the EET is semi-realistic; it displays circuits schematically, but includes light bulbs with dynamically changing brightness (as the amount of current through the bulb increases, the yellow area inside (and around) the bulb becomes larger and the colour tone of that yellow changes as well) and realistic measuring devices. The simulated model is authentic with some exceptions: unlike real circuits the wires have no resistance, the battery is always ideal (i.e. there is no change in the potential difference with time), connections are always proper and measurements always ideal. With the EET, students are able to construct various virtual DC circuits by using the mouse to drag wires, bulbs and resistors to the desired location in the circuits. After constructing the circuit or putting the circuit into a particular configuration, students can observe the effects of their actions and get instant feedback. They can, for instance, see how the current flows within the circuit, and whether and how brightly the bulbs are lit. They can also conduct electrical

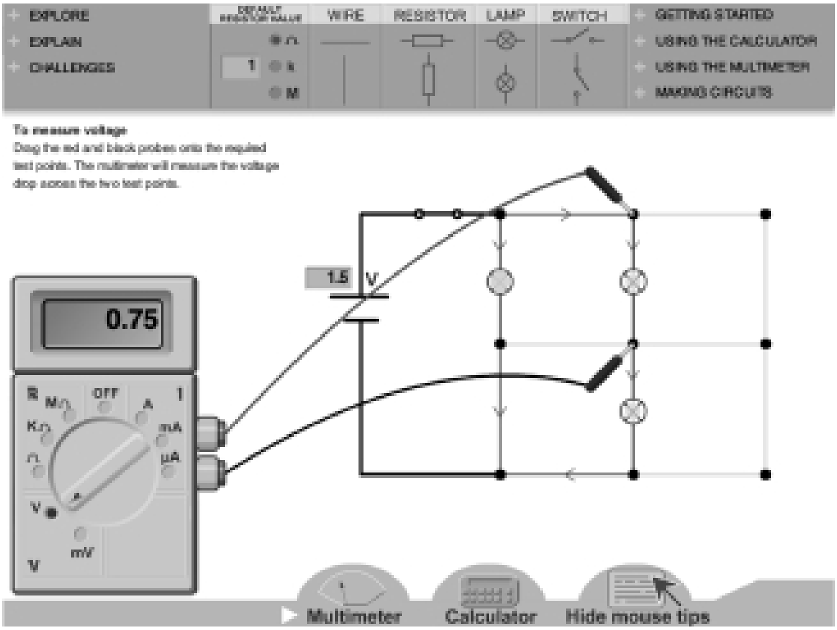


Figure 7.2 The Electricity Exploration Tool (© Digital Brain).

measurements with a multimeter by dragging its probes to the required testing points. In the present study the students asked to measure the bulb voltages under the assumption that observing the voltages across the bulbs in different configurations could help them to understand the variations in bulb brightness better.

Real circuits

The laboratory equipment kit (LEK) consists of real batteries, wires, bulbs and a voltmeter. It allows the students to construct various real DC circuits and conduct electrical measurements. In the LEK, each circuit component is attached to a base that displays the diagrammatic symbol of that component (see Figure 7.3). Inclusion of the diagrammatic symbols is believed to make it easier for the students to relate the real circuits and the virtual circuits and make translations from one representation to the other. The LEK was used only in the hybrid condition.

Worksheets

In all conditions the assignments and instructions were given in the form of worksheets that asked and guided the students to construct various circuits and conduct various electrical measurements with the simulation (EET) and the

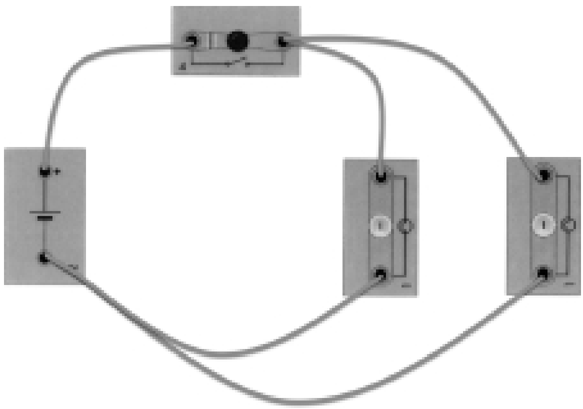


Figure 7.3 Example of a parallel circuit constructed with the laboratory equipment kit.

real circuits (LEK; hybrid condition only). The worksheets were designed to promote conceptual change; the students were asked to construct circuits that would (A) confront common misconceptions of electric circuits that have been identified in previous studies and (B) introduce an alternative (correct) explanation that is based on the scientific model of electric circuits. In total there were 12 worksheets and they became gradually more difficult; in the beginning the circuits that the students were asked to construct contained only a single bulb but later multiple bulbs with mixed configuration. Each worksheet focused on one topic.

Procedure

Pre-test and post-test

The students completed a subject knowledge assessment questionnaire individually, before and after the intervention. The questionnaire measured students' knowledge about the features that affect the lighting and the brightness of the bulb(s) in simple DC circuits.

Intervention

The actual intervention, when the pairs constructed various circuits in the four different learning conditions, lasted one session (completion of 12 worksheets without time limit) and took place in the school's computer suite. Since this study was the students' first formal introduction to the subject of electricity, the main aim was for the students to establish an understanding of the relationships between the observable variables, i.e. the number of bulbs, the circuit configuration, and the variations in bulb brightness, as well as the relationship between the bulb brightness and the voltage across the bulb. During the intervention the pairs were working one at the time in the school's computer suite because each pair was videotaped. They could take as long as they wanted to solve all the circuit assignments of the 12 worksheets by either using only the simulation (SI, SE) or the simulation in parallel with real circuits (HI, HE). The pairs received one worksheet at a time and were required to write down their answers on the worksheet. They could proceed to the next worksheet only when they had completed the previous worksheet correctly.

Data collection and analysis

The work of all 25 pairs was video-recorded during the intervention. One video-recording device captured the action on a computer screen and the other recorded students' actions, expressions and talk. These two video streams were combined into one video output layer in order to synchronise students' reactions with the related situation (see Figure 7.1). A detailed transcript of each videotape was constructed. This included students' conversational interactions, their answers to the worksheets, and their non-verbal interactions with the simulation and the LEK.

A method of content analysis was used to analyse the video data and video data transcripts. The focus of the analysis was on cognitive conflicts, self-explanations and analogical encodings. Two independent raters rated 20 per cent of the video data concerning cognitive conflicts and self-explanations.

Cognitive conflict. An incident was categorised as a cognitive conflict if a student explicitly expressed disbelief in the results of the simulation or real circuits and searched for an explanation for the discrepancy between their expectations and the results. A situation where a student was first surprised by the results of the simulation or real circuits, but accepted the new result immediately, or paid no additional attention to the matter, was *not* categorised as cognitive conflict. Inter-rater reliability (Cohen's Kappa) for cognitive conflicts was .88.

Self-explanation. A comment or a comment chain that contained domain-relevant articulation concerning the behaviour of a particular virtual or real circuit was categorised as a self-explanation. Inter-rater reliability for self-explanations was .88.

Analogical encoding. The analysis related to analogical encodings focused on the use of the simulation and the real circuits in parallel in the hybrid environment. Analogical encoding was defined as an event where the students linked and made explicit translations between the simulation and the real circuits. Because we did not attempt to compare any conditions, instances of analogical encodings were not categorised nor quantified. Instead, in the results section, we will provide illustrative examples of the situations where analogical encodings took place.

It should be noted that there were some problems relating to the video data. Sometimes the students were only whispering (although they were encouraged to speak at a normal conversation level), so it was impossible to make sense of the conversation. Although in most of the cases both the video and the audio were flawless, there were also some minor technical problems in the data: the audio or the video or both were sometimes momentarily disrupted (typically for one to three seconds). The worse case was in the hybrid implicit condition where almost half of one video and audio was corrupted (this was the only incidence with such a large data loss).

Results

We begin the results section by investigating the number of cognitive conflicts and self-explanations in each condition. This is followed by a section that focuses on instances of reasoning and learning by analogical encoding in the hybrid environment (HI, HE) by presenting excerpts of the translated video data transcripts.

Excerpts of the transcripts of the video data follow the following conventions:

- [ ] Words enclosed in square parentheses have been added to aid understanding and readability of the dialogue.
- ( ) Words enclosed in parentheses indicate non-verbal actions, e.g., measuring of the voltage across a bulb.

The names are pseudonyms.

Cognitive conflicts and self-explanations in different learning conditions

As we can see from Table 7.2, only about a quarter of the students in each condition experienced a cognitive conflict during the intervention. Only one student experienced more than one conflict during the intervention. All the conflicts were experienced with the simulation. Although explicit instruction (SE, HE) seemed to promote more self-explanations than implicit instruction

Table 7.2 Number of cognitive conflicts and average number of self-explanations (SD) in each learning condition.

	Simulation implicit (SI) N = 12	*Simulation explicit (SE)* N = 14	*Hybrid implicit (HI)* N = 12	*Hybrid explicit (HE)* N = 12
Number of cognitive conflicts (number of students experiencing a conflict)	4 (3)	3 (3)	3 (3)	3 (3)
Average number of self-explanations (SD)	1.33 (1.37)	2.50 (2.21)	1.92 (1.72)	2.75 (2.63)

(SI, HI), there was no statistical difference in the number of self-explanations generated during the intervention between the students who received explicit instruction and those receiving implicit instruction, Kruskall Wallis' ANOVA, $X^2(1) = 1.86$, $p = .17$. There were also no differences between the students who used only the simulation (SI, SE) and those who used the simulation and the real circuits in parallel (HI, HE), $X^2(1) = .286$, $p = .59$. Also, no differences were detected between the four conditions, $X^2(3) = 2.39$, $p = .495$.

Some of the students who experienced a cognitive conflict seemed to have misconceptions that were already quite deeply rooted. In excerpt 1 below, Paula has a constant current model. In this misconception current (in this case voltage) is always shared equally among the circuit components regardless of how the circuit has been configured; bulb brightness is believed to be negatively correlated with the number of elements.

Excerpt 1

Paula and Annukka work with the simulation. They have just constructed a series circuit with two bulbs. All the previous worksheets have included only single-bulb circuits. Now their task is to measure the voltages across the bulbs. Paula has earlier told Annukka that when her family was living overseas she constructed circuits with real equipment in school.

SEQUENCE 1

Paula (SCAFFOLDING): The voltage of those [two bulbs] is not 1.5 volts because there are two bulbs. That is because the magnitude of the battery, or whatever that is, is only 1.5 volts and that 1.5 needs to be divided so that there is enough for both.

The girls measure the voltage across each bulb and they obtain 0.75 volts for both.

SEQUENCE 2

Now the girls have moved to the next worksheet where they have been asked to construct a two-bulb parallel circuit.

PAULA: The bulbs are equally bright.

PAULA (MEASURES THE VOLTAGE ACROSS THE FIRST BULB): The voltage of the first bulb is one point [1.5] . . . no it's not.

PAULA: How can it be 1.5 because there is that other bulb? [If the voltage across the first bulb is 1.5 volts] then the other would get nothing.

PAULA: I don't understand this.

PAULA: The voltage can't be 1.5, because then the other bulb wouldn't get any voltage.

Paula stares for some time at the computer screen.

PAULA (GROANS): Let's put that as 1.5 (writes it down as an answer on the worksheet).

PAULA (CONTINUES WONDERING): I just don't get how this can be possible!

PAULA (POINTS TO THE VOLTMETER): If it is connected like that, maybe it measures the whole thing [the voltage across both bulbs].

PAULA: [If the voltage across each bulb is 1.5 volts] Shouldn't it be 3 volts that battery?

PAULA (TO TEACHER): This result is not possible!

TEACHER: If you take a look at the circuit configuration, it differs from the previous circuit [two bulbs in series]. Your results are correct. This circuit is just different.

In order to be sure, Paula takes the mouse and reconstructs the previous circuit (two bulbs in series).

Analogical encoding in the hybrid environment

In this section our aim is to illustrate from the video data transcripts from the hybrid environment (HI, HE) how the students link the use of two representations in parallel and how this might contribute to their understanding of electrical circuits.

Excerpt 2

Here the students have built a single-bulb circuit. The task is to measure bulb voltage and compare the voltage across the bulb to the battery voltage.

Joni uses the virtual multimeter to measure the bulb voltage in the simulation. The meter shows 1.5 volts.

Kalle watches from close by.

JONI: The bulb and battery voltages are identical.

Now the students shift to the real circuits.

JONI (AFTER HE HAS MEASURED THE REAL BULB VOLTAGE): 1 . . . about 1.5.

TEACHER (LOOKS AT THE VOLTMETER READING): Yes, about 1.3 volts [to be precise].

TEACHER: Do you have any idea why the voltage is slightly lower in the real circuit than in the simulation?

JONI (POINTING TOWARDS THE REAL BULB AND EXPLAINING): This wears down the battery.

TEACHER: Exactly. In the simulation we have a kind of ideal case; it's as if we always have a brand new battery.

Joni (going back to the simulation): I'd like to see what would happen if this [virtual] battery was not full. Joni tries to adjust the virtual battery to 1.3 volts to match the real battery. He manages to type in 1 volt instead of 1.3 volts exactly (in the simulation it is somewhat difficult to adjust the battery voltage because the virtual battery updates very quickly so you have to type in very fast). Nevertheless, the boys observe the change in bulb voltage and brightness when Joni changes the battery voltage.

Excerpt 3

Here Atso and Stiina have built a series circuit with two bulbs. They have measured the voltages across each bulb alone (0.75 v) and together (1.5 v) in the simulation. Now they have rebuilt the same circuit with the real equipment.

ATSO: [The overall brightness of the two bulbs connected in series gives] pretty dim light.
TEACHER: That's right. Now you can measure the voltages across the bulbs.
ATSO (AFTER MEASURING THE FIRST BULB): 0.6[v].
ATSO: Do we also have to measure the second bulb?
TEACHER: You can measure the voltage across both bulbs.
ATSO (AFTER MEASURING): So that is equal to the total [battery voltage] (Atso had measured earlier that the real battery voltage was 1.2 volts).
TEACHER: Yes. Good.

In excerpt 2, Joni first learned from the simulation that in a single-bulb circuit the potential difference across the bulb equals the battery voltage. His theory was further strengthened when he switched to the real circuits and realised that the same rule applied there, even though the voltages were slightly different due to the friction and resistance. In the final step, when he moved back to the simulation and adjusted the battery voltage, he saw once more that the voltage across the bulb was identical to the battery voltage. In a similar fashion, in excerpt 3, Atso had first learned from the simulation that when two bulbs are connected in series the voltage across each bulb is halved and the voltage across both is equal to the battery voltage. When he then moved to the real circuits, he learned that, despite slightly lower voltages, the same rules applied.

Excerpt 4

SEQUENCE 1

Here Meri and Sanna have built a series circuit with three bulbs using the simulation and they have measured the voltages across each bulb.

MERI: All the bulbs are equally bright.
SANNA: No they aren't.

MERI: They all have the same voltage.
SANNA: They don't.
MERI: Yes they do. Look. (Shows her answers in the worksheet) 0.5, 0.5, 0.5.
SANNA: Oh yes. You're right.

Now the students move to the real circuits. They are puzzled because one of the bulbs does not seem to light up or the light seems to be extremely dim (almost invisible). In the simulation above they had just concluded that all the bulbs were equally bright.

SANNA: This doesn't light up.
MERI (BENDS FORWARD TO EXAMINE THE CIRCUIT CLOSER): Yes it does.
SANNA: Does it?
MERI: Just a tiny [light].
TEACHER: It is very dim – it should be dim.
TEACHER: Try another battery.
MERI (AFTER SHE HAS CHANGED THE BATTERY): now it's even worse [dimmer].
TEACHER: Try to measure the bulb voltages.
MERI (MEASURES THE VOLTAGE ACROSS THE FIRST BULB): Just a little.
MERI (MEASURES THE SECOND BULB): It is the same … it is 0.5 [volts].
MERI (MUMBLES HAPPILY AFTER SHE HAS MEASURED THE THIRD BULB): Mmm-m …

SEQUENCE 2

Here, just a few moments ago, Meri and Sanna had created a rather complex circuit using the simulation where the bulb brightness can be expressed as A > B > C = D. They obtained the following results when they measured the voltages across each bulb: A = 0.9 v, B = 0.6 v, C&D = 0.3 v. Now the girls have recreated the same circuit using the real equipment. While constructing, they constantly used the simulation as a point of reference. Because the voltage across each bulb is low, the bulbs are again very dim. While they inspect the real circuit, they use the simulation as a point of reference.

SANNA: These two [A & B] light up.
MERI: These two [C & D] may be lit as well. You might not be able to see 0.3.
MERI: Let's test it [with the voltage meter].
MERI (MEASURES ONE BULB): This would be 0.6 [bulb B].

The girls measure the next bulb.

MERI: Yes, 0.3 [bulb C].
MERI (THIRD BULB): Yeah [bulb D; the same as bulb C].
MERI: Then the last one … Yeah [0.9v, bulb A], pretty good [that we were able to complete this assignment].

At the beginning of sequence 1 the students learned from the simulation that when three bulbs are connected in series the voltage across each is identical, and Sanna in particular understood the relationship between voltage and bulb brightness when she stated that because the bulbs share the same amount of voltage they must be equally bright. When the students moved to the real circuits, their understanding of the relationship was further elaborated; they understood the relationship between the bulb voltage and the relative brightness of the bulb (or the whole circuit). This issue is nicely demonstrated in sequence 2 where two of the bulbs did not light up but Meri stated explicitly that you don't necessarily see light if the voltage across the bulb is only 0.3 v. In both sequences the simulation played an important role as it was used as point of reference when the students constructed real circuits and interpreted their functions.

Discussion

The aim of this chapter is to explore from the video data the issues that could explain why combining and linking the use of virtual external representations (a computer simulation) with concrete external representations (laboratory activities) seems to promote students' conceptual understanding of electrical circuits so effectively. The focus was on the following three issues: cognitive conflicts, self-explanations and analogical encoding.

During the intervention we found no difference in the amount of cognitive conflict between the four conditions. In all conditions only about a quarter of the students experienced a conflict and only one student experienced more than one conflict. The fact that all the conflicts were experienced with the simulation is probably due to the fact that in the hybrid environment (HI, HE) the students were asked to construct each circuit first with the simulation. The other explanation is the general lack of conflict and the fact that students' initial models concerning electrical circuits were mostly immature and fragmentary; the overall resistance to change was therefore relatively low because this was the students' first formal introduction to electric circuits. The fragmented nature (cf. diSessa, 1993) of the students' initial models in the present study becomes evident when we look at the reliability of the subject knowledge assessment questionnaire: Cronbach's alpha for the pre- and post-test was .667 and .822, respectively (Jaakkola et al., in press). The lower pre-test alpha level means that the students' knowledge of electricity was less accurate and systematic before the intervention than after the intervention. In other words, at this early stage of science learning, the students seem to have some correct prior knowledge about the functioning of electrical circuits, but that knowledge is incomplete. Consequently, learning could be regarded more as gap filling or enriching than as conceptual change (Chi, 2008). If we consider the proportion of correct models in Table 7.1, this finding suggests that it is indeed beneficial to try to promote students' conceptual understanding of electric circuits as early as the elementary school level. At this early stage of science learning the students do not have deeply rooted misconceptions because their ideas about the functioning of electrical circuits are not yet coherent and consistent. In line with many previous findings (e.g., McDermott & Shaffer 1992; Reiner et al., 2000), excerpt 1 suggests that as the students acquire more experiences with the electrical circuits, and their ideas become more coherent, their resistance to new ideas increases accordingly.

We also found no differences in the number of self-explanations during the intervention between the four conditions or the two factors (environment and instruction). However, there was a clear tendency suggesting that explicit guidance increased the amount of self-explanations. The fact that we did not find statistically reliable differences between any of the conditions might partially be a function of small sample size and skewed distributions; among the total sample of 50 students there were 14 students who did not provide a single self-explanation and 11 who provided only one self-explanation. At the other extreme, one student provided eight self-explanations. The students' young age and inexperience with electrical circuits may have contributed to the fact that they generally provided so few self-explanations during the intervention; at this early age it is not easy to articulate ideas about the functioning of electrical circuits explicitly. However, it is likely that the numbers of reported self-explanations are underestimations; sometimes the pairs were just whispering, and due to a few technical problems some pieces of data were missing (see data analysis section). The worst situation was in the hybrid implicit condition (HI) where two of the pairs were mostly whispering throughout the intervention. If we look at Table 7.2, we can see that among the students who received implicit instruction (SI, HI), there seems to be a nascent trend suggesting that those students who used the simulation and laboratory equipment in parallel (HI) seemed to generate more self-explanations than the students who used only the simulation (SI). Without the above kind of problems, this trend could be more pronounced.

The video data excerpts provided clear evidence about the existence and the benefits of analogical encodings in the hybrid environment. In all three excerpts that were presented, the students clearly benefited from the fact that they could compare simultaneously virtual and concrete representations of electrical circuits. In excerpts 2 and 3 the students had to deal with discrepant results of voltmeter readings between the virtual circuits and the real circuits. In order to understand the alteration in voltmeter readings caused by friction and resistance, they needed to focus on those features that could be generalised across the two representations. This meant that the students had to first discover the rules governing each representation based on the data (i.e., voltmeter readings) and then infer a further abstraction from these rules that would apply to both representations. In excerpt 4 the simulation played an important role as it was used as a point of reference when the students constructed real circuits and interpreted their functioning; the simulation helped the students to deal with the discrepant results and understand the functioning of real circuits better. It also appears that the real circuits introduced more details and deepened students' understanding. In sequence 1, for instance, the students first concluded from the simulation that because the voltages across the bulbs were identical, then the bulbs' brightness must be identical as well. However, they paid no explicit attention to the overall brightness of the bulbs or the circuit. When they moved to the real circuits they paid attention spontaneously to the overall brightness of the bulbs and learned the relationship between the bulb voltage and the relative brightness of the bulb (or the whole circuit).

This finding suggests that the use of two representations can also cause slightly different aspects of the content to appear more salient than others (cf. Ainsworth, 2006). Had the students used only the simulation, their understanding of the relationship between the voltage and the bulb brightness might have remained at a more surface level (the bulbs with identical voltages are equally bright but no attention paid to the overall brightness). Had the students used only the real circuits, this might have led to a current consumption misconception;[6] in sequence 1, for instance, the students might just have concluded that one bulb would be dimmer than the other two because there is less current left for it. Now they knew from the simulation that all the bulbs should be equally bright and the voltage across each should be identical (they had learned from earlier tasks that laws they learn in the simulation also apply to the real circuits). The fact that analogical encoding did not increase notably the amount of self-explanations in the hybrid environment as compared with the simulation environment deserves some thought. As it was discussed above, it could just be that the students' young age and inexperience are factors impeding explicit self-explanations. Support for this interpretation comes from Gentner et al. (2003) who point out that 'Working through the comparison of two cases that share a common underlying principle can be illuminating even if the common principle is only partially understood in either case' (p. 394). This implies that learning in an analogical encoding situation is often implicit and does need to take the form of explicit comparison, as was the case in the excerpts that were provided in the results section.

To conclude, the results of this study show that it can be beneficial to try to promote students' conceptual understanding of electrical circuits at the early elementary school level because they do not yet have deeply rooted misconceptions that could hamper teaching and learning. In line with results from other domains (e.g., Gentner et al., 2003; Gick & Holyoak, 1983; Kurtz et al., 2001), the results further suggest that it is beneficial to use the simulation and the real circuits in parallel, because analogical encoding of two information resources can improve schema abstraction and deepen students' conceptual understanding of electrical circuits.

Acknowledgement

The present study was supported by Grant 116393 from the Academy of Finland.

Notes

1. As an example of a naturally hidden process, the existence of current cannot be observed in real circuits (you can, of course, measure it), but an electricity simulation can easily show whether or not there is a flow inside a circuit, the path of that flow and possibly even its magnitude.
2. The way the simulation and laboratory activities are combined in our studies is different from earlier attempts. In other studies (e.g., Zacharia, 2007), the students have been typically using each material at different times.
3. The decision to ask the students to construct each circuit first with the simulation is based on the assumption that constructing virtual circuits is easier than constructing real circuits (cf. Finkelstein et al., 2005), and the virtual circuit could then serve as a point of reference when the students reconstruct the circuit with the real equipment.
4. It would make little sense to ask the students to construct every circuit twice in the simulation condition. It is highly unlikely that such instruction (constructing the same circuit twice in a row with the same equipment) would be taken seriously by any student. In order to match the amount of circuits constructed in the hybrid conditions the students using the simulation would need to construct twice as many different (additional) circuits. This would result in unequal coverage of the content between the simulation and hybrid conditions. Furthermore, if the circuits become more and more complex, as in our study, it is questionable whether the students could construct twice as many circuits.
5. Since fifty is not divisible by four, prior to matching, two randomly selected students were allocated randomly to one of the four learning conditions.
6. According to this view, current flows from the battery and, while it travels through the circuit, it encounters obstacles (resistors) that gradually consume the current and slow it down.

References

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183–198.

Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *Handbook of research on conceptual change* (pp. 61–82). Hillsdale, NJ: Erlbaum.

Chi, M. T. H., & VanLehn, K. A. (1991). The content of physics self-explanations. *Journal of the Learning Sciences, 1*, 69–105.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Erlbaum.

de Jong, T. (2006). Computer simulations: Technological advances in inquiry learning. *Science, 312*, 532–533.

diSessa, A. (1993). Towards an epistemology of physics. *Cognition and Instruction, 10*, 105–225.

Electricity Exploration Tool [Adobe® Flash application]. (2003). Brighton, UK: digitalbrain plc.

Finkelstein, N. D., Adams, W. K., Keller, C. J., Kohl, P. B., Perkins, K. K., Podolefsky, N. S., Reid, S., & LeMaster, R. (2005). When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment. *Physical Review Special Topics – Physics Education Research, 1*, 1–8.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*, 393–408.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1–38.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences, 14*, 69–110.

Hennessy, S., Deaney, R., & Ruthven, K. (2006). Situated expertise in integrating use of multimedia simulation into secondary science teaching. *International Journal of Science Education, 28*, 701–732.

Jaakkola, T., & Nurmi, S. (2008). Fostering elementary school students' understanding of simple electricity by combining simulation and laboratory activities. *Journal of Computer Assisted Learning, 24*, 271–283.

Jaakkola, T., Nurmi, S., & Veermans, K. (in press). A comparison of students' conceptual understanding of electric circuits in simulation only and simulation-laboratory contexts. *Journal of Research in Science Teaching*. DOI 10.1002/tea.20386.

Kurtz, K. J., Miao, C., & Gentner, D. (2001). Learning by analogical bootstrapping. *Journal of the Learning Sciences, 10*, 417–446.

Lee, Y., & Law, N. (2001). Explorations in promoting conceptual change in electrical concepts via ontological category shift. *International Journal of Science Education, 23*, 111–149.

Lehtinen, E., & Rui, E. (1996). Computer supported complex learning: An environment for learning experimental methods and statistical inference. *Machine Mediated Learning, 5*(3–4), 149–175.

Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction, 11*, 357–380.

McDermott, L. C., & Shaffer, P. S. (1992). Research as a guide for curriculum development: An example from introductory electricity: Part I. Investigation of student understanding. *American Journal of Physics, 60*, 994–1013.

Merenluoto, K., & Lehtinen, E. (2004). Number concept and conceptual change: Towards a systemic model of the processes of change. *Learning and Instruction, 14*, 519–534.

Rasch, T., & Schnotz, W. (2009). Interactive and non-interactive pictures in multimedia learning environments: Effects on learning outcomes and learning efficiency. *Learning and Instruction, 19*, 411–422.

Reiner, M., Slotta, J. D., Chi, M. T. H., & Resnick, L. B. (2000). Naive physics reasoning: A commitment to substance-based conceptions. *Cognition and Instruction, 18*, 1–34.

Ronen, M., & Eliahu, M. (2000). Simulation – A bridge between theory and reality: The case of electric circuits. *Journal of Computer Assisted Learning, 16*, 14–26.
Strike, K. A., & Posner, G. J. (1982). Conceptual change and science teaching. *European Journal of Science Education, 4*, 231–240.
Vosniadou, S. (2002). On the nature of naive physics. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 61–76). Dordrecht, The Netherlands: Kluwer.
Zacharia, Z. C. (2007). Comparing and combining real and virtual experimentation: An effort to enhance students' conceptual understanding of electric circuits. *Journal of Computer Assisted Learning, 23*, 120–132.

Chapter 8

Using static and dynamic visualisations to support the comprehension of complex dynamic phenomena in the natural sciences

Peter Gerjets, Birgit Imhof, Tim Kühl, Vanessa Pfeiffer, Katharina Scheiter, and Sven Gemballa

Introduction

Comprehension difficulties in the natural sciences often seem to result from the complexity, speed or scale of dynamic phenomena under consideration and from the necessity to relate the observation of the concrete objects involved in these dynamic phenomena to the underlying abstract scientific concepts and theories. To address this type of comprehension difficulty, educators regularly use external visual representations like graphs, pictures, videos or animations when teaching scientific concepts and phenomena. From a psychological perspective, this approach can be justified by referring to the finding that instructional visualisations enable a direct and parsimonious access to visuospatial information (Larkin & Simon, 1987) and facilitate inferences grounded in perception (Goldstone & Son, 2005). It can thus be postulated that instructional visualisations are particularly well suited to conveying an understanding of complex visuospatial relations that are an important characteristic of many scientific domains. Another line of reasoning pertains to the claim that the use of instructional visualisations is particularly helpful when the entities under consideration are difficult or even impossible to observe in the real world. Many scientific phenomena involve this type of entities, which are not directly accessible to perception because they are too small, too big, too fast, too slow, or too complex (Park & Gittelman, 1992). In this paper, however, we will not cover the general issue of whether and when visualisations are useful to support (science) learning. Instead, we will address a more specific question, namely, under what conditions various types of visualisations turn out to be beneficial (e.g., Tversky, Bauer-Morrison, & Bétrancourt, 2002). In particular, we will focus on the comparison between static and dynamic visualisations with regard to their potential to support the comprehension of complex dynamic phenomena in the natural sciences. We will first introduce some theoretical considerations and empirical findings with regard to the drawbacks and advantages of static and dynamic visualisations. In the remainder of the chapter, we will provide an overview of four studies that we conducted in the biological domain of fish locomotion to compare static and dynamic visualisations with regard to varying learning objectives.

Drawbacks and advantages of static and dynamic visualisations

From a theoretical point of view, dynamic visualisations seem to have a strong potential to support learners in understanding dynamic changes, mainly because they can reduce specific processing demands that might otherwise (e.g., when only static visualisations are available) be imposed on learners (e.g., inferring dynamic properties of the depicted phenomenon). However, there are also drawbacks in learning with dynamic visualisations under certain conditions, namely when the dynamic visualisations themselves impose additional processing demands on learners (e.g., memorising configurations of objects that disappear in the visualisation over time). The processing demands that are relevant for analysing the potential drawbacks and benefits of dynamic visualisations can be classified according to their relation to transience, visual complexity, mental animation and the direct depiction of dynamic features. We assume that our theoretical reasoning with regard to these different types of processing demands will not only apply to instructional visualisations in the natural sciences, but also pertain to dynamic phenomena in other domains.

Transience

A potential drawback of using dynamic as compared with static visualisations for instructional purposes is their transience. Due to the short-lived nature of dynamic representations, it may be a perceptually and cognitively demanding process for learners to identify and extract relevant information and crucial states from the visualisation before this information disappears or changes (e.g., Lowe, 1999). Moreover, learners may be challenged by the necessity to memorise crucial states and relate them to each other (e.g., Van Gog et al., 2009). The potential detrimental effect of transience on learning, however, may vary depending on specific characteristics of the learning objective and the instructional visualisation used. For instance, if the learning objective addressed does not involve the identification of crucial states of a process, or if the dynamic visualisation used is presented repeatedly so that relevant information is available several times, processing demands due to transience will be less severe. The four studies summarised in this chapter differ with regard to the learning objectives they address, but they share the feature of presenting dynamic visualisations repeatedly in order to counteract the negative effects of transience. When comparing static and dynamic visualisations with regard to their transience-related processing demands, it should be noted that multiple static pictures that are presented sequentially one after the other will also be of a short-lived nature and, thus, suffer from transience. However, identifying and extracting relevant information and crucial states from each single picture of sequentially presented static visualisations should be perceptually and cognitively less demanding compared with dynamic visualisations.

Visual complexity

Dynamic visualisations often turn out to be more complex than static visualisations because usually several elements at different spatial locations change at the same time. Accordingly, learners can experience difficulties in following the visualisation properly as they have to split their attention between several elements that change simultaneously. With regard to both transience and visual complexity, dynamic visualisations seem to easily violate the so-called apprehension principle, which advocates that the structure and content of visualisations should be designed in a way that the relevant information can be readily and accurately perceived and understood (Tversky et al., 2002). As a result, dynamic visualisations might 'overwhelm' learners (Ayres & Paas, 2007; Lowe, 2003, 2004; Van Gog et al., 2009), eventually leading to an incomplete processing. In general, an appropriate processing of dynamic visualisations presupposes that learners can cope with the transience and visual complexity of dynamic information. In particular, competent users need the ability to identify, memorise, and compare crucial elements and states in dynamic visualisations and to split their attention across several relevant aspects at the same time (Hegarty, 2004; Lowe, 1999; Rieber, 1990; Van Gog et al., 2009). Thus, successful knowledge acquisition from dynamic visualisations can be considered as a resource-intensive process, which requires the simultaneous and optimised availability of different learning resources (e.g., competencies of learners, processing capacities of the cognitive system, appropriate designs of visualisations; Mayer & Anderson, 1992; Mayer & Moreno, 2002; Scheiter, Gerjets, & Catrambone, 2006). As we argued for transience, the detrimental effects of visual complexity will be moderated by learning objectives (e.g., do the learning objectives imply to understand the relation between changes at different spatial locations) and the design of the visualisation (e.g., are learners supported in their attention distribution over time, for instance by means of cueing). In the four studies summarised in this chapter, we kept in mind not to increase unnecessarily the visual complexity of the visualisations used.

Mental animation

Dynamic visualisations deliver information on how objects and their positions change over time (motion; Rieber, 1990) as well as on the directions of these changes (trajectory; Rieber, 1990). Because dynamic – in contrast to static – visualisations explicitly display the continuity of changes, learners do not have to infer these changes on their own by means of resource-demanding processes of mental animation (Hegarty, 1992), which harbour the risk of building less elaborated or even incorrect mental models of the trajectory, especially for learners with lower learning prerequisites. Accordingly, dynamic visualisations can be considered to act as an external substitute for these internal cognitive processes (cf. supplantation, Salomon, 1979), thereby allowing for cognitive offloading (Scaife & Rogers, 1996). This is what Schnotz and Rasch (2005) call the enabling or facilitating function of dynamic visualisations, respectively. Thus, if the learning objectives require an understanding of the continuous aspects of a movement (which is clearly the case in the four studies summarised in this chapter) dynamic visualisations should be superior to static ones.

Direct depiction of dynamic features

An important advantage of dynamic visualisations is that they are able to directly represent specific dynamic features of moving objects such as their velocity or acceleration. Therefore, they should be particularly apt to conveying knowledge with regard to these dynamic features. In contrast, static visualisations do not allow the property to depict the velocity and acceleration of objects. Moreover, these dynamic features cannot even be inferred directly from static visualisations – so that in this respect static visualisations are not informationally equivalent to dynamic visualisations. Rather, in learning with static visualisations, this kind of information has to be provided by means of additional external representations (e.g., text). The required integration of static visualisations with verbal information on dynamic features into a coherent mental model is probably a demanding and resource-intensive process. Thus, whenever the understanding of specific dynamic features is required, as, for instance, in Study 1 of this chapter, we expect dynamic visualisations to be clearly superior to static visualisations. This assumption can be justified by means of the so-called congruence principle that recommends ensuring a correspondence between the desired structure and content of a learner's internal representation and the structure and content of an external instructional representation (Tversky et al., 2002).

Taken together, our theoretical considerations suggest the expectation that dynamic visualisations should outperform static ones with regard to learning effectiveness when learning objectives are addressed that require a deeper understanding of continuous aspects of a movement (which might be hard to animate mentally) or that involve specific dynamic features like velocity or acceleration (that cannot be represented directly without dynamic visualisations). Furthermore, dynamic visualisations should be more effective when problems of transience and visual complexity are kept tractable, for instance, by means of an improved instructional design.

Empirical findings

Up to now, the empirical question of whether dynamic visualisations are better suited than static ones to conveying knowledge still remains unanswered. Different meta-analyses have come up with different results. In a review by Tversky et al. (2002), most of the studies failed to show any advantages of dynamic compared with static visualisations. Moreover, in a study by Mayer et al. (2005), the authors even found static visualisations to be superior to dynamic visualisations. A recent meta-analysis by Höffler and Leutner (2007), on the other hand, revealed a medium-sized overall advantage of dynamic compared with static visualisations. Although the results of this meta-analysis are in favour of using dynamic visualisations, the large heterogeneity among the empirical findings also highlights the importance of looking more closely at the conditions under which dynamic visualisations will aid learning (cf. Bétrancourt, 2005; Hegarty, 2004; Schnotz & Lowe, 2008). Höffler and Leutner (2007) identified several factors that moderate the effectiveness of dynamic visualisations. In particular, the to-be-achieved learning objective appears to play an important role in this respect. The strongest effects in favour of dynamic visualisations were observed for those studies that involved procedural-motor knowledge as a learning objective (see also Ayres et al., 2009; Van Gog et al., 2009; Wong et al., 2009).

In the four studies summarised in this chapter we used varying learning objectives to compare learning from static and dynamic visualisations. All studies were related to the same biological domain of fish locomotion in order to keep the overall domain constant. Within this domain, we addressed learning objectives like the conceptual understanding of the physical principles underlying fish locomotion, the classification of different fish locomotion patterns, or the identification of different fish species based on important static and dynamic features. These learning objectives differ with regard to the task demands they impose (e.g., how difficult and important is it to animate a movement mentally or to extract different multiple features simultaneously), and with regard to the knowledge tests that are appropriate for their assessment (e.g., factual knowledge tests, recognition tests, or transfer tests based on verbal or pictorial stimuli).

Another important consideration with regard to available empirical comparisons between dynamic and static visualisations concerns the fact that different types of static visualisations are often pooled into a single category, whose effects are then compared with dynamic visualisations. However, as will be argued next, it can be expected that the effectiveness of static visualisations will depend on their presentation format (e.g., single versus multiple static visualisations, sequential versus simultaneous presentation of multiple static visualisations). Nevertheless, there are only a few studies available that investigate, for instance, the relative effectiveness of a sequential versus simultaneous presentation of multiple static visualisations (e.g., Boucheix & Schneider, 2009). In a sequential presentation, multiple static pictures are presented one after another, whereby each picture is replaced with its successor in the sequence of events. In a simultaneous presentation, all pictures are presented on one display, representing different discrete states of a movement at the same time. As the third Experiment summarised in this chapter investigates different presentation formats for static visualisations, we will provide some additional reflections on the differences between a sequential versus simultaneous presentation of multiple static visualisations in the following paragraph.

From a theoretical perspective, a main advantage of a *sequential presentation* will be that it supports the identification of corresponding visual elements across multiple pictures, because these elements will appear in (almost) the same spatial position on the screen for each picture. However, as each picture is replaced by its successor in a sequential presentation, this format shares some of the problems of transience with dynamic visualisation. It has to be considered, though, that in contrast to dynamic visualisations, the frame rate of sequential static presentations is usually much lower (for instance, in the studies summarised in this chapter, the frame rate for sequential presentations of static visualisations is about 1 frame every four seconds or even slower, whereas it is approximately 15 frames per second in dynamic visualisations). A main advantage of a *simultaneous presentation*, on the contrary, is that the depicted information remains visible on the screen. Thus, this format might enhance learners' ability to compare discrete steps of a process in greater detail. Additionally, learners presented with multiple static visualisations simultaneously can easily control the pacing of their cognitive processing, whereas in a sequential presentation format the transience of the information will affect the pacing of the processing. In line with this reasoning, Boucheix and Schneider (2009) demonstrated in a mechanical domain that simultaneous static visualisations, allowing for comparisons among discrete states, improved learning compared with sequential static visualisations and were as effective for learning as dynamic visualisations. To test whether simultaneous static visualisations are also better suited than sequential ones to support learners in understanding complex dynamic phenomena in the natural sciences, we compared dynamic visualisations with both types of multiple static visualisations in the third study reviewed in this chapter.

A final issue in comparing dynamic and static visualisations empirically is related to the question of how well these two types of instructional visualisations are suited to supporting learners not only in rather artificial learning experiments in the laboratory, but also in realistic and situated learning scenarios where the use of instructional materials is intertwined with real-life experiences and *in vivo* observations of the to-be-taught phenomenon. According to Stone-Romero (2002), experimental laboratory settings are obviously useful to investigate basic psychological processes, as they inform about causal connections that also occur in field settings. However, it remains often unclear, whether effects found in our laboratory settings will also show up in more realistic and applied settings. To investigate the effectiveness of visualisation formats, not only in the laboratory, but also in a situated learning scenario, the fourth study reviewed in this chapter compares static and dynamic learning materials for identifying fish species in the context of a snorkelling excursion in the Mediterranean Sea.

In sum, based on both theoretical considerations as well as empirical data, it can be concluded that dynamic as well as static visualisations have their specific benefits and drawbacks, depending on the concrete learning objectives addressed, on the task demands imposed by them, and on the knowledge tests used for their assessment. Accordingly, it is plausible that mixing up studies addressing different learning objectives in a single meta-analysis might yield equivocal and unclear overall results. From our perspective, it seems to be more promising to use a task-analytic approach by first analysing for a particular piece of instruction, which of the potential advantages and drawbacks of the different types of visualisation may be crucial, before deciding on an instructional format that seems to be effective. In particular, it has to be considered how important factors like transience, visual complexity, mental animation, depiction of dynamic features, identification of corresponding objects, and comparisons between discrete process steps are for this particular piece of instruction. We will exemplify this approach for different learning objectives in the biological domain of fish locomotion.

Overview of experiments

In the remainder of this chapter we will review four experiments that compared static and dynamic visualisations for teaching different aspects of fish locomotion. The first three studies were conducted in the laboratory, whereas the fourth study was a field study, in which a classroom setting was combined with learning in the field. Fish locomotion is a scientific domain that is characterised by concrete, complex, partially unperceivable, fast and dynamic processes, including changes of hidden aspects (e.g., moving body parts that are visible only from a particular perspective), which may be difficult to observe during real-world observations. Thus, in teaching fish locomotion the use of instructional visualisations can, in general, be expected to be helpful. Knowledge about fish locomotion is particularly important for biologists because, on the one hand, knowledge about different forms of fish locomotion can be used to classify different fish families or species, and, on the other hand, the various movement patterns are related to several important principles in biology (evolutionary adaptation, ecosystems) and principles in other sciences (e.g., physics). Thus, the locomotion behaviour of fish is a complex and dynamic science content that reflects both dynamic and non-dynamic aspects.

In our studies we addressed three different types of learning objectives related to fish locomotion: Study 1 has a conceptual learning objective, namely understanding the physical principles underlying fish locomotion; Studies 2 and 3 have a perceptual learning objective, that is, classifying different fish locomotion patterns according to their characteristic perceptual features, and Study 4 has an applied learning objective, namely identifying different fish species in a situated scenario based on important static and dynamic features. According to the theoretical considerations outlined in this chapter, we expected dynamic visualisations to be superior to static visualisations for those learning-outcome measures that have a strong focus on dynamic aspects of the content domain. This is expected because dynamic aspects may be difficult to animate mentally or to infer from static visualisations (e.g., speed differences in different phases of a locomotion cycle, classifying fish species according to their movement pattern). On the other hand, learning-outcome measures that do not necessitate a dynamic mental model of the phenomenon under consideration may not benefit from dynamic visualisation (e.g., answering factual knowledge on scientific terms, understanding biodiversity).

Study 1: Conceptual understanding of the physics of fish locomotion

In the first study (Kühl Scheiter, Gerjets, & Edelmann, in press), we addressed the effectiveness of static versus dynamic visualisations for supporting a conceptual understanding of the physical principles underlying fish locomotion. Eighty undergraduate students were asked to acquire knowledge on how an undulatory (i.e., wave-like) fish movement generates forces under water. This topic requires the understanding of physical concepts like force vectors in relation to movement characteristics such as trajectories, velocity, and acceleration. The learning materials consisted of seven instructional segments, each lasting for 45 seconds. Each instructional segment in the condition with dynamic visualisations contained an animation showing an undulatory movement of a fish in a recursive fashion, so the movement of the swimming fish was looped. In the static visualisation condition, for each segment, nine key frames from the corresponding animations were extracted. The key frames were displayed sequentially one after another. The nine static key frames represented two repetitions of an undulatory movement, so that each learner had the chance to see each frame twice. Both the animations and the sequential static visualisations were system-paced and accompanied by verbal explanations. The total presentation time was identical for the conditions. Learning outcomes were measured by means of a verbal factual knowledge test, a pictorial test as well as by transfer tasks. For the transfer tasks, learners had to apply their dynamic mental model of fish locomotion to new situations and problems (for instance, a new locomotion pattern was shown and they had to predict in which direction a fish with this pattern would swim and why). Furthermore, processing demands after learning were assessed by means of a rating scale for perceived difficulty ('How difficult was it for you to understand the contents?'). No differences between the visualisation conditions were expected for factual knowledge or for pictorial tasks, as these tasks do not focus on the dynamic aspects of fish locomotion. However, for the transfer tasks we hypothesised dynamic visualisations to be superior to static visualisations. The subjective processing demands measure was expected to be in accordance with learning outcomes, in that higher processing demands are related to lower learning outcomes. In line with our hypotheses, the results of this study showed that there was no difference between static and dynamic visualisation with regard to either factual knowledge ($F < 1$) or pictorial tasks ($F(1,76) = 1.05$, *ns*), whereas for transfer tasks, dynamic visualisations proved to be superior ($F(1,76) = 9.86$, $p < .01$). Moreover, subjects rated processing demands as being a bit higher when learning with static compared with dynamic visualisations ($F(1,76) = 3.41$, $p = .07$). Thus, the results confirm our general assumption concerning the superiority of dynamic visualisations for achieving conceptual learning objectives that require a deeper understanding of a complex dynamic phenomenon that is difficult to animate mentally. An important characteristic of the animations (as well as of the static visualisations) used in this study might be that showing the same cyclic locomotion pattern several times reduced the problem of transience associated with both dynamic and sequential static visualisations.

Study 2: Perceptual classification of fish locomotion

In a second study (Gerjets, Scheiter, & Imhof, 2007), we investigated the effectiveness of dynamic and sequential static visualisations for enhancing students' ability to classify perceptually different fish according to their locomotion pattern. Eighty university students participated in this study. In all experimental conditions, seven different locomotion patterns were illustrated by means of non-interactive visualisations that were accompanied by auditory text. Half of the participants received a dynamic visualisation for each of the seven locomotion patterns, whereas the other half received nine key frames from each of the dynamic visualisations as static visualisation. The series of static visualisations was presented successively by the system for two times each. The total presentation time was system-paced and the same for all conditions (72 seconds for each of the seven locomotion patterns). As a second variable we manipulated the degree of realism of the visualisations between subjects (realistic videos versus computer-generated schematic animations and the respective key frames). However, as there were no differences between corresponding conditions with realistic and schematic visualisations, and as realism is not at issue in this chapter, we will not go into any further detail with regard to this second factor. Thus, subgroups with schematic versus realistic visualisations were collapsed for the statistical analyses referred to in this chapter. Learning outcomes were measured by means of a factual knowledge test, a pictorial recognition test as well as by transfer tasks. The factual knowledge test asked for scientific terms and characteristic features associated with different locomotion patterns. In the recognition test, participants had to recognise the locomotion patterns of fish seen during learning, whereas in the transfer test, learners had to classify the locomotion patterns of novel fish. Both the recognition test and the transfer test require a sophisticated mental representation of a complex dynamic pattern. To assess processing demands concerning the learning phase, learners were asked, immediately after completing the learning phase, to rate the mental activity required during learning ('How much mental activity was required? That is, how demanding was the learning task for you?'), the effort they had to invest to understand the content ('How hard did you have to work in your attempt to understand the contents of the learning environment?') and the stress during learning ('How stressed (insecure, discouraged, annoyed) did you feel during the learning task?'). Each question had to be answered on a rating scale from very low to very high. It was hypothesised that static and dynamic visualisations would not differ for the factual knowledge test, whereas for recognition and transfer tasks learners receiving dynamic visualisations should outperform those with static visualisations. As expected, there were no effects for factual knowledge acquisition ($F < 1$). However, dynamic visualisations improved recognition ($F(1,72) = 10.30$, $p < .01$) and transfer performance ($F(1,72) = 4.92$, $p < .05$) compared with static visualisations. The conditions did not differ with regard to the assessed processing demands, namely mental activity ($F < 1$), effort ($F < 1$) and stress ($F(1,72) = 1.60$, *ns*). Thus, this study demonstrates that dynamic visualisations can outperform static visualisations, not only for conceptual but also for perceptual learning objectives, if the content domain is characterised by complex dynamic patterns and the learning objectives require a sophisticated mental representation of these patterns. Learning objectives in the same content domain, which do not require such a representation (e.g., factual knowledge questions on characteristic features of different locomotion pattern), however, will not benefit from dynamic visualisations.

Study 3: Sequential versus simultaneous presentation of static visualisations

One possible reason for the inferiority of static visualisations in the first two studies might have been that multiple static visualisations were presented sequentially in a system-paced way. Thus, characteristic drawbacks of dynamic visualisations, like their transience, may have applied to the static conditions in these studies as well. Moreover, specific processing advantages of presenting multiple static visualisations, for instance the fact that they may facilitate comparisons among discrete steps, may be better implemented by presenting multiple static visualisations simultaneously. Therefore, we conducted a third study (Imhof, Scheiter, & Gerjets, 2009) with learning materials similar to the ones used in Study 2, with the exception that only four of the seven locomotion patterns were taught. In addition to the two conditions from Study 2, with either dynamic or sequential static visualisations, the effectiveness of simultaneous static visualisations was investigated to test whether dynamic visualisations would not only be superior to sequentially presented, but also to simultaneously presented static visualisations, and whether simultaneous representations would show the expected superiority compared with sequential ones. Similar to

Study 2, the degree of realism was manipulated between participants as a second factor. Half of the participants were presented with realistic visualisations and the other half of the participants received schematic visualisations. As in Study 2, there were no differences between the subgroups with schematic versus realistic visualisations and, thus, the data of these subgroups were collapsed again for the statistical analyses referred to in this chapter. One hundred and twenty university students participated in this study. The presentation mode for the dynamic and sequential static visualisations conditions was the same as in Study 2. In the simultaneous static visualisation condition the nine key frames extracted from the dynamic visualisations were presented at the same time all together on a single screen. The total presentation time of the learning materials was system-paced and the same across conditions (72 seconds per each locomotion pattern). As learning outcome measures, the recognition and transfer tests from Study 2 were adapted. Moreover, processing demands were assessed after the learning phase with the same three items as in Study 2 (mental activity, effort, and stress). Furthermore, eye-tracking data were obtained during learning in order to investigate the cognitive and perceptual processes occurring. Particularly in the static conditions, eye-tracking data can give important information about how learners use the external representations, for instance, how often and how long they look at particular static visualisations. Data analyses revealed overall main effects for recognition ($F(2,108) = 7.18$, $p < .01$) and transfer performance ($F(2,108) = 8.08$, $p < .01$). *Post hoc* tests showed that the dynamic condition outperformed the sequential static condition for recognition ($p < .01$) and for transfer ($p < .001$), thereby replicating the results of Study 2. The simultaneous static condition, however, did not differ significantly from either the dynamic or the sequential static condition. Thus, although simultaneously presented static visualisations were not significantly better than sequentially presented ones, they were as good for learning as the dynamic visualisations. With regard to processing demands, there was an overall effect for effort ($F(2,108) = 3.23$, $p < .05$), a marginal effect for stress ($F(2,108) = 2.85$, $p = .06$), but no overall effect for mental activity ($F(2,108) = 1.79$, *ns*). *Post hoc* tests revealed that sequential static visualisations led to a higher rating of effort ($p < .05$) and a marginally higher rating of stress ($p = .07$) than dynamic visualisations, whereas simultaneous static visualisations did not differ from the other two conditions in this respect. Finally, analyses of the eye-tracking data revealed that there were processing differences for sequential and simultaneous static visualisations. For instance, the number of fixations on the visualisations revealed that learners in the simultaneous static condition looked more often at specific pictures than in the sequential static condition, although the overall number of fixations for all nine pictures did not differ between the two groups. Thus, it seems that learners in the simultaneous static condition used a different strategy to study the pictures, which was particularly effective for learning the locomotion patterns. The results imply that not only dynamic visualisations but also a simultaneous presentation of multiple static visualisations might be an appropriate way of teaching learners how to classify perceptually different locomotion patterns.

Study 4: Combining mobile learning with real-life experiences

In the fourth study (Pfeiffer et al., 2009), we investigated how static and dynamic visualisations interact with real-life experiences. Therefore, we changed the learning scenario in Study 4 into a combination of classroom learning and learning in the field, where the learning objective was to identify different Mediterranean fish species. The standard process of fish species identification is based on using field guides which rely solely on static features (e.g., a combination of morphological features and colour patterns), despite the fact that for fish species identification, not only static features, but also dynamic features (e.g., swimming style, typical behaviours or interactions), may be helpful. We hypothesised that providing additional dynamic representations would be advantageous for this task, because they can provide information on behavioural and locomotory features in an explicit way. Moreover, when observing fish in their natural habitat, while diving or snorkelling, dynamic features are often more salient than static ones and hence reflect the real-world experience more appropriately. In this study, characteristic static and dynamic features of the 18 most common coastal fish species of a specific Mediterranean region were presented to 35 students by means of either digital videos or single static key frames taken from the videos. In the dynamic condition, learners were shown digital realistic videos presenting the species in their natural habitats. Each video pauses either once or several times while playing in order to emphasise the species' most relevant features. During these pauses static frames were presented in combination with written information of two to four words describing the important features visible in the static frames. In the static condition, we only used these still frames from the videos and presented them sequentially to learners. The overall duration of the sequential presentation of the still frames for each species was equivalent to the duration of the corresponding video in the dynamic condition (including the embedded static frames). All visualisations were accompanied by auditory explanations, which were the same in both conditions. The static as well as the dynamic visualisations were presented on portable DVD players. During an initial learning phase (90 min.) in a classroom setting, participants studied the DVDs containing either the static or the dynamic version of the visualisations. Subsequently, students' performance with regard to fish identification was measured by a first post-test, which asked students to identify fish species from unknown videos. Subsequently, a real-world learning experience took place, where students went snorkelling in the Mediterranean Sea (240 min.). They were told to identify as many as possible of the 18 fish species introduced to them in the DVD material. For snorkelling, the dynamic and static groups were assigned to two remote, but nearly equivalent, diving spots. They were instructed to change locations after half of the snorkelling time.

During the snorkelling phase, students were allowed to verify their fish species observations by using the DVD material on the beach and to collaborate with each other in the water and on the beach. Moreover, experts helped students in the water to find suitable places for fish observation, but did not help with fish identification. Immediately after snorkelling, students' learning outcomes were tested again by administering the post-test a second time. The results showed a significant knowledge gain between the first and the second post-test due to the snorkelling experience ($F(1,33) = 213.12$, $p < .001$), with learners achieving higher scores in the second post-test. There was no overall difference between learners who had studied either the static visualisations or the dynamic visualisations (including the embedded static frames, $F < 1$). However, a significant interaction showed that students' knowledge improvement from the first to the second post-test was significantly greater in the dynamic group than in the static group ($F(1,33) = 6.27$, $p < .05$). This interaction could be traced back to a marginal superiority of dynamic over static visualisations for post-test 2 ($p < .10$), whereas no differences could be observed for post-test 1 between the two conditions. Although no main effect was found for visualisations format, the interaction between both factors suggests that providing learners with additional dynamic visualisations, which have been shown to be effective in the first three studies, also seems to possess a higher instructional potential in the context of a real-world learning scenario.

Discussion

The objective of the four studies summarised in this chapter was to test the relative effectiveness of dynamic and static visualisation formats for various kinds of learning objectives in the biological domain of fish locomotion. It could be shown that the dynamism of visualisations did not influence factual knowledge acquisition (Studies 1 and 2). This finding confirms a statement by Bétrancourt and Tversky (2000), claiming that dynamic visualisations are not more advantageous than static visualisations in conveying factual knowledge. However, the results showed that there was a superiority of dynamic visualisation formats on transfer tasks (Studies 1, 2, and 3), for instance in questions on the underlying physical principles (Study 1), and on perceptual tasks, such as recognising locomotion patterns and fish species (Studies 2, 3, and 4). These results confirm our general assumption that dynamic visualisations are an effective instructional device when conceptual or perceptual learning objectives are addressed that require a deeper understanding of dynamic phenomena which are difficult to animate mentally or to infer from static visualisations.

However, in contrast to sequentially presented static visualisations, simultaneous static visualisations were not inferior to dynamic visualisations. This replicates the findings of Boucheix and Schneider (2009), who demonstrated in a mechanical domain that simultaneously presented static visualisations improved performance compared with sequentially presented visualisations and were instructionally equivalent to dynamic visualisations. Thus, on the one hand, a sequential presentation of multiple static visualisations does not seem to use the full potential of static pictures (i.e., being able to compare important steps of a process), on the other hand, this format even shows some of the disadvantages of dynamic visualisations (i.e., information transience).

The results of our fourth study showed that the potential of dynamic learning materials can even unfold when the instructional scenario is enriched with real-world experiences. That is, enhancing a real-world learning experience like snorkelling with dynamic visualisations (and embedded static frames) improved learning to a greater extent than using only static visualisations in this situation. Even if the dynamic visualisations did not lead to an improved knowledge acquisition initially, as reflected in the first post-test before snorkelling, they nevertheless facilitated learning from the subsequent real-world experience as indicated by the superiority of the dynamic condition for the second post-test. A possible explanation for this finding might be that fish observed during snorkelling could be linked to the videos more easily than to the static frames, as behavioural patterns and locomotion behaviour (i.e., aspects not visible in the static visualisations) could be used as cues for memory access. Thus, knowledge acquired from dynamic visualisations may remain rather inert, as long as it is not used and strengthened in the context of a real-world experience (cf. Resnick, 1987). Overall, Study 4 stresses the importance of testing the effectiveness of instructional materials not only in the laboratory, but also in the context of situated learning scenarios, which may act as moderators for their instructional effectiveness.

The results of our studies are in line with the general claim that research should not ask for the most suitable visualisation format in general; rather, it should investigate the conditions under which the specific advantages of various visualisation formats unfold best (e.g., Bétrancourt, 2005; Hegarty, 2004; Schnotz, 2002; Tversky et al., 2002).

References

Ayres, P., Marcus, N., Chan, C., & Quin, N. (2009). Learning hand manipulative tasks: When instructional animations are superior to equivalent static representations. *Computers in Human Behavior, 25*, 348–353.

Ayres, P., & Paas, F. (2007). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology, 21*, 695–700.

Bétrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 287–296). New York: Cambridge University Press.

Bétrancourt, M., & Tversky, B. (2000). Effect of computer animation on users' performance: A review. *Le Travail Humain, 63*, 311–330.

Boucheix, J.-M., & Schneider, E. (2009). Static and animated presentations in learning dynamic mechanical systems. *Learning and Instruction, 19*, 112–127.

Gerjets, P., Scheiter, K., & Imhof, B. (2007, August). Understanding locomotion of fish from static and dynamic visualisations. In R. Lowe (Chair), *Building dynamic* *mental models from visualisations* (II). Symposium conducted at the 12th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI), Budapest, Hungary.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences, 14*, 69–110.

Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1084–1102.

Hegarty, M. (2004). Dynamic visualisations and learning: Getting to the difficult questions. *Learning and Instruction, 14*, 343–351.

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*, 722–738.

Imhof, B., Scheiter, K., & Gerjets, P. (2009). Realism in dynamic, static-sequential, and static-simultaneous visualisations during knowledge acquisition on locomotion patterns. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2962–2967). Austin, TX: Cognitive Science Society.

Kühl, T., Scheiter, K., Gerjets, P., & Edelmann, J. (in press). The influence of text modality on learning with static and dynamic visualizations. *Computers in Human Behavior. doi:10.1016/j.chb.2010.05.008.*

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

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education, 14*, 225–244.

Lowe, R. K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction, 13*, 247–262.

Lowe, R. K. (2004). Interrogation of a dynamic visualisation during learning. *Learning and Instruction, 14*, 257–274.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between word and pictures in multimedia learning. *Journal of Educational Psychology, 84*, 444–452.

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied, 11*, 256–265.

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review, 14*, 87–99.

Park, O., & Gittelman, S. S. (1992). Selective use of animation and feedback in computer-based instruction. *Educational Technology, Research & Development, 40*, 27–38.

Pfeiffer, V. D. I., Gemballa, S., Jarodzka, H., Scheiter, K., & Gerjets, P. (2009). Situated learning in the mobile age: Mobile devices on a field trip to the sea enhance knowledge on fish biodiversity. *ALT-J Research in Learning Technology, 17*, 187–199.

Resnick, L. D. (1987). Learning in school and out. *Educational Researcher, 16*, 13–20.

Rieber, L. P. (1990). Animation in computer-based instruction. *Educational Technology Research and Development, 38*, 77–86.

Salomon, G. (1979). *Interaction of media, action and learning.* San-Francisco: Jossey-Bass.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies, 45*, 185–213.

Scheiter, K., Gerjets, P., & Catrambone, R. (2006). Making the abstract concrete: Visualizing mathematical solution procedures. *Computers in Human Behavior, 22*, 9–26.

Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14*, 101–120.

Schnotz, W., & Lowe, R. K. (2008). A unified view of learning from animated and static graphics. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation: Research and design implications*. New York: Cambridge University Press.

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology: Research and Development, 53*, 47–58.

Stone-Romero, E. F. (2002). The relative validity and usefulness of various empirical research designs. In S. G. Rogelberg (Ed.), *Handbook of Research Methods in Industrial and Organizational Psychology*. Malden, MA: Blackwell Publishers.

Tversky, B., Bauer-Morrison, J., & Bétrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies, 57*, 247–262.

Van Gog, T., Paas, F., Marcus, N., Ayres, P., & Sweller, J. (2009). The mirror neuron system and observational learning: Implications for the effectiveness of dynamic visualisations. *Educational Psychology Review, 21*, 21–30.

Wong, A., Marcus, N., Ayres, P., Smith, L., Cooper, G. A., Paas, F., & Sweller, J. (2009). Instructional animations can be superior to statics when learning human motor skills. *Computers in Human Behavior, 25*, 339–347.

Chapter 9

The role of external representations in learning combinatorics and probability theory

Bas Kolloffel, Tessa H. S. Eysink, and Ton de Jong

Introduction

There are many ways to represent information in educational settings: textual descriptions, formulae, photographs, drawings and so on. A good match between the type of representation and learning demands can greatly support learning and contribute to enhanced levels of performance and understanding (Ainsworth, 2006; Greeno & Hall, 1997). Often, more than one type of representation appears to qualify for being used in a learning situation. An informed choice for one type of representation or another can be made on several grounds (Ainsworth, 2006; Scaife & Rogers, 1996). For example, a representation can be used because it causes less cognitive load compared with other representations. Representations can also be selected on the basis of the extent to which they promote clarity or reduce ambiguity (Stenning & Oberlander, 1995). Furthermore, combining two or more representational formats is assumed to have some additional effects on knowledge construction processes (Ainsworth, 1999, 2006). Different formats can complement each other or constrain the interpretation of the other (Ainsworth, 1999, 2006; van der Meij & de Jong, 2006).

External representations can be presented to students, but students can also construct representations themselves. Cox (1999) argues that the process of constructing a representation helps students to improve their knowledge, because the interaction between their internal representation and the external representation they construct, can make them aware of gaps in their internal representations they had not noticed before. Examples of activities in which students construct an external representation are: writing a summary (Hidi & Anderson, 1986), creating a drawing (Van Meter & Garner, 2005), or constructing a concept map (Gijlers & de Jong, submitted; Nesbit & Adesope, 2006; Novak, 1990).

Whether it be external representations presented to students or representations constructed by students, in either way a clear-cut recipe for which representational format to use when does not exist. Moreover, some researchers argue that the effects of representations found in one domain cannot readily be generalised to other domains (Cheng, Lowe, & Scaife, 2001). In this chapter a concise overview is presented of a series of studies that focused on the effects of representations on knowledge construction in the domain of combinatorics and probability theory. The effects of different representational formats were investigated and compared with each other in the context of simulation-based inquiry learning. In Study I, the effects of different formats used to represent the subject matter in computer simulations were investigated. In Study II, the focus was on the effects of format on learning outcomes when learners (individually or collaboratively) construct domain representations themselves. (For more details about the studies, see e.g., Kolloffel, 2008; Kolloffel, Eysink & de Jong, 2010; Kolloffel et al., 2009.)

Representations in combinatorics and probability instruction

There are several ways of representing information in the domain of combinatorics and probability theory. Some of the most commonly used formats will be discussed on the basis of the following problem, which is typical for the domain:

> Your bank distributes a random four-digit code as a personal identification number (PIN) for its credit card. What is the probability that a thief finding the card and trying to get money with it will guess the correct code in one go, and will be able to plunder your account?

One of the most common ways to represent the steps towards solving this type of problem is by means of a diagram. In a diagram, information is indexed by a two-dimensional location in a plane, explicitly preserving information about topological and structural relations (Larkin & Simon 1987) (see Figure 9.1).

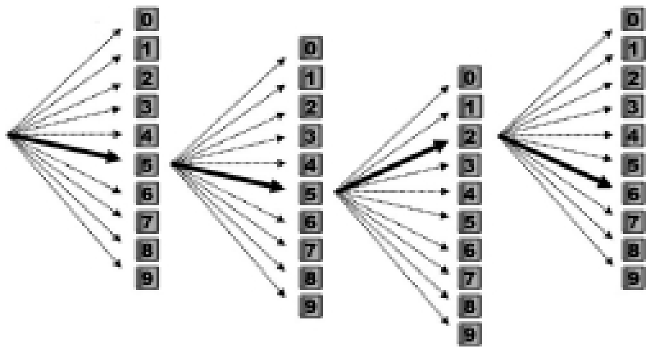


Figure 9.1 Diagram representing solution PIN-code problem.

> When selecting the first digit of a PIN-code, one can choose from ten digits: 0, 1, 2, up to 9. The chance that 5 will be selected as the first digit is equal to one out of ten. When selecting the second digit of the PIN-code, one can choose from ten digits again, because the digit that was selected the first time, can be selected again.
>
> The chance that 5 is selected as second digit of the code is therefore equal to one out of ten possible digits.
>
> The chance that 2 is selected as the third digit of the code is also equal to one out of ten possible digits, and so is the chance that 6 is selected as fourth digit.

Figure 9.2 Text representing PIN-code problem.

Diagrams are considered a powerful tool for teaching combinatorics and probability theory (e.g., Fischbein, 1987; Greer, 2001). They are especially effective in assessing the probability of various options (Fischbein, 1987).

Another way of representing the steps in the PIN-code problem is displayed in Figure 9.2.

Textual representations emphasise other relational features than those emphasised by diagrams. The use of natural language facilitates relating information in the text to everyday experiences and situations. In textual representations information is organised sequentially, preserving temporal and logical relations rather than topological and structural relations (cf. Larkin & Simon, 1987). Whereas diagrams allow simultaneous access, accessing and processing a text requires the reader to keep certain elements of the text highly activated in working memory, which is thought to burden working memory considerably (Glenberg, Meyer, & Lindem, 1987).

A third way to represent the PIN-code problem is by using an arithmetical representation (see Figure 9.3).

In this representation the underlying principle or concept is not as explicit as in a diagram or text, and therefore many learners tend to view mathematical symbols (e.g., multiplication signs) purely as indicators of which operations to perform on adjacent numbers (see, e.g., Cheng, 1999).

Each representation presented above represents the same information, that is, all of the information in one representation can be inferred from the others and vice versa. This is called *informational equivalence* (c.f. Larkin & Simon, 1987). Informational equivalence does not necessarily imply that the information can

$$p(\text{PIN} = 5526) = \frac{1}{10} \times \frac{1}{10} \times \frac{1}{10} \times \frac{1}{10}$$

Figure 9.3 Equation representing PIN-code problem.

be extracted equally easily and quickly from each representation. Therefore, the so-called *computational efficiency* (c.f. Larkin & Simon, 1987) can differ between informationally equivalent representations, which, in turn, can affect learning and knowledge construction processes.

Types of knowledge

In order to solve problems, learners need to acquire appropriate cognitive schemas. Sweller (1989, p. 458) defined a schema as 'a cognitive construct that permits problem solvers to recognise problems as belonging to a particular category requiring particular moves for solution'. A complete schema therefore rests on three pillars: conceptual knowledge, procedural knowledge, and situational knowledge. *Conceptual knowledge* is 'implicit or explicit understanding of the principles that govern a domain and of the interrelations between units of knowledge in a domain' (Rittle-Johnson, Siegler, & Alibali, 2001, p. 364). Conceptual knowledge develops by establishing relationships between pieces of information or between existing knowledge and new information. As learners' conceptual knowledge becomes sufficiently advanced, well-integrated, and automated, their ability to assess meaningful situations and predict the outcomes of complex events can develop so quickly that learners become incapable of verbalising the reasoning on which their assessments are based. This kind of conceptual knowledge is called *intuitive knowledge* and is assumed to be particularly fostered by simulation-based inquiry learning (for an extensive discussion, see Swaak & de Jong, 1996). Henceforth, the following terminology will be used: if a specific distinction is made between non-intuitive and intuitive aspects of conceptual knowledge, the former will be called *formal conceptual knowledge,* the latter *intuitive knowledge*. The term *conceptual knowledge* refers to conceptual knowledge in general. *Procedural knowledge* is 'the ability to execute action sequences to solve problems' (Rittle-Johnson et al., 2001, p. 346). *Situational knowledge* (de Jong & Ferguson-Hessler, 1996) enables students to analyse, identify, and classify a problem, to recognise the concepts that underlie the problem, and to decide which operations need to be performed to solve the problem.

Study I: which format is most effective in a computer simulation?

The aim of the first study was to investigate the effects of different representational formats used to present information to students. Five conditions were compared:

- Diagram
- Arithmetic
- Text
- Text + Arithmetic
- Diagram + Arithmetic

In theory, more combinations of representations would have been possible (e.g., Text + Diagrams), but due to screen size limitations, these combinations were not feasible in the electronic learning environments without severely hampering the readability.

The main focus was on the differential effects of formats on schema construction and cognitive load. It was assumed that the Diagram format would emphasise conceptual domain aspects and would cause low levels of cognitive load. The Arithmetic format was hypothesised to emphasise procedural domain aspects and, because of the complexity of this format, to cause relatively high levels of cognitive load. The Text format was anticipated to stress situational and conceptual domain aspects and, intrinsic to comprehensive reading, to cause much cognitive load. In the case of multiple representations it was assumed that cognitive load would be relatively low and that Text + Arithmetic would emphasise conceptual, situational and procedural domain aspects, and Diagram + Arithmetic to stimulate conceptual and procedural knowledge.

Method

Participants

A total of 123 students participated in the study: 61 boys and 62 girls. The average age of the participants was 15.61 years ($SD = 0.59$). Three participants were excluded from the analyses because their post-test scores deviated by more than 2 SDs from the mean scores within their condition. The students participated in the experiment during regular school time, so that participation was obligatory.

Design

The experiment employed a between-subjects pre-test post-test design, with the representational format in which the domain was presented (diagram ($n = 24$), arithmetic ($n = 25$), text ($n = 24$), a combination of text and arithmetic ($n = 24$), or a combination of diagram and arithmetic ($n = 23$)) as the independent variable.

Measures

KNOWLEDGE MEASURES

Two knowledge tests, a pre-test and a post-test, have been developed and used by several international research teams. The tests were validated in a series of pilot studies (see Eysink et al., 2009). The pre-test aimed at measuring the prior knowledge of the participants. The post-test contained 44 items and was specifically designed to measure the different knowledge types: formal conceptual knowledge (12 items), intuitive (13 items), procedural (14 items), and situational knowledge (5 items).

COGNITIVE LOAD

Cognitive load principles may support or even determine decisions as to which representational format to use (Leung, Low, & Sweller, 1997). Three types of cognitive load are distinguished: intrinsic, extraneous, and germane load. *Intrinsic load* is generated by the complexity of the learning material. *Extraneous load* is determined by the way in which the material is organised and presented (e.g., diagrams or text formats). *Germane load* refers to load caused by mental activities relevant to schema acquisition, such as organising the material and relating it to prior knowledge (Paas, Renkl, & Sweller, 2004; Sweller, van Merriënboer, & Paas, 1998).

After each section in the learning environment students were presented with a cognitive load questionnaire consisting of six items (see Table 9.1). The students indicated their amount of mental effort on 9-point Likert scales. Each time the cognitive load questions were presented, they appeared in a different order.

Instruction

The instructional approach used in this study is based on inquiry learning (de Jong, 2005, 2006). Computer-based simulation is a technology that is particularly suited for inquiry learning. Computer-based simulations contain a model of a system or a process. The student is enabled to induce the concepts and principles underlying the model by manipulating the input variables and observing the resulting changes in output values (de Jong & van Joolingen, 1998).

Table 9.1 Cognitive load items.

Type of cognitive load	*Item*
Intrinsic load (IL)	How easy or difficult do you consider probability theory at this moment?
Extraneous load 1 (EL1)	How easy or difficult is it for you to work with the learning environment?
Extraneous load 2 (EL2)	How easy or difficult is it for you to distinguish important and unimportant information in the learning environment?
Extraneous load 3 (EL3)	How easy or difficult is it for you to collect all the information that you need in the learning environment?
Germane load (GL)	How easy or difficult was it to understand the simulation?
Overall load (OL)	Indicate on the scale the amount of effort you had to invest to follow the last simulation.

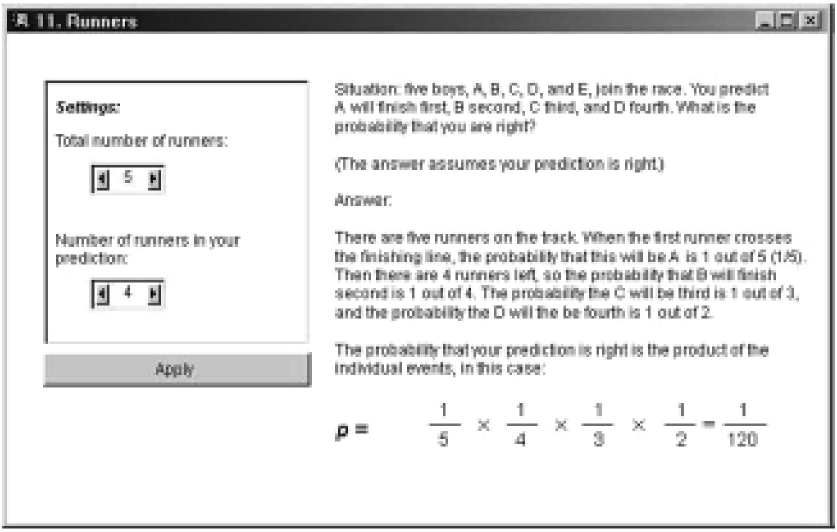


Figure 9.4 Screen dump Probe simulation (displayed is the Text + Arithmetic version of Probe).

The learning environments used in the current study, called Probe (see Figure 9.4), were created with SIMQUEST authoring software (van Joolingen & de Jong, 2003).

Probe contained simulations and a series of questions (both open-ended and multiple-choice items) and assignments. In the case of the multiple-choice items, the students received feedback from the system about the correctness of their answers. If the answer was wrong, the system offered hints about what was wrong with the answer.

Procedure

The experiment was carried out in a real school setting in one 3-hour session including a 15-minute break. Students worked individually. They were told that they could work at their own pace, and that the three hours would be more than enough to complete all assignments. The participants started the session by logging on to the electronic pre-test. It was announced that the post-test would contain more items of greater difficulty than the pre-test, but that the pre-test items nonetheless would give an indication of what kind of items to expect on the post-test. After completing the pre-test the participants received a printed introductory text about the domain. Along with the introductory text the participants received information about how to enter the learning environment. After finishing the last section of the learning environment, the participants received logon instructions for the post-test environment.

Results

Test results

A one-way ANOVA performed on the pre-test scores established that there were no differences between conditions, $F(4, 119) = 0.63$, *ns*. The post-test results are displayed in Figure 9.5. All post-test measures were analysed by one-way ANOVAs with representational format as factor.

Analysis of *post-test total scores* showed a significant effect of representational format, $F(4, 119) = 2.57$, $p < .05$. A *post hoc* least significant difference (LSD) analysis revealed that participants in the Text + Arithmetic condition outperformed participants in the Diagram condition ($p < .01$), and participants in the Diagram + Arithmetic condition ($p < .05$). With regard to *formal conceptual knowledge*, no main effect of format, $F(4, 119) = 1.56$, *ns*, was found. Neither was a main effect of format on *intuitive knowledge* observed, $F(4, 119) = 1.91$, *ns*, was found.

Analysis of *procedural knowledge* revealed a significant main effect of representational format, $F(4, 119) = 2.52$, $p < .05$. A *post hoc* LSD analysis showed that participants in the Text + Arithmetic condition outperformed participants in the Diagram condition ($p < .01$), and participants in the Diagram + Arithmetic condition ($p < .05$).

With regard to *situational knowledge* items, no main effect of representational format, $F(4, 119) = 0.66$, *ns*, was found.

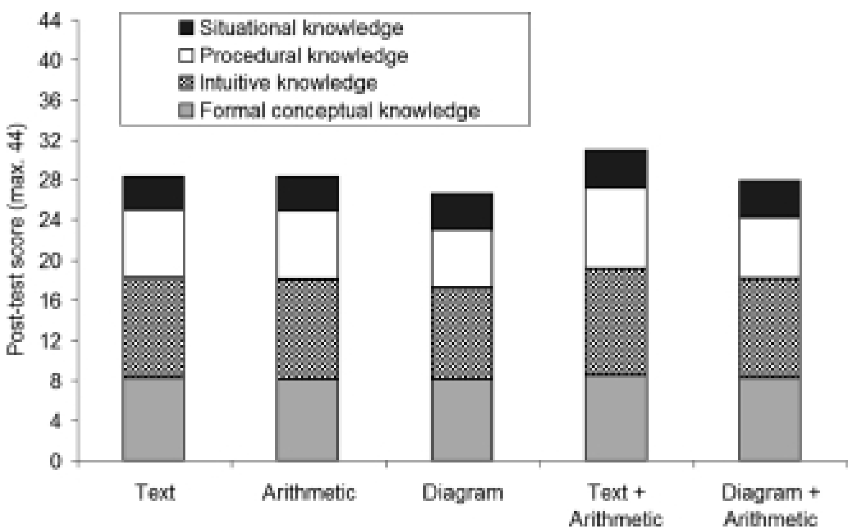


Figure 9.5 Post-test scores.

Cognitive load

Cognitive load measures are displayed in Figure 9.6.

All cognitive load measures were analysed by one-way ANOVAs with condition as factor.

Regarding *intrinsic load*, a main effect of representational format was found, $F(4, 118) = 2.70$, $p < .05$. *Post hoc* LSD analyses showed that participants in the Diagram condition experienced more intrinsic load than participants in the Text + Arithmetic condition ($p < .01$) and participants in the Diagram + Arithmetic condition ($p < .05$). Furthermore, participants in the Textual condition reported higher levels of intrinsic load compared with participants in the Text + Arithmetic condition ($p < .05$).

Differences between conditions with regard to extraneous load were observed as well, $F(4, 118) = 2.89$, $p < .05$. Participants in the Diagram condition experienced higher levels of extraneous load compared with participants in the Arithmetic condition ($p < .01$), the Text + Arithmetic condition ($p < .01$), and the Diagram + Arithmetic condition ($p < .01$). Analysis of the extraneous load sub-measures (EL 1, EL 2, and EL 3) showed that the differences between conditions with regard to extraneous load, in general, could be entirely attributed to EL 1 ($F(4, 118) = 3.55$, $p < .01$); participants in the Diagram condition reported more EL 1 compared with participants in the Arithmetic condition ($p < .01$), the Text + Arithmetic condition ($p < .01$), and the Diagram + Arithmetic condition ($p < .01$).

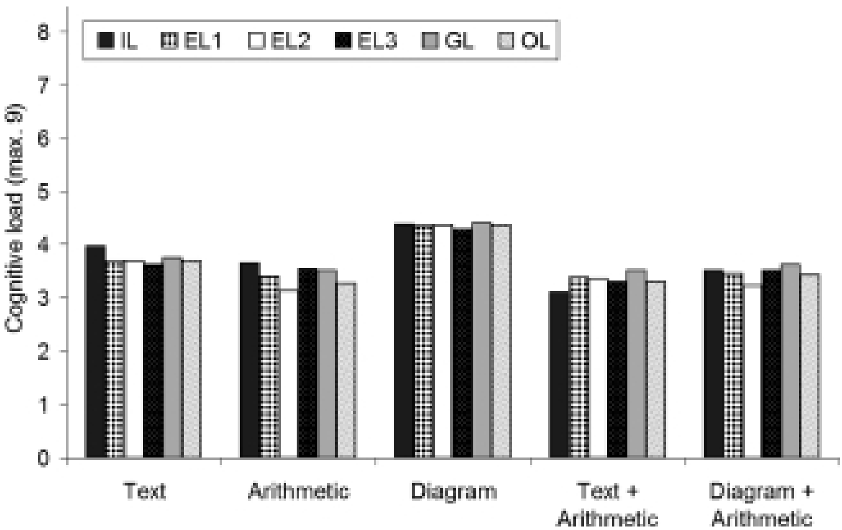


Figure 9.6 Cognitive load measures.

With regard to *germane load*, differences were found between conditions, $F(4,118)=2.89$, $p < .05$. Participants in the Diagram condition reported having more germane load as compared with participants in the Arithmetic condition ($p < .01$), the Text + Arithmetic condition ($p < .01$), and the Diagram + Arithmetic condition ($p < .05$).

Furthermore, a difference between conditions was found concerning the *overall load*, $F(4,117)=2.85$, $p < .05$. Participants in the Diagram condition had to invest more effort than participants in the Arithmetic condition ($p < .05$), and the Text + Arithmetic condition ($p < .01$).

Conclusion

The findings of the study show that the best learning outcomes were obtained by students who were presented with a combination of text and equations. Their post-test scores with regard to procedural knowledge and their post-test overall scores were significantly higher than those of students presented with tree diagrams or with a combination of tree diagrams and equations.

Students presented with tree diagrams reported the highest levels of cognitive load compared with students in the other conditions: they considered the domain more difficult (intrinsic load), found the learning environment more difficult (extraneous load), had more difficulty understanding the simulations (germane load), and had to invest more mental effort to complete their learning task. The students were all familiar with the conventions of tree diagrams, but they did not have much prior knowledge about the domain when they entered the instruction. Possibly, tree diagrams are more suited for people who already have domain knowledge (e.g., math teachers) rather than learners without domain knowledge. Tabachneck-Schijf, Leonardo, and Simon (1997) argue that experts, in particular, benefit from diagrams because for them diagrams serve as an aid to access information stored in long-term memory. Moreover, diagrams help to decrease the expert's working memory load, because elements of information that are displayed in the diagram do not have to be kept activated in working memory all the time. Furthermore, diagrams could be less suited for learners because the reasoning steps depicted in tree diagrams are quite implicit and require advanced knowledge to infer them. This might be a disadvantage in domains like combinatorics and probability theory, where problem solving requires a set of reasoning steps to be taken in a specific order. This could explain the advantage of the Text + Arithmetic format. In the textual part of the representation the learners are led step by step through an explicit, sequential line of reasoning, followed by an equation repeating these steps concisely in an arithmetical (and also sequential) way.

Study II: do representational tools support understanding in individual and collaborative learning?

Besides presenting external representations to students, students can also construct representations themselves. This activity has been found to foster deeper understanding (Cox, 1999). The format (e.g., concept maps, textual summaries, tables) used to construct a representation has been found to influence knowledge construction. For example, constructing a concept map is assumed to focus the attention of learners on the identification of concepts and their mutual relations (Nesbit & Adesope, 2006). With regard to student-constructed representations of information in the domain of probability theory, it has been found that students avoid using formal ways of representing the probability of events and prefer to use alternative forms of representation, ranging from textual statements to numerical representations (Tarr & Lannin, 2005). This finding indicates that not all formats may be equally suitable for students trying to express their knowledge. Moreover, in collaborative learning settings it was found that the format in which students constructed a representation influenced the focus of students' discourse and collaborative activities (e.g., Suthers & Hundhausen, 2003; van Drie et al., 2005). A question that has yet to be answered is which additional learning effects can be expected from constructing representations in a collaborative learning setting as compared with constructing representations in an individual learning setting, and which role representational format plays here.

In this study the effects of three, commonly used formats in the domain of combinatorics and probability theory have been compared: a conceptual, arithmetical, and a textual format. The following questions guided this study. Does the representational format of a tool affect the likelihood that students engage in constructing representations? Does format have differential effects on the quality of the representations students construct? Does the construction of a representation of a domain lead to better learning outcomes than not constructing a representation? The effects were studied in both an individual learning setting and in a collaborative learning setting.

Constructing a *conceptual representation* like a concept map is assumed to direct the students' attention to identifying concepts and their mutual relations. Therefore, it is hypothesised that this form will stimulate in particular the acquisition of conceptual knowledge rather than procedural or situational knowledge. Constructing a concept map usually does not cause students much difficulty.

Constructing an *arithmetical representation* is expected to foster the acquisition of procedural knowledge more than conceptual or situational knowledge. It is also expected that many students will have difficulty with this formal format.

The *textual format* is particularly suited to describing the domain in one's own words. This is hypothesised to stimulate the acquisition of especially situational and conceptual knowledge, though the conceptual aspects might not be as emphasised as with constructing a conceptual representation. It is anticipated that students find text an easy way to express themselves.

Method

Participants

The data collection of the study in the collaborative setting took place exactly one year after the study with individual students at the same school. In the *collaborative learning study*, in total 128 secondary education students entered the experiment, and the data of 61 pairs could be analysed. The average age of these 56 boys and 66 girls was 14.62 years ($SD = .57$). In the *individual learning study*, 95 secondary education students, 50 boys and 45 girls, participated. The average age of the students was 14.62 years ($SD = .63$). The students attended the experiment during regular school time; therefore, participation was obligatory.

Design

The experiments employed a between-subjects design with the format of the provided representational tool (conceptual, arithmetical, or textual) as the independent variable. Students were randomly assigned to conditions. Of the 61 pairs in the collaborative setting, 22 pairs were in the Conceptual condition, 19 pairs in the Arithmetical condition, and 20 pairs in the Textual condition. Of the 95 students in the individual learning setting, 33 were in the Conceptual condition, 30 in the Arithmetical condition, and 32 in the Textual condition.

Learning environment

The simulation-based learning environment used in these studies was the Text + Arithmetic version of Probe that was used previously in Study I.

Representational tools

Students were encouraged to construct a representation in which they summarised what they considered to be the main points of the domain (e.g., principles, variables and their mutual relationships). Each student/dyad could use a *representational tool* to create the representation. A representational tool is an electronic on-screen tool designed to construct, discuss, and share external representations (Suthers & Hundhausen, 2003). There were three types of representational tools, one for each experimental condition: (a) a conceptual

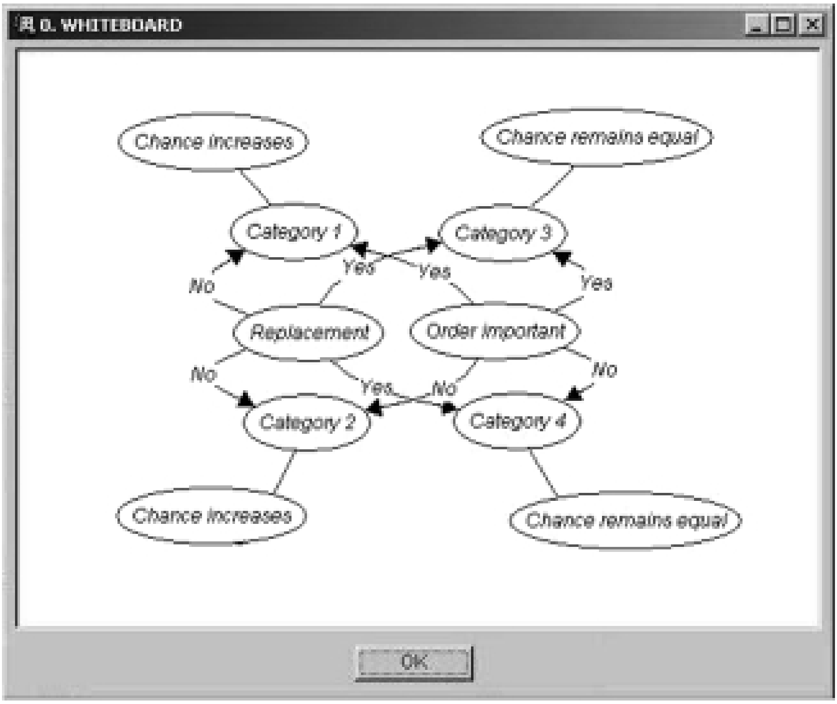


Figure 9.7 Conceptual representational tool.

representational tool, (b) an arithmetical representational tool, and (c) a textual representational tool. The *conceptual representational tool* (see Figure 9.7) could be used to create a concept map of the domain. Students could draw circles representing domain concepts and variables. Keywords could be entered in the circles. The circles could be connected to each other by arrows indicating relations between concepts and variables. The nature of these relations could be specified by attaching labels to the arrows.

In the *arithmetical representational tool* (see Figure 9.8), students could use variable names, numerical data, and mathematical operators (division signs, equation signs, multiplication signs, and so on) in order to express their knowledge.

Finally, the *textual representational tool* (see Figure 9.9) resembled simple word processing software, allowing textual and numerical input.

In the current studies, the representational tools were intended as means to *support* students while learning. Therefore, the use of the representational tools was not obligatory, although students were strongly advised to use the tool and they were informed that using the tool would help them to better prepare themselves for the post-test.

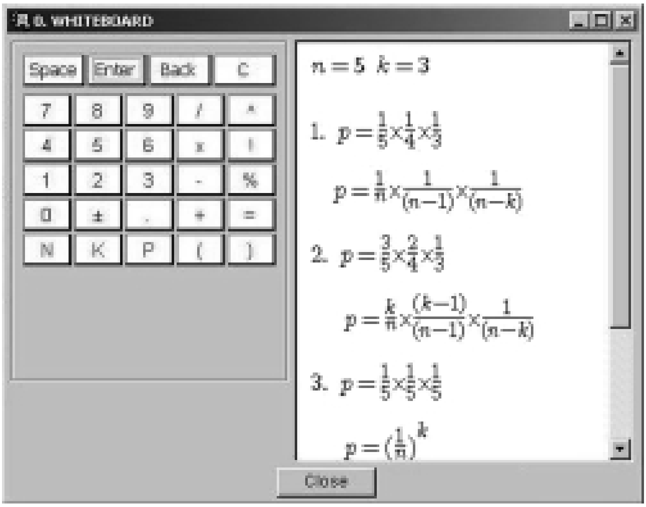


Figure 9.8 Arithmetical representational tool.

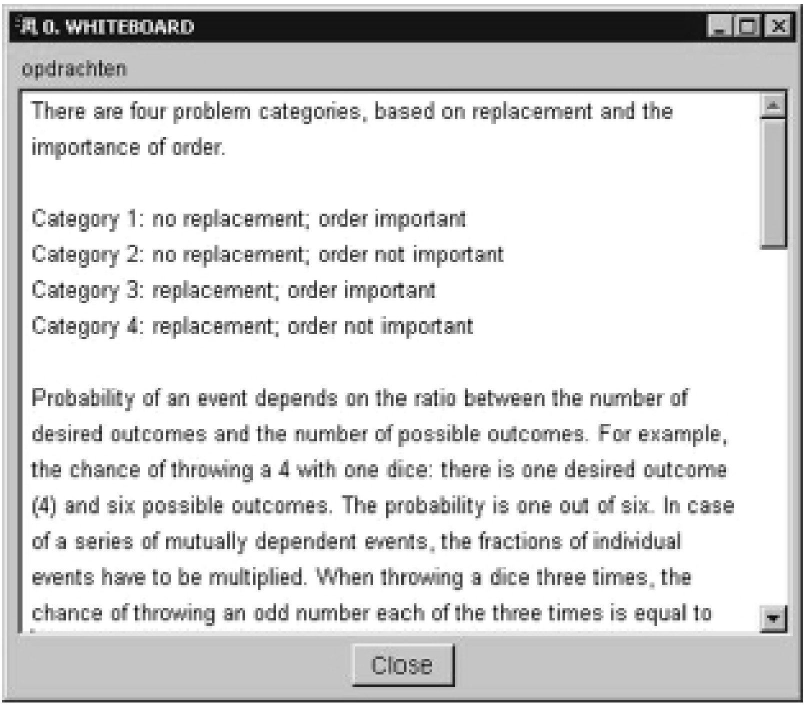


Figure 9.9 Textual representational tool.

Knowledge measures

Two knowledge tests were used in this experiment: a pre-test and a post-test. The tests contained 12 and 26 items respectively. The post-test was a shortened version of the one used in Study I.

Procedure

The experiments were carried out in a real school setting in three sessions, each separated by a one-week interval. The procedures in both the individual and collaborative setting were identical.

The first session started with a short introduction followed by the pre-test. In both the individual and the collaborative setting, students completed the pre-test individually. At the end of the pre-test the students received a printed introductory text in which the domain was introduced. The duration of the first session was limited to 50 minutes. During the last 15 minutes of the session, the students received an explanation of how their representational tool could be operated and they could practise with the tool.

During the second session, students worked with the learning environment which contained simulations and a series of questions and assignments. Each learning environment contained a representational tool that could be used to construct a domain representation. The duration of this session was set at 70 minutes. Students in the individual learning setting worked alone. In the collaborative learning setting students were allowed to choose their partner themselves. Communication between students was on a face-to-face basis: the collaborating students were sitting next to each other, using the same computer terminal. Despite the possibility of following a non-linear path through the learning environment, students were advised to keep to the order of sections because they built upon each other.

The third session was set at 50 minutes. First, students were allowed to use the learning environment for 10 minutes in order to refresh their memories with regard to the domain. Then all students had to close their domain representations and learning environments, and had to complete the post-test. In both the individual and the collaborative setting, students completed the post-test individually.

Data preparation

The domain representations constructed by the students were scored by means of a scoring rubric. This rubric revolved around the principle that scoring of the domain representation should not be biased by the representational format of the representational tool, that is, all types of representations should be scored on the basis of exactly the same criteria. The maximum number of points that could be assigned on the basis of the rubric was eight points. The rubric was used to assess whether domain representations reflected the concepts of replacement and order, presented calculations, referred to the concept of probability, indicated the effect of size of (sub)sets on probability, and the effects of replacement and order on probability.

Results

Use of representational tools

The percentages of students in each condition who did or did not use the representational tool to construct a domain representation are displayed in Figure 9.10.

The overall picture is that about 50 per cent of the students (both individuals and dyads) provided with a conceptual or textual tool used the tool. Of students (both individuals and dyads) provided with the arithmetical tool, about 20 per cent actually used the tool. A Chi-Square analysis showed that these differences between conditions were significant ($p < .01$). Compared with students in the Arithmetical condition, students in the Conceptual condition used their tool more often ($p < .01$) and so did students with a textual tool ($p < .05$). No difference was observed between the Conceptual and the Textual condition.

In Figure 9.11 the average quality scores of the constructed representations are displayed. In the case of representations constructed by pairs, the representations are considered a group product and therefore the quality scores are assigned to pairs and not to individuals. All representations were scored by two raters who worked independently from each other. The inter-rater agreement was .89 (Cohen's Kappa) for the individual setting and .92 for the collaborative setting.

A two-way ANOVA with setting (individual vs. collaborative learning) and condition as factors showed that with regard to quality scores there was

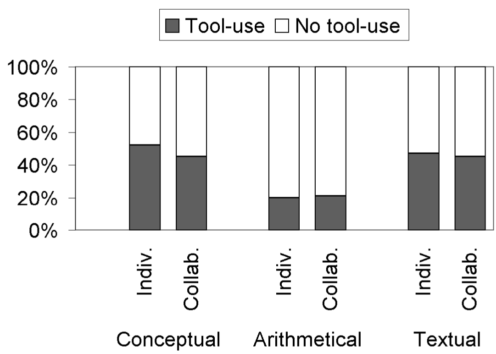


Figure 9.10 Percentage of individuals/dyads in each condition who did or did not use their representational tool.

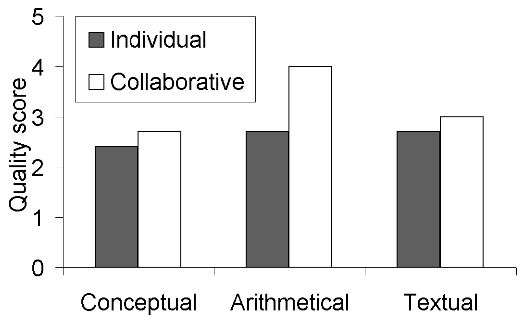


Figure 9.11 Quality scores of constructed representations.

no main effect of setting, no main effect of condition, and no interaction effect.

Knowledge measures

Two measures of knowledge were obtained: prior knowledge (pre-test score) and post-test score. Both in the collaborative and in the individual setting students completed the tests individually. The tests were validated in a series of pilot studies (Eysink et al., 2009). The reliability, Cronbach's α, of the pre-test was .64. This reliability is low, but sufficient for the purpose of verifying that students did not have too much prior knowledge and that there were no differences between settings and/or conditions. The reliabilities of the formal conceptual, intuitive, procedural, and situational knowledge scales were respectively: .76, .74, .74, and .67. Furthermore, students were asked for their latest school report grade in mathematics. This grade, which can range from 1 (very, very poor) to 10 (outstanding), was interpreted as an indication of the student's general mathematics achievement level.

Three-way ANOVAs with setting (individual or collaborative), condition (Conceptual, Arithmetical, Textual), and tool-use (Tool-use or No-tool-use) as factors were performed to test for a priori differences with respect to math grade (general mathematics achievement level) and pre-test score (prior knowledge).

A difference regarding *math grade* was observed with respect to setting ($p < .05$) and tool-use ($p < .01$). No main effect of condition was found. An interaction between setting and tool-use ($p < .01$) was observed. On average, the math grades of students in the collaborative learning setting were somewhat higher compared with the individual students. Furthermore, in the individual learning setting it was observed that students who used their representational tool had higher math grades compared with individuals who did not use the tool. The math grades of individuals who used the tool were equal to those of students in the collaborative setting.

With regard to *pre-test scores*, no main effects were found for setting, condition or tool-use. No interaction effects were observed either.

Therefore, in the analyses of post-test measures only math grade was entered as a covariate. In Figure 9.12 an overview of post-test measures is provided.

MANOVAs with setting (individual or collaborative), condition (Conceptual, Arithmetical or Textual; not separately displayed in Figure 9.12) and tool-use as factors and math grade as covariate were applied to post-test measures.

The outcomes of the analyses showed a difference with regard to *setting* ($p < .001$). Students in the collaborative learning setting showed higher scores with respect to intuitive knowledge ($p < .001$), situational knowledge ($p < .01$) and the post-test overall score ($p < .001$). No difference between individuals and dyads was found with respect to formal conceptual or procedural knowledge.

No differences were observed for *condition*.

An interaction was observed between setting and tool-use ($p < .05$). This interaction concerned situational knowledge ($p < .01$) and the post-test overall score ($p < .05$). The interaction indicates that with regard to situational knowledge and post-test overall scores, students in the collaborative setting in general outperformed individual students, but in cases where individual students constructed a representation their scores equalled those of students in the collaborative setting.

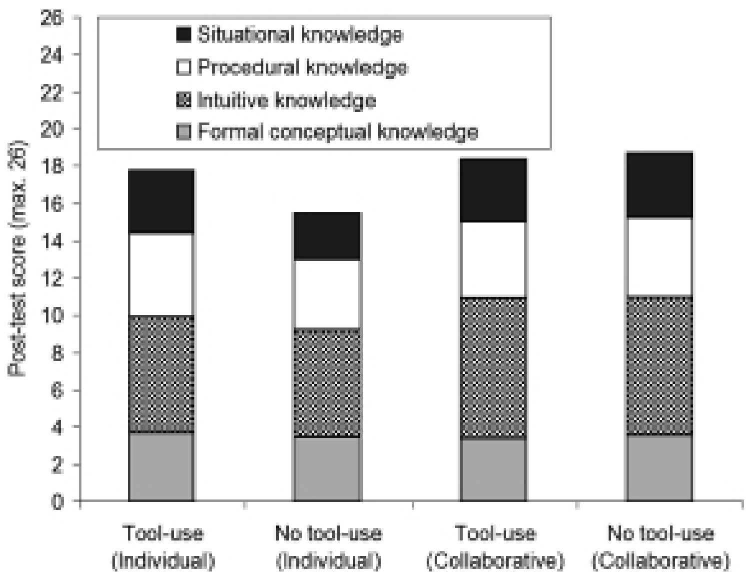


Figure 9.12 Post-test measures.

Conclusion

With respect to their inclination to use a representational tool, students in both the collaborative and individual setting were found to be very like-minded: representational tools with a conceptual or textual format were used much more readily (around 50 per cent use) compared with the arithmetical format (around 20 per cent use). In both settings the formats of the tools did not lead to differential effects on the quality of the constructed representations. Perhaps the textual and the conceptual formats are more close to the code in which students think (and/or talk) and explain the domain to themselves, or maybe students consider those formats more suited to expressing their knowledge to the outside world.

With regard to learning outcomes, it was found that, in general, post-test scores in the collaborative setting were higher than in the individual setting. Only individual students who engaged in constructing a domain representation in general equalled post-test scores of students in the collaborative setting. It was observed that students learning collaboratively and students constructing domain representations in the individual setting both showed enhanced levels of situational knowledge. This type of knowledge is a prerequisite for going beyond the superficial details of problems in order to recognise the concepts and structures that underlie the problem (e.g., Fuchs et al., 2004). Furthermore, students in the collaborative learning setting showed enhanced levels of intuitive knowledge. The observation that collaborative students (regardless of whether or not they constructed a representation) outperformed individuals (even those who did construct a representation), implied that (a) intuitive knowledge was enhanced by collaborative learning and (b) the activity of constructing representations was not sufficient for individual students to equal the levels of intuitive knowledge of students who worked collaboratively.

Theoretical implications and general conclusions

The overall aim of the studies presented here was to establish if and how representational format affects knowledge construction. It was assumed that knowledge of the domain of combinatorics and probability has three components: conceptual knowledge, procedural knowledge, and situational knowledge. In the studies presented here it was found that each of these types of knowledge could be influenced.

Enhanced levels of *conceptual knowledge* were observed when students learned in a collaborative setting. In particular, they showed enhanced levels of intuitive knowledge. Not much is known about how learners acquire intuitive knowledge, but there are indications that it is fostered particularly by processes of interpretation and sense-making (Gijlers & de Jong, submitted; Zhang et al., 2004). However, more research is needed before solid conclusions can be drawn about the relation between collaborative learning and the acquisition of intuitive knowledge.

It was observed that the acquisition of *procedural knowledge* is influenced by the representational format in which the domain is presented to the students. A combination of a textual and an arithmetical format significantly improved levels of procedural knowledge.

It was observed that the representational format used to present the domain to students does not affect *situational knowledge*, but it was found that constructing a domain representation, regardless of the format, is associated with significantly higher levels of situational knowledge. It was also observed that collaboration is beneficial for situational knowledge. In that case, it did not make a difference whether or not students constructed a domain representation; their situational knowledge scores were equal anyway to individual students who had constructed a domain representation.

Acknowledgement

The authors gratefully acknowledge the Netherlands Organisation for Scientific Research (NWO) for funding this study (project number 411-02-162).

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*, 131–152.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183–198.

Cheng, P. C. H. (1999). Unlocking conceptual learning in mathematics and science with effective representational systems. *Computers & Education, 33*, 109–130.

Cheng, P. C. H., Lowe, R. K., & Scaife, M. (2001). Cognitive science approaches to understanding diagrammatic representations. *Artificial Intelligence Review, 15*, 79–94.

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

de Jong, T. (2005). The guided discovery principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 215–229). Cambridge (UK): Cambridge University Press.

de Jong, T. (2006). Technological advances in inquiry learning. *Science, 312*, 532–533.

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

de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*, 179–202.

Eysink, T. H. S., de Jong, T., Berthold, K., Kolloffel, B., Opfermann, M., & Wouters, P. (2009). Learner performance in multimedia learning arrangements: An analysis across instructional approaches. *American Educational Research Journal, 46*, 1107–1149.

Fischbein, E. (1987). *Intuition in science and mathematics: An educational approach*. Dordrecht, The Netherlands: Reidel.
Fuchs, L. S., Fuchs, D., Prentice, K., Hamlett, C. L., Finelli, R., & Courey, S. J. (2004). Enhancing mathematical problem solving among third-grade students with schema-based instruction. *Journal of Educational Psychology, 96*, 635–647.
Gijlers, H., & de Jong, T. (submitted). Facilitating collaborative inquiry learning with shared concept maps and proposition tables.
Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26*, 66–83.
Greeno, J. G., & Hall, R. P. (1997). Practicing representation: Learning with and about representational forms. *Phi Delta Kappan, 78*, 361–367.
Greer, B. (2001). Understanding probabilistic thinking: The legacy of Efraim Fischbein. *Educational Studies in Mathematics, 45*, 15–33.
Hidi, S., & Anderson, V. (1986). Producing written summaries: Task demands, cognitive operations, and implications for instruction. *Review of Educational Research, 56*, 473–493.
Kolloffel, B. (2008). *Getting the picture: The role of external representations in simulation-based inquiry learning*. Doctoral dissertation. Enschede, The Netherlands: University of Twente (Electronic version available at: http://purl.org/utwente/60341).
Kolloffel, B., Eysink, T. H. S., & de Jong, T. (2010). The influence of learner-generated domain representations on learning combinatorics and probability theory. *Computers in Human Behavior, 26*, 1–11.
Kolloffel, B., Eysink, T. H. S., de Jong, T., & Wilhelm, P. (2009). The effects of representational format on learning combinatorics from an interactive computer-simulation. *Instructional Science, 37*, 503–517.
Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65–99.
Leung, M., Low, R., & Sweller, J. (1997). Learning from equations or words. *Instructional Science, 25*, 37–70.
Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*, 413–448.
Novak, J. D. (1990). Concept maps and Vee Diagrams: Two metacognitive tools to facilitate meaningful learning. *Instructional Science, 19*, 29–52.
Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science, 32*, 1–8.
Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*, 346–362.
Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies, 45*, 185–213.
Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science, 19*, 97–140.
Suthers, D. D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *Journal of the Learning Sciences, 12*, 183–218.

Swaak, J., & de Jong, T. (1996). Measuring intuitive knowledge in science: the development of the what-if test. *Studies in Educational Evaluation, 22*, 341–362.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology, 81*, 457–466.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

Tabachneck-Schijf, H. J. M., Leonardo, A. M., & Simon, H. A. (1997). CaMeRa: A computational model of multiple representations. *Cognitive Science, 21*, 305–350.

Tarr, J. E., & Lannin, J. K. (2005). How can teachers build notions of conditional probability and independence? In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 215–238). New York, NY: Springer.

van der Meij, J., & de Jong, T. (2006). Supporting students' learning with multiple representations in a dynamic simulation-based learning environment. *Learning and Instruction, 16*, 199–212.

van Drie, J., van Boxtel, C., Jaspers, J., & Kanselaar, G. (2005). Effects of representational guidance on domain specific reasoning in CSCL. *Computers in Human Behavior, 21*, 575–602.

van Joolingen, W. R., & de Jong, T. (2003). SimQuest: Authoring educational simulations. In T. Murray, S. Blessing, & S. Ainsworth (Eds.), *Authoring tools for advanced technology educational software: Toward cost-effective production of adaptive, interactive, and intelligent educational software* (pp. 1–31). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review, 17*, 285–325.

Zhang, J. W., Chen, Q., Sun, Y. Q., & Reid, D. J. (2004). Triple scheme of learning support design for scientific discovery learning based on computer simulation: experimental research. *Journal of Computer Assisted Learning, 20*, 269–282.

Chapter 10

Symbolising and the development of meaning in computer-supported algebra education

Koeno Gravemeijer, Michiel Doorman, and Paul Drijvers

Introduction

One of the most salient characteristics of mathematics is its logical structure; one would therefore expect mathematics to be a subject that is easy to teach. Reality, however, shows that this is not the case. Mathematics proves to be hard to learn for many students. A common explanation is that mathematics is too abstract. Sfard (1991) offers a deeper explanation by showing that there is more to mathematics than just the rules of logic. She points to the epistemological nature of mathematical knowledge. Reflecting upon the history of mathematics, she construes a dual nature of mathematical conceptions: a structural conception, and an operational conception.[1] The first concerns mathematical objects, the latter mathematical processes. We may start by elucidating the structural conception that encompasses the notion of mathematical objects. Sfard (1991) observes that, although it is commonplace to speak of 'a function such that . . .' in a similar manner as a physicist speaks about the existence of certain subatomic particles, mathematical objects are very different from physical or material objects. Mathematical constructs such as functions are inaccessible to our senses.

> [E]ven when we draw a function or write down a number, we are very careful to emphasise that the sign on the paper is but one among many possible representations of some abstract entity [. . .].
>
> (Sfard, 1991, p. 3)

She goes on to say that, the fact that such immaterial mathematical objects are experientially real for mathematicians, while novices may be unable to 'see' these objects, may be one of the reasons why mathematics appears practically inaccessible. In contrast to the static structural conception, the operational conception concerns processes, algorithms and actions. Sfard (1991) argues that one may speak of the dual nature of mathematics because operational and structural conceptions are both incompatible and complementary. Following Sfard, we may elucidate the dual nature with two examples, number, and function. The operational conception of number concerns the activity of counting. The structural aspect of number concerns number as a quantity, or the ‘quantity number’ that describes a property of a set (or the class of all sets of the same finite cardinality). A function may operationally be seen as a computational process, as a recipe to transform one number into another (or to link one number to another). Structurally, a function can be thought of as a set of ordered number pairs.

Sfard (1991) illustrates the operational and the structural conception of function by discussing three different representations of a function shown in Figure 10.1. The computer program may be associated with an operational conception rather than with a structural conception, since it presents the function as a computational process. The graph, on the other hand, can be seen as a concise depiction of the corresponding set of ordered number pairs. In this manner, the whole set of number pairs can be grasped as an integrated whole. Moreover, the graph may be used to talk about certain characteristics of a fourth-order polynomial function, while treating the function as an object. The algebraic representation can be interpreted in both ways, as a description of a computation, or as a relation between two magnitudes. Sfard (1991) observes that, historically, the operational aspect preceded the structural aspect, and argues that the same should be the case for the individual learning process, because the structural approach is more abstract than the operational. To strengthen her point, she refers to Piaget who argues:

> the (mathematical) abstraction is drawn not from the object that is acted upon, but from the action itself. It seems to me that this is the basis of logical and mathematical abstraction.
>
> (Piaget, cited by Sfard, 1991, p. 17)

Other scholars come to similar conclusions, e.g., Dubinsky (1991), who speaks of ‘encapsulation’, Tall (1991), who introduces the term ‘procept’ to stress the process element that is present in most concepts, and Freudenthal (1991), who speaks of cumulative process of common sense experiences that are organised in

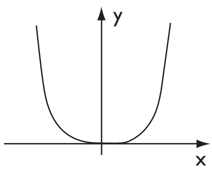
Graph	Algebraic expression	Computer program
	$y = 3x^4$	10 INPUT X 20 Y = 1 30 FOR I = 1 TO 4 40 Y = Y * X 50 NEXT I 60 Y = 3 * Y

Figure 10.1 Different representations of a function. (Copied from Sfard (1991, p. 6).)

rules, which become common sense of a higher order. For reasons of readability, however, we will stick to Sfard's terminology. History shows that the transition from an operational conception to a structural conception is very difficult. Here, the long history of transition of the so-called rhetorical algebra into the syncopated and symbolic algebra may serve as an example. According to Sfard (1991), we may discern three stages in the transition from an operational conception to a structural conception:

- interiorisation, which refers to becoming proficient with executing the process, and becoming able to carry it out in your mind;
- condensation (curtailing and compressing);
- reification (the actual shift to the structural conception, or to the mathematical object).

To be able to reason with mathematical objects, one must be able to work with the products of certain processes without bothering about the processes themselves. This requires a profound insight in those processes, and a certain amount of proficiency with those processes. Finally, the transition from computational operations to abstract objects is a long and inherently difficult process. There has to be a transition stage in which one starts to treat a process as an object before it has become an object. One must try, for instance, to manipulate a function as an object-like entity in order to come to see it as an object. Sfard observes that this requires a struggle that may appear too challenging for some.

Visual representations

In this chapter we will advance the view that visual representations and models can play a significant role in supporting students in this struggle. We want to stress, however, that this role of visual representations differs significantly from the conventional manner in which external representations are put to use. In fact, the use of visual representations and models that we propose can be seen as a reaction to the problems with the more conventional role of models in mathematics education. We will therefore start by discussing the use of conventional models and their limitations.

For a long time, mathematics educators have been designing and using external representations to make the formal, abstract mathematics accessible for students. These, what we may call 'didactical', representations have played a key role in the kind of mathematics education in which mathematics is seen as an objective body of knowledge that students have to acquire. Tacit and visual models were seen as powerful means to support learning for understanding. By acting with well-designed concrete models, students were expected to discover the mathematics that was embedded in the models. In relation to this, Cobb, Yackel, and Wood (1992) speak of a representational view. Mathematics educators, who use tactile models and visual representations in this manner, implicitly or explicitly hold the view that learning is characterised as a process in which students construct mental representations that mirror the mathematical features of external representations. But the problem with this approach is that the meaning of external representations is dependent on the knowledge and understanding of the interpreter. This creates a problem that is known as 'the learning paradox' (Bereiter, 1985), which can be captured with the question: how is it possible to learn the symbolisations that you need to come to grips with new mathematics, if you need to have mastered this new mathematics in order to be able to understand those very symbolisations?

If we follow Sfard's (1991) line of reasoning, mathematics educators who have a structural conception of functions, for instance, conceive functions as mathematical objects. That is to say, they experience them as real objects that you can act and reason with – just like most adults experience numbers as real objects. With the construction of various mathematical objects mathematicians create a body of knowledge that they experience as real. Cobb et al. (1992) argue that since mathematics educators experience mathematics as an objective body of knowledge, they are inclined to project their own mathematical knowledge and understanding into external representations that are to make that mathematics accessible for students. In this manner, they continue, a dualism is erected between mathematics in students' heads and mathematics in the external environment, which then leads to the aforementioned learning paradox. They go on to say that the representational view is problematic, since

> the assumption that students will inevitably construct the correct internal representation from the materials presented implies that their learning is triggered by the mathematical relationships they are to construct before they have constructed them.
>
> (Cobb et al., 1992, p. 5)

Or, to use Sfard's terms, the students are to 'see' mathematical objects, which they have not yet constructed.

We may argue, however, that the learning paradox dissolves when we adopt a more dynamic view of learning that is in line with a learning process in which operational conceptions transform into structural conceptions. Within such a view, mathematical symbols and models may be developed in a bottom-up manner (e.g., Doorman, 2005). Meira (1995) observes that, in the history of mathematics, mathematical symbols did not suddenly appear in their full-fledged form. Instead, these symbols grew out of informal situated forms of symbolising that developed over time in a dialectic process in which symbolisations (visual representations) and meaning co-evolved. Therefore, he suggests that students should develop symbols and meaning in a similar process and advocates for

> [an] activity-oriented view that takes cultural conventions, such as notational systems, to shape in fundamental ways the very activities from which they emerge, at the same time that their meanings are continuously transformed as learners produce and reproduce them in activity.
>
> (Meira, 1995, p. 270)

Following Meira (1995), we may envision a dynamic learning process in which symbolisations and meaning co-evolve, and in which the ways that symbols are used and the meanings they come to have, are seen to be mutually constitutive.

A similar observation is made by Latour (1990), who speaks of a 'cascade of inscriptions'. He introduces the word inscription to make a distinction between the material part of a visual representation and the meaning that is attributed to that inscription in some social practice. Inscriptions will mostly be marks on paper or on a computer screen, but may also consist of tactile objects. By making this distinction between inscription and meaning explicit, Latour underscores that symbols or inscriptions do not come with meaning. In our view, this distinction is more powerful than the common distinction between internal and external representations. For the notion of an external representation brings with it the idea of meaning as something that is represented. In contrast, the label inscription points to the fact that the inscription as such is devoid of meaning. Meaning is attributed to inscriptions by actors who use a given inscription in a certain social practice. As a consequence, the same inscription may have different meanings in different social practices. The letter 'x', for instance, has a different meaning in language lessons or algebra. Moreover, the letter 'x' may have different meanings in various practices in algebra. When solving $8x = 7$ for x, for instance, x may be thought of as a given, although unknown, number, which multiplied by 8 will get you 7. In another social practice, in a mathematics classroom that is reasoning about 'p' as a parameter in $f(x) = px + 12$, x will have to be thought of as a variable that encompasses all possible values of a domain set. A complicating element is that the actors may participate on different levels in such a social practice of mathematics in school. Even if the discourse is about finding an unknown x in early algebra, some students may (already) be thinking of a variable that may assume a certain value for which the expression '$8x = 7$' is true. In another instructional practice with somewhat older students, some of them may be struggling to make sense of parameters because they only have at their disposal a conception of x as a placeholder for an unknown number.

We may conclude this analysis by stating that there is a need for a bottom-up approach to helping students develop meaningful symbols. Following Meira (1995) and others, such as Roth and McGinn (1998), we may conclude that the way symbols and meaning emerge has to take the form of a dialectic process in which symbols and meaning co-evolve.

Emergent modelling

A heuristic that aims at supporting the design of such a process is that of 'emergent modelling' (Gravemeijer, 1999). As a point of clarification, we want to note that the emergent modelling instructional design heuristic does not aim at modelling as a goal in and of itself. Instead, modelling is employed in the service of a long-term learning process in which mathematical meaning and mathematical forms of representation are developed. Consequently, emergent modelling differs from what is typically referred to as 'mathematical modelling' (as a synonym of 'applied mathematical problem solving'). Mathematical modelling may be described as translating contextual problems into mathematical models by employing the mathematical tools that are ready at hand for the problem solver, whereas the kind of modelling that characterises emergent modelling is better described as a process of organising. The modelling activity, typically, has the character of coming to grips with problem situations and the embedded mathematics by way of organising and structuring them. In the emergent modelling process there is no such distinction between the model and the situation model – at least initially. The idea is that, in the first phase of the learning process, informal ways of modelling emerge while students are trying to come to grips with contextual problem situations and the embedded mathematics. Later, subsequent ways of modelling will start to serve as a basis for developing more formal mathematical knowledge. Finally, this more formal mathematical knowledge may eventually be experienced as mathematics that is ready to hand in a process of mathematical modelling.

The long-term process of emergent modelling may be described in the following manner. First a model is constituted as a context-specific model of acting in a given situation. Then, gradually, the students are stimulated to shift their attention towards the mathematical relations involved. As a result, the students may start to build a framework of mathematical relations. Then, the model begins to derive its meaning for the students from this emerging framework of mathematical relations, and the model becomes more important for them as a base for reasoning about the mathematical relations involved than as a way to symbolise mathematical activity in a particular setting. In this sense, the role of the model gradually changes as it takes on a life of its own. As a consequence, the model can become a referential base for more formal mathematical reasoning.

Within the above developmental progression, we can discern four types of activity, which we may denote as levels even though they do not involve a strictly ordered hierarchy (Gravemeijer, 1999):

- activity in the task setting, in which interpretations and (situation-specific) solutions depend on understanding of how to act in that setting;
- referential activity, in which models-of refer to activity in the setting described in instructional tasks;

– general activity, in which an orientation on mathematical relations and strategies makes it possible to act and reason independently of situation-specific imagery, and models start to function as models for more formal mathematical reasoning;
– formal mathematical reasoning that is no longer dependent on the support of models-for mathematical activity.

In conclusion, we may observe a transition from a model of informal situated activity to a model for more formal mathematical reasoning. Thus, we may characterise the emergent models heuristic as a model-of/model-for transition, in which 'the model' is understood as a global overarching concept. In the concrete elaboration of an instructional sequence, this overarching model takes on various manifestations, which we may call sub-models. The idea is that the students will use those sub-models as tools, and that each activity with a newer sub-model or tool is experienced as a natural extension of the activity with the earlier sub-model/tool. Consequently, the formal mathematical symbols that will eventually be introduced will be rooted in concrete activities of the students. The dynamic character of this process justifies the term emergent models. The meaning of the label, however, is broader. It refers both to the process by which models emerge, and to the process by which these models support the emergence of more formal mathematical knowledge.

In summary, we may observe that there are three interrelated processes. First, there is the overarching model, which first emerges as a model of informal mathematical activity, and then gradually develops into a model for more formal mathematical reasoning. Second, the model-of/model-for transition involves the extension of the mathematical reality, which can be called formal in relation to the original starting points of the students. This extension (which is new for the students, not for the teachers) consists of mathematical objects that derive their meaning from a network of mathematical relations. In other words, the model-of/model-for transition coincides with a shift towards a structural conception (Sfard, 1991). Third, in the concrete elaboration of the instructional sequence, there is a series of sub-models, which build on each other. And, although the term 'emergent modelling' may suggest differently, these sub-models or inscriptions will – in practice – not be invented by the students. Instead, the emergent modelling process is organised by an instructional sequence and by the teacher who introduces each new sub-model when he or she thinks the development of the students allows for it. That is to say, their mathematical understanding has progressed to a level at which they could, in principle, invent that sub-model themselves. So even though the goal is to support a bottom-up process, reality may be different. In practice there may be a tension between the planning of the instruction, which – at least in part – will be done in advance, and the actual development of the students. The teacher therefore has to carefully monitor whether the students experience each new sub-model as fitting their own current thinking. In this respect, Gravemeijer, Bowers, and Stephan (2003) introduced the term 'imagery'; the students have to be able to 'see' the earlier activity of working with the former sub-model in their way of working with the new model (e.g., Doorman & Gravemeijer, 2009).

Emergent modelling applied to functions

In the following we will discuss briefly how the emergent modelling heuristic may be applied to the design of instruction on mathematical functions. In doing so, we will zoom in on a part of a potential learning route in which a computer applet is used as a means of support for fostering the transition from a conception of a function as a calculational recipe to that of a set of ordered number pairs (Drijvers et al., 2007).

An important guideline that the emergent modelling heuristic offers instructional designers is the advice to analyse the potential endpoints of the instructional sequence under consideration in terms of mathematical objects and a corresponding framework of mathematical relations. To use Sfard's (1991) terms, we will be aiming for a transition from an operational to a structural conception. So the question arises, what is the mathematical object we are aiming for? In the case of functions, a static description is offered by Sfard's historical analyses, which show that we are aiming for an object that fits the formal definition of a function as a set of ordered number pairs. She further found that the structural conception manifests itself in the way one operates with function as an object. According to the exposition on the emergent modelling design heuristic presented above, eventually the students will have to reach a level of more formal mathematical reasoning, which is no longer dependent on the support of models-for mathematical activity. We may, however, have a closer look at what the latter implies. In some cases, the more elaborate sub-models slide under some formal notation. In the instructional sequence for addition and subtraction up to 100, for instance, a tactile representation of numbers with beads on a so-called arithmetic rack[2] is at the core of the sequence (Gravemeijer, 2008). Towards the end of the sequence, actions with the beads on the arithmetic rack, and the corresponding forms of reasoning, are described schematically on paper. Then, gradually, those inscriptions on paper start to function as the next sub-model, which, finally, is replaced by a description with standard numerical expressions. In other cases, some graphical representations are never completely abandoned. The latter will also be the case with functions, since it is quite common to revert to tables or graphs to support one's reasoning, even if one is working at a very abstract level.

When we start to look for an overarching model, we find that the notion of a function is intimately tied to three different representations, algebraic expressions, tables and graphs (Janvier, 1987). It will be clear that all three forms of representation have to be part of the learning process, and part of what it means to have a structural understanding of functions. At first sight this seems to ask for three emergent models. In our view, however, tables are rather straightforward representations that do not require an elaborated learning process. This leaves us with two models. We will, therefore, need two learning routes, one on the topic of algebraic expressions and one on graphs. The latter, which concerns Cartesian graphs as a tool to reason about the covariation of two variables, may start with empirical functions, that is to say, with the representation of empirical data – with common examples such as graphing the outside temperature, a child's height or a company's sales against time. Although we will not go into detail here, we want to note that this encompasses the development of both the notion of variable and the notion of covariation. In relation to this, we point to the fact that the idea of measures as a possible value on a variable does not come naturally. Instead, students tend to see measures as attributes of individuals (Hancock, Kaput & Goldsmith, 1992). Nevertheless, notions of variable and covariation may be further developed when the students start graphing functions that are defined by mathematical prescriptions.

The second learning strand concerns the development of algebraic expressions. Here, the starting point may lay in functions that are described by series of arithmetical operations. Those series of operations may be represented with so-called arrow chains in which each arrow signifies one operation, as in Figure 10.2.

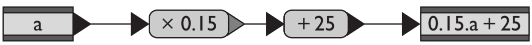


Figure 10.2 Arrow chain.

Working with arrow chains as tools for repeated calculations may constitute the first phase of such a learning route. Early in such a learning route students will be asked to investigate arrow chains that are used for solving contextual problems. Further topics of investigation will be the order of the operations, curtailing arrow chains, looking for input–output patterns and linking arrow chains to algebraic expressions, tables and graphs.

In the second phase, more condensed algebraic expressions will replace the chains of arrows representing individual operations. Then the attention may shift towards exploring the role of parameters and structuring algebraic expressions into smaller units that can be treated as variables in and of themselves. This is what Wenger (1987) calls the global substitution principle, according to which an expression such as

$$2 \cdot v \cdot \surd(u-1) + v \cdot \surd(u+1) = 0$$

for instance, can be transformed into

$$2 \cdot A \cdot v + B \cdot v = 0$$

We assume that such explorations will foster the development of a structural conception. In contrast with Sfard (1991) we do not link the reification of the structural conception of number to one specific moment that is tied to the introduction of an adequate definition. We tend to see reification as a gradual process in which input–output relationships are explored extensively and become part of a network of mathematical relations in connection with which functions may acquire an object-like entity. It is against this background that we want to discuss the partial sequence that involves the computer applet 'AlgebraArrows' in more detail.

Computer applet

The computer tool AlgebraArrows allows for the construction of chains of operations, which are represented by individual arrows that signify 'machines' that each carry out one operation (Boon, 2008). Students can create arrow chains in such a manner that they can be used to execute a given calculational procedure, and, if desired, to repeat the same calculation for a series of input values. The applet has a variety of features to support investigations. One allows for creating tables that show a series of input and output values. Another feature allows for the chains to be extended, linked, compared and compressed. Further, a valuable feature of the computer tool is that the input and output values can be labelled. This enables the students to tie the calculation with the arrow chain explicitly to the contextual problem for which a solution is sought. Moreover, these labels already signify categories, and, as such, refer to values that may vary. Thus when the students start to reason about variables, the names for those variables are already ready to hand. Finally, a graph can be created, with the input value on the horizontal axis and the output value on the vertical axis.

We will illustrate the potential use of the computer applet with a short instructional sequence of eight lessons that was tried out in a teaching experiment. The sequence starts with asking the students to solve contextual problems that require them to specify the series of operations with which the answers can be calculated. Next, arrow chains are introduced as a means for describing such series of operations. Then the AlgebraArrows applet is introduced as a tool that can carry out a series of calculations, which a student may define by putting arrow chains together in the window of the AlgebraArrows applet. This makes it easy for them to produce output values for a series of input values. Then the students are made familiar with the option of producing input and output tables with the AlgebraArrows applet. This enables them to inspect such tables for patterns. In relation to this, students are asked to compare the input and output tables of two arrow chains to solve contextual problems. Here, the students may be asked to look for critical values – such as break-even points – or trends. Subsequently the students are acquainted with the tool option to produce graphs that correspond with such input and output tables. This then allows for tasks that ask for reasoning with graphs as visual representations of sets of ordered number pairs that are linked by some calculational recipe.

In line with the emergent modelling heuristic, we discern a series of sub-models that the students are expected to use as tools. What is aimed for is that the students experience each activity with a newer sub-model or tool as a natural extension of the activity with the earlier sub-model. The idea is that the students will 'see' the earlier activity of working with the former sub-model in their way of working with the new model. We have elaborated this in more detail for this sequence in the first two columns of Table 10.1. The next two columns of the table describe the activity for the students and the mathematical issues that constitute the goals aimed for.

This short sequence will have to be followed by a variety of activities that involve the use of graphs, which is seen as instrumental for developing the notion of a function as an object (Slavit, 1997). Other activities may focus on investigating the arrow chains themselves, e.g., looking at the order of the operations, or curtailing arrow chains. Later on, a feature of the applet may be exploited, that allows for a letter to be inserted in the input box, which results in an algebraic expression in the output box. This opens up a variety of possibilities for exploring the relation between algebraic expressions and arrow chains. We cannot describe a complete instructional program here, but we will limit ourselves to noting that we gradually move into an area where the students may use computer algebra as a tool for further investigations (see, e.g., Kieran & Drijvers, 2006). We would argue, however, that there is a need for tailor-made computer algebra systems for mathematics education that

Table 10.1 Cascade of sub-models.

Tool/sub-model	*History/imagery*	*Activity*	*Mathematical issues*
Written calculation		Finding output values	Calculational recipe & notion of variable
Machine/arrow language	Written calculation	Finding output values	Calculational recipe & notion of variable
Arrow chain in AA-applet	Machine/arrow language	Identifying input & output variables and finding output values	Calculational recipe & notion of variable
Tables in AA-applet	Arrow chain in AA-applet	Investigating dependency relationships	Notion of variable relation between input–output values
Graphs in AA-applet	Tables and arrow chains in AA-applet	Investigating relations between functions	Treating a function as an object (set of ordered number pairs) characteristics of input–output relationships

would, for instance, allow for Wenger's (1987) 'global substitution' in a more straightforward manner.

The teaching experiment was conducted with three grade eight (13–14-year-olds) classes at three different schools. Teaching sessions and group work in two classes were videotaped, and screen-audio videos of three pairs of students working with the computer tool were collected. Students' answers to the computer activities were saved on a central server. In addition, the researchers collected students' written work and the results of a written assessment at the end of the teaching experiment. The data analyses started with organisation, annotation and description of the data in a multimedia data analysis tool. Initially, the tasks in the instructional sequence served as the unit of analysis for clipping the videos. Codes were used to organise and document the data, and to produce conjectures about patterns in the teaching and learning process following the principles of grounded theory (Strauss & Corbin, 1998).

Central in the instructional sequence are mathematical tasks that involve the repeated execution of arithmetical operations. In the first phase of the sequence, the students are expected to find out how the repetition – that is inherent in such tasks – can be organised most effectively. In this manner, functions will come to the fore as recipes for transforming one number into another – which corresponds with the operational conception of functions. The objective of the next phase is to shift the attention of the students, with the help of the computer tool, from the calculations that have to be executed to the relation between input and output values. This shift is facilitated by the computer tool that takes care of the calculations.

In the following, we will give a brief impression of how this instructional sequence actually evolved during the teaching experiment. Activities with the computer applet were preceded with some activities to orient the students on repeatable calculational procedures. It took the students some time to figure out which quantity they could use as the independent variable.[3]

The first activities with the computer concerned contextual problems that the students were asked to solve with the help of the computer applet. To do so, they had to translate the contextual problem into a series of operations that could be represented with an arrow chain. It was found that it was not self-evident for the students in which order the operations had to be sequenced. When designing the arrow chain for the costs of a mobile phone, for instance, most students were inclined to start with the fixed costs. Only when working this out did they realise that this would not work because in this manner the fixed costs would be multiplied with the variable costs. Another complication arose when they had to compose the arrow chain in such a manner that it could be used for a series of calculations with a variety of input values. To make this possible, the arrow chain had to start with the variable, which meant that the students had to identify the independent variable in the contextual problem.

Figure 10.3 Enacting a chain of operations.

This, too, required a small explorative learning process. This process was supported with whole-class activities in which carton board operations were used that each signified an individual operation (or 'machine' that would carry out that operation), which had to be placed in the order of execution that would fit a given task (see Figure 10.3).

An important step in the sequence involved tasks, which required the students to make comparisons between two functions, such as the following.

To get some jobs done in the house we can choose from two companies. Company Pietersen charges €92 starting costs and an hourly rate of €30. Company TweeHoog charges €45 starting costs and an hourly rate of €32.75.

QUESTION **a**: A job takes 9 hours. Which company do you choose?
QUESTION **b**: After how many hours is Pietersen cheaper?

Question **a** asks the students to calculate and compare the output values of two functions for a given value. In question **b** the students are asked to look at a range of values. We present an excerpt of the protocol of Lisa and Romy, after they created the corresponding arrow chains, added tables of input and output values, and inserted '9' in the input box to answer question **a.** For question **b** they enter 'rate' in the input box and ask for the table tool (see Figure 10.4).

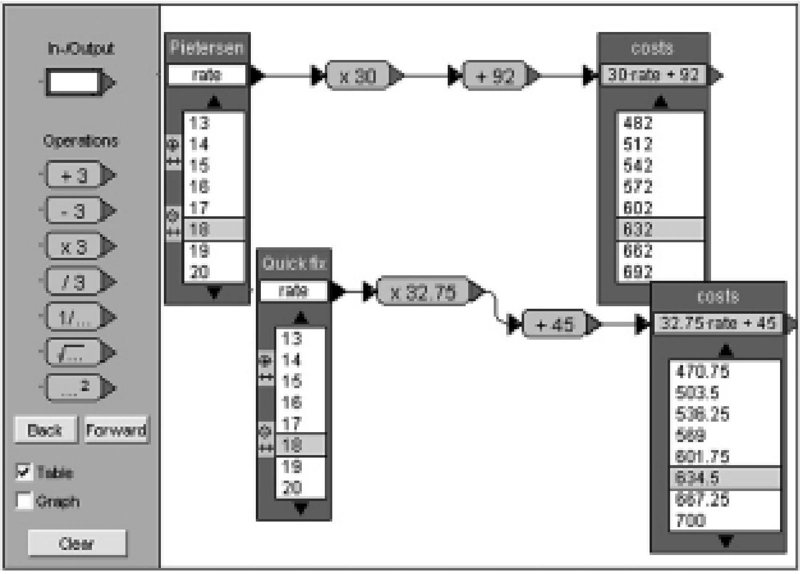


Figure 10.4 ArrowApplet with chains for two contractors.

[They read question **b**, after how many hours is Pietersen cheaper?]
[They start to scroll up in the table of the input values.]
R: No, here he is still more expensive.
L: Oh, yes.
[She starts to scroll down.]
R: I think it is after the nine . . .
L: I should TweeHoog . . . oh yes, of course . . . [she looks again at the output values of 9, and then quickly scrolls down]
R: I look at Pieters . . .
L: Here he is still more expensive [scrolls along 13, 20].
L: Here he is cheaper.
R: . . . cheaper . . .
[They scroll up again.]
L&R: And there, and there too . . . [select 19, 18, 17]
L: No, here he is more expensive.
L&R: From 18 hours. [They return to 18 and type their answer 'from 18 hours' in the answer box of question **b**.]

Note that the arrow chains fulfil two different roles when the students answer questions **a** and **b**. To answer question **a**, the students have to design an arrow chain that fits the calculation that has to be carried out. In question **b**, their attention shifts towards comparing series of input and output values. We believe this to be an important shift. Using Saxe's (2002) terminology, we may speak of a 'form–function shift'. Initially, the arrow chain is constructed as a means to calculate an output value for a given input value, whereas later the arrow chain is taken as a given, when the students investigate the patterns of input and output pairs. It is exactly this kind of shift that we try to foster with the AlgebraArrows applet. A systematic analysis of the work of 155 students involved in a second teaching experiment, with basically the same goal and design as the first one, showed that a substantial number of them made this shift (Doorman et al., submitted).

Given these results, we may argue that the short instructional sequence with the AlgebraArrows applet matched the kind of instruction we aimed for, in that it fostered a dynamic learning process in which symbolisations and meaning co-evolved, and in which the ways that symbols were used and the meanings they came to have were mutually constitutive.

Conclusion

We observed that one of the reasons why it is so difficult to learn mathematics originates in what Sfard (1991) calls the dual nature of mathematics. Mathematical concepts have both a procedural and a structural nature. The history of mathematics shows that new mathematical concepts emerge as procedures first and then gradually transform into objects. Trying to teach mathematics on the basis of the structural conceptions will cause problems if students only have a procedural conception of the topic that is being taught. Such instruction asks of the students to reason in terms of mathematical objects they have not constructed yet. The alternative is to foster a bottom-up process that starts with a procedural conception of a given topic and tries to support the transition to a structural conception of that topic. Nevertheless, learning mathematics in this manner is not easy either, since students eventually have to begin to treat procedures as objects before they have constructed those objects.

We observed that the problems caused by the dual nature of mathematics are conflated with issues that relate to the use of symbolic representations in mathematics. We argued that the so-called learning paradox cannot be overcome by concretising mathematical concepts by means of didactical models. For as long as the students have not developed a structural conception, they will not be able to interpret concrete representations in a structural manner. It showed that the alternative here is, again, a bottom-up approach. Following Meira (1995), one may try to take the historical process in which symbolic representations and meaning co-evolve in a dialectic process, as a source of inspiration for the design of instruction. An instructional design approach that can be seen as an instantiation of this recommendation is the instructional design heuristic of emergent modelling.

We elaborated the emergent modelling approach for instruction on algebraic functions, and we illustrated this with a brief sketch of a teaching experiment that involved the use of a computer applet that enabled the students to represent calculations that consisted of a series of operations in the form of a series of arrows in an arrow chain, in which each arrow signified one operation. Analysing the way the students interacted with the tool, we observed a form–function shift (Saxe, 2002). This form–function shift can be seen as a first step in the development of a structural conception of algebraic functions, and we can argue that the activities with the computer applet supported this first step. In relation to this, we want especially to highlight the role of visual representations in the shape of an arrow chain, which the students first put together as a series of operations for specific calculations, but which gradually took on the role of a machine that produced output values for arbitrary input values and that afforded the construction of the notion of input–output relationships.

Acknowledgements

We thank our colleagues Peter Boon, Sjef van Gisbergen, and Helen Reed, and the teachers and students of the experimental sites, for their collaboration, as well as the Netherlands Organisation for Scientific Research (NWO) for supporting the research study (project no 411-04-123).

Notes

1. We will follow Sfard (1991) in the distinction she makes between a concept (a mathematical idea in its 'official' form) and a conception (the internal, subjective cluster of internal representations and associations).
2. The arithmetic rack is a scaffolding device that allows for representing basic number facts in such a manner, that they can function as means of support for deriving more advanced number facts in addition and subtraction up to 20.
3. Although we use the term 'independent variable' here, this term was not used in the teaching experiment which had the character of an informal exploration.

References

Bereiter, C. (1985). Towards a solution of the learning paradox. *Review of Educational Research, 55*, 201–226.

Boon, P. (2008). AlgebraArrows. Retrieved June 9, 2008, from http://www.fi.uu.nl/wisweb/en/welcome.html.

Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics Education, 23*, 2–33.

Doorman, L. M. (2005). *Modelling motion: From trace graphs to instantaneous change.* Utrecht, The Netherlands: CD-Beta Press.

Doorman, L. M., & Gravemeijer, K. P. E. (2009). Emergent modeling: Discrete graphs to support the understanding of change and velocity. *ZDM International Journal on Mathematics Education, 14*, 199–211.

Doorman, M., Boon, P., Drijvers, P., Van Gisbergen, S., Gravemeijer, K., & Reed, H. (submitted). Tool use and conceptual development: an example of a form-function-shift.

Drijvers, P., Doorman, M., Boon, P., Van Gisbergen, S., & Gravemeijer, K. (2007). Tool use in a technology-rich learning arrangement for the concept of function. In D. Pitta-Pantazi & G. Philipou (Eds.), *Proceedings of the Vth Congress of the European Society for Research in Mathematics Education CERME5* (pp. 1389–1398). Cyprus: Larnaca.

Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. In D. O. Tall (Ed.), *Advanced mathematical thinking* (pp. 95–123). Dordrecht, The Netherlands: Kluwer.

Freudenthal, H. (1991). Revisiting mathematics education. Dordrecht, The Netherlands: Kluwer.

Gravemeijer, K. (1999). How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning, 1*, 155–177.

Gravemeijer, K. (2008). RME theory and mathematics teacher education. In D. Tirosh & T. Wood (Eds.), *International handbook of mathematics teacher education: Vol. 1. Knowledge and beliefs in mathematics teaching and teaching development* (pp. 283–302). Rotterdam, The Netherlands: Sense Publishers.

Gravemeijer, K., Bowers, J., & Stephan, M. (2003). Continuing the design research cycle: A revised measurement and arithmetic sequence. In M. Stephan, J. Bowers, & P. Cobb (with K. Gravemeijer) (Eds.), *Supporting students' development of measuring conceptions: Analyzing students' learning in social context. Journal for Research in Mathematics Education Monograph* No. 12 (pp. 103–122). Reston, VA: National Council of Teachers of Mathematics.

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist, 27*, 337–364.

Janvier C. (Ed.). (1987). *Problems of representation in the teaching and learning of mathematics.* Mahwah, NJ: Erlbaum.

Kieran, C., & Drijvers, P. (2006). The co-emergence of machine techniques, paper-and-pencil techniques, and theoretical reflection: A study of CAS use in secondary school algebra. *International Journal of Computers for Mathematical Learning, 11*, 205–263.

Latour, B. (1990). Drawing things together. In M. Lynch & S. Woolgar (Eds.), *Representations in scientific practice* (pp. 19–68). Cambridge: MIT-press.

Meira, L. (1995). The microevolution of mathematical representations in children's activities. *Cognition and Instruction, 13*, 269–313.

Roth, W.-M., & McGinn, M. K. (1998). Inscriptions: Toward a theory of representing as social practice. *Review of Educational Research, 68*, 35–59.

Saxe, G. B. (2002). Children's developing mathematics in collective practices: A framework for analysis. *Journal of the Learning Sciences, 11*, 275–300.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*, 1–36.

Slavit, D. (1997). An alternate route to the reification of function. *Educational Studies in Mathematics, 33*, 259–281.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory.* Thousand Oaks, CA: SAGE Publications.

Tall, D. (Ed.) (1991). *Advanced mathematical thinking.* Dordrecht, The Netherlands: Kluwer.

Wenger, R. H. (1987). Cognitive science and algebra learning. In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 217–251). Mahwah, NJ: Erlbaum.

Chapter 11

The 'numbers are points on the line' analogy

Does it have an instructional value?

Xenia Vamvakoussi

Problems of conceptual change in the development of the rational number concept

In the past few years, we have conducted a series of studies investigating the development of the number concept from a conceptual change perspective (Vamvakoussi & Vosniadou, 2004, 2007, 2010). A key assumption of our theoretical framework is that students, before they are exposed to instruction about non-natural numbers, have already formed complex and relative coherent explanatory frameworks for numbers; these are tied around their knowledge and experience with natural numbers, which are shaped initially in the context of their socio-cultural environment and further confirmed and strengthened during the first years of instruction which focuses on natural number arithmetic. Among the background assumptions underlying students' initial number concept is the idea that numbers are discrete, i.e. they obey the successor principle (see also Gelman, 2000; Smith, Solomon, & Carey, 2005). It is amply documented that the idea of discreteness constrains students' understanding of the density property of rational and real numbers not only at elementary and secondary education (Hartnett & Gelman, 1998; Merenluoto & Lehtinen, 2002), but at tertiary education level as well (Giannakoulias, Souyoul & Zahariades, 2007; Tirosh et al., 1999).

However, it is not the only constraint on students' understanding of the density of numbers. Consider the following examples, coming from individual interviews with ninth-graders, who were asked how many numbers there are between 3/8 and 5/8 (Vamvakoussi & Vosniadou, 2004).

> Example 1. '*There is one number*, 4/8... *Just a moment! There is also 4.0/8, isn't it? And* $\sqrt{16}/4$, *right? There are many, many numbers in between*.'

> Example 2. '*There is no other number, because if you simplify 4/8 you get 1/2. And this is not in between*.'

Then consider the example of a ninth-grader who answered that there are nine numbers between 0.001 and 0.01, but stated without hesitation that there are infinitely many numbers between 3/8 and 5/8. Following a prompt by the interviewer, he explained:

> Example 3. '*Between 0.001 and 0.01 there are nine numbers. Or maybe ten – I'm not so sure about that. But if you convert them to fractions, you can find more numbers in between. You can find infinitely many numbers.*'

These students have been exposed to intensive instruction about decimals and fractions, including ordering, operations and conversions, starting from the third grade and throughout the elementary school. They have been introduced to the term 'rational' and 'real' numbers in the eighth grade, and have been using variables with real values ever since. However, the above excerpts show that when they are faced with a relatively unfamiliar task their judgements are constrained by natural number knowledge interference, which leads them to refer to a finite number of intermediate numbers; in addition, there is also a clear issue of interpretation of rational number notation. All three students seem to treat different symbolic representations of the same numbers as if they were different numbers (for similar results in the context of ordering see Markovits & Sowder, 1991). This tendency was further investigated in subsequent studies (Vamvakoussi & Vosniadou, 2007, 2010) and the results corroborated the above: students were reluctant to accept that there can be decimals between fractions and vice versa, even when they answered that there are infinitely many intermediate numbers; they also gave different patterns of responses for different symbolic representations (e.g., infinitely many intermediates between decimals, but a finite number between fractions). As Kilpatrick, Swafford, and Findell (2001) also note, it seems that students deem rational numbers a set of different, unrelated 'sets' of numbers (integers, decimals, fractions).

In treating the different symbolic representations as if they were different numbers, students seem to confuse between the representations and the mathematical object (Duval, 2006). In a pure or advanced mathematical context, a rational or real number can be viewed as an abstract, not tangible entity, which takes its meaning within a formal system, i.e. as a mathematical object. This perspective is, typically, not fostered in school contexts, where students struggle to construct meaning for decimals and fractions (Dörfler, 1995). Besides violating the expectations based on their initial number concept, these new constructs are presented in different contexts and accompanied by a large number of different concrete, graphical and symbolic representations. For example, the number 'one half' can be represented graphically in ways that emphasise the part-whole meaning of the fraction – like the pie – or its measure aspect, like the number line; depending on the context, it can be represented symbolically as a percentage (e.g., '50 per cent discount'), in the form a/b, or in various decimal forms (e.g., €0.50 or 0.500 kg). It can also be related to a collection of discrete objects, via the 'one out of two' interpretation of its meaning. For these different meanings and representations to be coordinated in an integrated whole, students have to realise that they are all connected to the same quantitative relation – or, at a more abstract level, the same (rational) number. This requires conceptual restructuring within the initial number concept, involving representational changes (see also Ni & Zhou, 2005; Smith et al., 2005). Given that the initial number concept does not support the development of the rational number concept, what are the learning mechanisms via which efficient learning can be accomplished?

Cross-domain mapping: a mechanism to foster conceptual restructuring

Analogical reasoning, and, in particular, cross-domain mapping, is considered an important mechanism for conceptual restructuring (Vosniadou, 1989). Gentner et al. (1997) describe the comparison between two domains (usually termed base and target domain) as a process that may highlight their common features and reveal unnoticed commonalities. They suggest that the projection of inferences adds to the knowledge of the target domain and that in the process, re-presentation of the target or of both domains may occur, to improve the match. They also propose that the use of analogies may lead to conceptual restructuring, i.e. a large-scale rearrangement of the target domain. They present examples drawing on detailed analysis of the use of analogies by scientists, in the context of discovery (see also Nersessian, 1992). Carey and her colleagues (Carey, 2004; Smith et al., 2005) view cross-domain mapping as a 'bootstrapping process' which supports learning when what is to be learned transcends what is already known not merely in a quantitative way, but also in some qualitative way. They focus on the rational number concept, arguing that the bootstrapping process through which this concept is created involves modelling numbers in terms of representations of physical quantity, i.e. a cross-domain mapping between matter and number. Their findings suggest that there is a close relation between elementary school children's understanding of matter and of number as infinitely divisible, the first slightly preceding the second. Smith et al. also refer to the work of Moss and Case (1999) as a promising example of a bootstrapping curriculum to foster the development of the rational number concept. The intent of this curriculum was to help fourth-graders to coordinate between an assumed global representation of proportions related to physical quantity and a numerical structure supporting splitting and halving. Students began representing, via percentages, the notions of full, nearly full, half-full and nearly empty, as these applied to a beaker of water, for instance. Then they were led to associate their intuitive understanding with numerical halving strategies, before they were introduced to decimal and fractional notation. Smith et al. argue that the mapping between number and physical quantity might be instrumental in children's appreciating the existence of rational numbers, e.g., between 0 and 1, and that any rational number is repeatedly divisible.

Taking a step forward in terms of the age of students, as well as the level of abstraction of the intended mathematical meanings, one could pose the following question: what would be an appropriate cross-domain mapping that would help secondary school students to (a) conceptualise numbers as individual entities, invariant under different forms of representation, and (b) re-represent them as a dense array, instead of sequences of numbers that preserve locally the ordering of natural numbers?

I will argue that the comparison between number and the geometric line entails a cross-domain mapping which is worth investigating in this respect. There are two reasons to support this claim: first, exploring the relations between the two domains has been instrumental in the historical development of both mathematical concepts (i.e. number and the line). Second, a product of this long interplay between number and the line, namely the number line, is commonly used in school settings to model the real numbers.

Number and the line: a case of long-term cross-domain mapping

Number and the line began as two inherently different objects of study and ended up as two objects between which, in certain mathematical contexts, there is no useful distinction (see, e.g., Schechter, 2006). In ancient Greek mathematics, the line was deemed inherently continuous, whereas number was viewed as a collection of units, discrete in nature. Moreover, the unit was viewed as a generator of numbers, but not a number itself (Klein, 1968/1992). The exploration of relations between number and the line was initiated by a theoretical question regarding measurement: can *any* length be captured by a number? The answer to this question depends crucially on what counts as a *number*. Within ancient Greek mathematics, ratios of commensurable and incommensurable quantities were not deemed *numbers*, thus the possibility of a correspondence between lengths and numbers was not thinkable.

The concept of number as a collection of units proved extremely robust for many centuries to follow; indeed, although ratios (rational as well as irrational) were used extensively for calculation purposes, it proved very difficult to conceptualise them as individual entities, let alone as members of the same family as the natural numbers. It is quite revealing that Stevin, the sixteenth-century mathematician who borrowed and adjusted the Arabic digital and positional system, felt obligated to state explicitly in his *Arithmétique* that (a) the unit is a number, (b) the fractional parts of the unit are also numbers, and (c) every 'root' is a number (Klein, 1968/1992). There emerges the possibility of assigning to number, properties that have been attributed before only to continuous quantities, precisely in attempts to describe continuity. These include the notion of infinite divisibility (already expressed by Aristotle) and the possibility to always find a third point between any two that implies the infinity of intermediate points, which was stated explicitly by Occham, a fourteenth-century philosopher

(Bell, 2005). From the point of view of today's mathematics, infinite divisibility, as well as the infinity of intermediate points, relates to the notion of density and is not sufficient to describe continuity. A clear differentiation between density and continuity was made possible by comparison with the domain of number, as late as the nineteenth century, through the attempts to describe continuity in terms of number (Boyer, 1959).

However, the fact that neither continuity nor the real numbers were adequately defined did not prevent the development of mathematical domains and tools that made use of both. In particular, intuitions of geometric continuity and assumptions as to the relation between geometrical magnitude and number were at the heart of calculus and analytic geometry.

This complex interplay between number and the line reached a crucial point with the foundation of real numbers by Dedekind and Cantor in the nineteenth century. It is interesting to see in Dedekind's own words (quoted in Dantzig, 1930/2005) how he refers to the cross-domain mapping that sustained his conceptualisation of continuity, which led him to define the real numbers so as to have this property:

> The comparison of the domain of rational numbers with a straight line has led to the recognition of the existence of gaps, of a certain incompleteness or discontinuity, in the former; while we ascribe to the straight line completeness, absence of gaps, or continuity. (p. 177)

Rather than lengths, numbers are now viewed as points on the line; the fact that the points corresponding to rational numbers do not exhaust the totality of the points is acknowledged (thus density is clearly differentiated from continuity).

However, the question remained: how can one be sure that irrational numbers indeed fill all the remaining 'gaps', i.e. that there is indeed a one-to-one correspondence between the real numbers and the points of the line? It turns out that the one-to-one correspondence between points and real numbers had to be postulated by what is now known as the Cantor–Dedekind axiom.

The school number line: affordances and limitations

The number line is a representation of the real numbers commonly used in school settings, grounded on the analogy 'numbers are points on the line'. As such, it calls for a reconceptualisation for numbers, which, arguably, might help students conceive of rational numbers as individual entities, and also facilitate their understanding that, for instance, 0.5 and 1/2 are interchangeable representations of the same number, rather than different numbers, since they correspond to the same point. It may also facilitate conceptualising infinite decimals, and in particular irrational numbers, as entities rather than unending processes (Sirotic & Zazkis, 2007). This could promote students' understanding of rational (and real) numbers as a unified number system and facilitate their understanding of their structure in this respect (Kilpatrick et al., 2001).

In addition, the geometrical features of the number line could be used to confront students' belief that numbers are discrete in nature. More specifically, it would be worth investigating whether an understanding of density could be fostered in a geometric context (i.e. as density of points) and then be transferred to an arithmetic context via the 'numbers are points' analogy. Besides preceding the emergence of the notion of the density of numbers in the course of historical development, density of points could be more accessible to students, due to the fact that it does not involve any symbolic notation, which proved to be an additional burden for students in our previous studies.

This should not be taken to imply that the 'numbers are points' analogy is deemed transparent to students, nor that the conceptualisation of the segment as a dense array of points is immediately triggered by the intuitively perceived continuity of the line. This point has been put forward by Núñez and Lakoff (1998), who make a sharp distinction between the notion of the 'holistic line', which can be conceptualised as the trace of a moving object (e.g., the pencil, when it does not leave the paper) and is continuous in an everyday sense, and the notion of the line as a set of points, which they characterise as a mathematically elaborated metaphor of the line. In fact, describing the line as a set of points is compatible with the idea of a series of points laying the one immediately next to the other.

There is yet another distinction that has to be taken into consideration, namely the distinction between abstract, geometrical objects and their physical representations. There is the possibility that students assign to the line, as well as to the points, properties such as width, or length, which do not apply in the case of their idealised counterparts. Conceptions of points as material spots are at odds with the infinite number of points on a line segment. Such conceptions may underlie the conception of a segment as 'a necklace of beads' (Sbaragli, 2006) and the belief that longer segments have more points (Fischbein, 1987, pp. 138–139). In addition, it has been shown that conceiving of actual infinity, as well as infinite divisibility, in a geometrical context may be constrained by pragmatic considerations (e.g., Tirosh & Stavy, 1996). On the other hand, Smith et al.'s (2005) findings suggest that children conceive matter as infinitely divisible slightly before assigning this property to number.

Students' early experiences with the number line at elementary school, when the number line represents the natural numbers, may also have an adverse effect. Quoting Dufour-Janvier, Bednarz, and Belanger (1987), English (1993) points out that students tend to see the number line as a series of 'stepping stones' with an empty space in between, commenting that this could explain why so many secondary students say that there are no numbers, or at the most one, between two whole numbers. Another common metaphor for the number line in school settings, namely that of the ruler (e.g., Doritou & Gray, 2007), may also convey the idea that the amount of intermediate numbers is finite.

Investigating the effect of the number line on students' judgements about the number of numbers in an interval

We asked two ninth-graders who had responded that there are no other numbers between two given decimals, to place them on the number line (unreported data from the Vamvakoussi & Vosniadou (2004) study). Both students drew a line and placed the two decimals, allowing for a small, yet obvious distance between them. The question was again posed: are there any numbers between the two given ones?

STUDENT A: 'No, there are no numbers in between. There is space, but there are no numbers.'

STUDENT B: 'There are . . . I don't remember how they're called . . . Thousandths? You know these little lines . . . Like there are on the ruler.'

In Vamvakoussi and Vosniadou (2007), we tested for possible effects of the presence and the order of appearance of the number line on students' judgements on the number of numbers in an interval. We used a questionnaire consisting of two parts. Both parts consisted of items asking how many numbers there are between two rational numbers a, b. In one of the parts, the numbers a and b were demonstrated on a number line. All participants received both parts, but half of them received the items with the number line first, whereas the other half received first the items without the number line. In both cases, the first part was withdrawn before the second one was administered. We found no significant effects of the presence and the order of appearance of the number line on students' performance. Sometimes the number line helped students to move from a 'no other number answer' to a 'there is a finite number of numbers' one (like Student B), but only rarely did they move to an 'infinitely many numbers' answer. Moreover, the presence of the number line sometimes had an adverse effect, leading students to move from the infinite to the finite side.

In order to investigate this phenomenon further, we designed a questionnaire consisting of 16 multiple-choice items. The first 15 were three item blocks, targeting (a) the number of numbers in an interval, (b) the number of numbers between two given numbers placed on the number line, and (c) the number of points on a segment. The 16th item was posed as a thought experiment. The scenario was that a mathematics teacher asked her students to imagine that they had the possibility of unlimited magnification of a line segment AB. How would they 'see' with their mind's eye the point A and its neighbouring points? The segment before the magnification was presented in a figure and then our participants were offered three choices, in the form of drawings produced by these imaginary students: in the first the magnified segment had grown in length and also in width, in the second it was presented as a 'necklace of beads', whereas in the third it remained continuous, increasing in length, but not in width. We administered this questionnaire to 229 secondary school students, from seventh to eleventh grade (Vamvakoussi & Chatzimanolis, 2008). Similar to the Vamvakoussi and Vosniadou (2007) study, the results showed no significant difference in students' performance with or without the number line. On the other hand, all age groups performed significantly better in the items of the third block relating to points, that is, students were more apt to accept that there are infinitely many points on a segment, than numbers in an interval, either with or without the number line. However, this finding should not be taken to imply that students showed a firm understanding of the infinity of points in an interval. On the contrary, they were susceptible to variations in the length of the segments, i.e. they moved from an 'infinite' to a 'finite' answer for shorter segments. For example, 58 students (25.33 per cent in the sample) answered that there are infinitely many points in the case of the longest segment presented in the questionnaire, but 25 out of them answered that there is a finite number of points in the case of the shortest segment. Only 24 students (10.5 per cent in the sample) were found to answer consistently that there are infinitely many intermediate points in all five items. When we examined these students' responses to the thought experiment, we found that 12, i.e. half of them, described the segment as consisting of discrete points. Eight of these students (33.3 per cent) described the segment as a physical object that grows in width as it gets magnified.

These findings suggested that the geometrical features of the line could be a good starting point for introducing the idea of density, since students were found on the 'infinite' side more often in the case of points than in the case of numbers. Obviously, this does not imply that they have a deep understanding of the infinity of points. The fact that they were susceptible to variations of the length of the presented segments, suggests that 'infinitely many points' may be interpreted by students as 'a very large amount of very small spots', which is compatible with the conceptualisation of points as material spots (Fischbein, 1987). What is interesting is that, whatever understanding students have of the number of points on a segment, it does not transfer to the domain of numbers, even when these are presented on the number line. Revisiting the example of Student A, it seems that either the presence of intermediate points on the segment is not acknowledged if not specifically asked for, or the correspondence between numbers and points is not evoked. The first could be explained assuming, as Núñez and Lakoff (1998) do, that from the students' point of view, the 'holistic' line has primacy over the points, which are called into existence only as 'locations' on the line.

Finally, our findings indicated that the 'infinity of points' on a segment might not necessarily imply their 'denseness', from the students' point of view. Indeed, students who were consistently on the ‘infinite side’ with respect to the number of points on a segment, regardless of its length, still described the segment as an array of points lying the one next to the other, given the possibility of unlimited magnification. This could again be attributed to a misinterpretation of the term ‘infinity’ to mean a very large amount. But there is also the possibility that these students believe that the points are infinitely many, yet discrete. Data obtained by a questionnaire cannot provide enough information to support this claim. However, we had the opportunity to investigate this assumption further in a microgenetic study, where we provided tenth- graders with information about the infinity of numbers in a specific interval and also about the numbers-to-points correspondence, and studied how they employed this information in solving density-related tasks (Vamvakoussi, Christou, & Vosniadou, in preparation). Consider the following excerpt from the individual interview of student C, who is dealing with two tasks. The first task asks if it is possible to know how many numbers there are between a, b $\in \Re$, a < b, without knowing which these numbers are. The second task asks whether it is possible to specify the first (smallest) value that the variable a $\in \Re$ takes, given that a > 10.

The student answers that it doesn’t matter whether or not we know which numbers a and b stand for, since there are infinitely many intermediates. She explains that ‘*if we place a and b on the number line, no matter where, this segment, big or small, always consists of infinitely many numbers*’. However, when it comes to the second task, she says:

STUDENT C: I would say 11, but there are more numbers between 10 and 11: 10.5, 10.1, 10.01, 10.001 and so on.
RESEARCHER: So what do you think, is there such a number?
STUDENT C: There is, but I do not know precisely which. It is what we were talking about before, there are infinitely many numbers in between, and so we may not be able to define the one immediately after 10.
RESEARCHER: So, let me draw a line and place 10 here. You say that 11 is not immediately after 10, since there are infinitely many others in between. But are you saying that there is *some* number immediately next to 10?
STUDENT C: Yes, it is the one that is immediately after 10, the immediately next point, the successive point.

This student refers to the infinity of numbers in an interval, as well as to the infinity of points on a segment. In fact, she uses the numbers-to-points correspondence to explain that there are infinitely many numbers in any interval. She also employs a recursive process in order to present examples of decimal numbers between 10 and 11, by adding more decimal digits. However, she seems convinced that the successor of 10 exists *in principle*. Moreover, she again draws on the number-to-points correspondence and refers specifically to the immediately next point. For this student, ‘infinitely many’ means something more than a ‘large amount’ – it presumably means an unending amount of numbers (or points), in the sense that there is always one more to be found. However, this process does not necessarily produce a dense array of numbers or points that can never be found the one immediately next to the other. This interpretation relates to the distinction between *potential* or *dynamic* and *actual* infinity. To illustrate the difference between these two aspects of infinity, one can compare the following sentences: (a) 'there are infinitely many natural numbers', (b) 'the cardinality of the set of natural numbers is aleph null'.[1] One can conclude the first by drawing on the possibility of always finding one more natural number by adding one unit, a process that can repeat ad infinitum. On the other hand, the second presupposes that the totality of natural numbers is conceptualised as a single entity. Both conceptual analysis and research-based evidence show that the notion of potential infinity is far more accessible, even to young children (e.g., Hartnett & Gelman, 1998; Singer & Voica, 2008), than the notion of actual infinity (e.g., Fischbein, 1987; Lakoff & Núñez, 2000; Tsamir & Tirosh, 1999). In line with this consideration, it could be argued that it might be easier to infer the infinity of number in an interval, than to conceptualise the interval as a dense array on numbers, i.e. that 'infinity' does not necessarily imply 'denseness' from the point of view of the learner.

To summarise, the findings of the Vamvakoussi and Chatzimanolis (2008) study, in line with findings from previous research, indicated that (a) students confuse between geometrical objects and their physical representations, (b) students may conceptualise points as 'locations' on the line that are present only when the location is indicated, and (c) the conception of the segment as an aggregate of points might come down to thinking of it as a 'necklace of beads'. In addition, these results showed that students were more apt to accept the infinity of points on a segment, than the infinity of numbers in an interval, indicating that the numbers-to-points correspondence was not employed, even in the presence of the number line. Finally, there were indications that two aspects of density, namely the 'infinity of intermediates' and the 'no successor' aspect, may not be equivalent from the learner's point of view.

Bridging the gap between discreteness and density with the 'rubber line'

It is probably evident by now that we are dealing with a situation where the idea of density is not adequately developed, neither in the domain of points nor in the domain of numbers. Nevertheless, there was evidence suggesting that the idea of density might be more accessible to students in a geometric than in an arithmetic context. The question arises: how can the gap between the segment as conceived by students and the segment as a dense array of points be bridged? A promising approach is the 'bridging analogy' one, developed by Clement and his colleagues (Brown & Clement, 1989; Clement, 1993), with the purpose of dealing with the problem of the gap that often exists between students' initial understanding of a situation and the intended scientific idea. They propose that this gap can be bridged via the interpolation of one or more intermediate anchoring situations which are accessible to students' current understanding and have the potential to trigger a correct intuition, i.e. one that can be developed toward understanding the target situation.

We employed the 'rubber line' as anchor, presenting the line as an imaginary rubber band that never breaks, no matter how much it may be stretched. The idea of an elastic line that one can stretch or shrink is not new. It has been employed, for example, to model multiplication and division on the number line (e.g., Kilpatrick et al., 2001). In the present case it serves as an intermediate anchoring situation between students' initial understandings of the line segment and the intended idea of the segment as a dense array of points: on the one hand, it is grounded on students' experience with a real-world object; on the other, it is grounded on a recursive process, which is likely to be supportive for students, as discussed in the previous section. The 'rubber line' aims at conveying the idea that no matter how close two points seem to be, there is always a line segment (thus, more points) to be found in between, by stretching the rubber line.

In order to test the added value of this approach, we designed a short, text-based intervention (Vamvakoussi, Katsigiannis, & Vosniadou, 2009). We assumed that if we provided students with explicit information about the infinity of numbers in an interval, they would improve their performance in similar tasks; however, such information would not be sufficient for them to deal with the 'no successor' aspect of density. We hypothesised that students exposed to the 'rubber line' analogy would perform better in items related to the 'no successor' aspect of density.

One could ask why we chose to base this short intervention solely on the use of texts, rather than creating a richer learning environment, allowing for interaction among the students, interventions by the teacher, and possibly for the use of manipulatives or virtual manipulatives. Of course, this would be a much better choice from an instructional design perspective. However, it would be very hard to tell which features of the environment could account for the possible learning gains. On the other hand, Greek students at the secondary level spend much of their time trying to deal with tasks on the basis of explicit information provided either by the teacher or by texts; dynamic representations and manipulatives are, typically, not used in mathematics classrooms at this level. Thus, this minimal design was close to their everyday classroom experiences.

We constructed three expository texts. All three texts had a common part referring to the number of numbers in the interval defined by 0 and 1, which provided the correct answer ('*there are infinitely many numbers between 0 and 1*'), differentiating between 'infinitely many' and 'a very large amount'. This common part also reminded students of the one-to-one correspondence between the infinitely many points of the line and the real numbers. The first text continued by evoking the notion of space between 0 and 1 on the number line and presented several examples of finite as well as infinite decimals lying in the interval. This last paragraph was varied in the second text with the insertion of two figures illustrating the interval and the given examples of intermediate numbers. Finally, the third text (thereafter, 'rubber line text') was identical to the first and included, in addition, a paragraph introducing the rubber line as a way to explain how it is possible for infinitely many numbers to lie on the line segment.

These texts were administered to three classes of eighth-graders (a total of 66 students) and three classes of tenth-graders (a total of 74 students), all in the same school in the suburbs of Athens. The students were pre-tested with a multiple-choice questionnaire, which, similar to the questionnaire used in the Vamvakoussi and Chatzimanolis (2008) study, asked about the number of numbers in an interval, with or without the number line, and the infinity of points on a segment. Then they received a text (students in the same class received the same type of text). Finally, they were administered the post-test, which consisted of two parts: the first was identical to the pre-test and the second included five items, asking students to evaluate a statement about the existence of two successive numbers or points, and to justify their answer. The procedure lasted about 45 minutes.

No significant performance differences were found between the students assigned to the different text-type conditions in the pre-test, nor in the first part of the post-test (within grade). As expected, students under all conditions improved their performance significantly after the intervention; in fact, there was a considerable shift from 'a finite number of intermediates' to 'infinitely many intermediates' answers after the intervention, under all conditions.

The added value of the 'rubber-line' text was manifested in students' responses to the additional items of the post-test. The students assigned to this condition outperformed the students in the other two conditions, whereas no significant difference was found between the other two texts. The rubber-line students were far more consistent in denying the possibility of any of the given pairs of numbers being successive; and similarly for points, even given the possibility of unlimited magnification. Figure 11.1 presents the percentage of students who managed to answer correctly all five additional items of the post-test, per grade. The students who were not exposed to the rubber-line text are grouped together and the percentages are calculated within each group (rubber line/ other), per grade. The great majority of the students who were not exposed to the rubber-line text were found at least once to deem two numbers or points successively. The reverse holds for the rubber-line students, even eighth-graders.

The findings of this study showed that the 'rubber line' helped students to deal with the 'no successor' aspect of density, which seems to be more difficult than the 'infinitely many intermediates' one. This is an encouraging result, given that, besides being short, the intervention was rather conservative in the sense that it was text-based and did not allow for any interaction in the classroom.

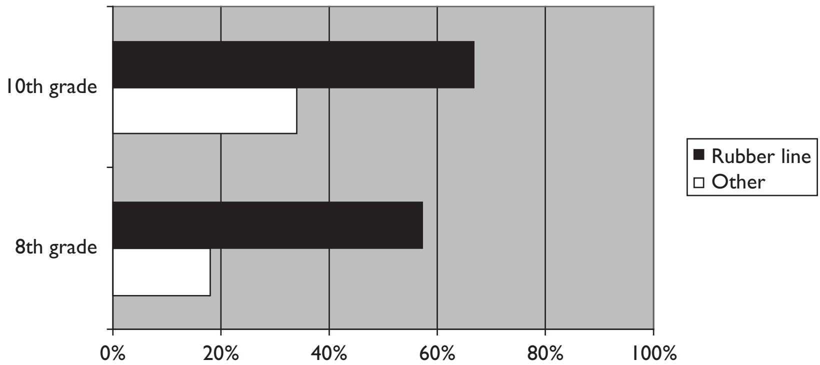


Figure 11.1 Percentage of students who responded correctly to all five items of the post-test, in the ‘rubber-line’ condition and in the other two conditions combined.

Concluding remarks

In this chapter I have presented an overview of studies that began with the investigation of secondary students’ understanding of density with respect to ordering, and concluded with the attempt to bring this concept within the grasp of students. Students’ difficulties with the density of numbers revealed the interference of natural number knowledge and the difficulty with the interpretation of rational number notation, both of which are closely related to the problem of conceptual change in the development of the rational number concept (see also Ni & Zhou, 2005; Smith et al., 2005). Cross-domain mappings could be instrumental in facilitating students’ attempts to construct meaning for rational number notation and perceive properties and relations that were not ‘visible’ before – and could not be through the lenses of their current number concept. Based on historical considerations and also some empirical evidence, we chose to invest on the cross-domain mapping between number and the line. Unlike the association between physical quantity and number, proposed by Smith et al. (2005) and a number of other researchers such as Moss and Case (1999) as instrumental in younger children’s understanding of the rational numbers, in this case both domains are abstract, and students’ intuitions of the denseness of points in a segment had to be built in what Clement (1993) would call an ‘anchoring situation’, namely the ‘rubber line’. This approach produced some encouraging results.

It should be stressed that an intervention, such as the one described in the previous section, is not deemed sufficient to bring about conceptual changes in students’ number concept. Rather, what is argued is that purposeful, long-term use in instruction of the number line, and in particular of the numbers-points correspondence, may be valuable – provided that it is accompanied by adequate explanations and representations that help students to develop and coordinate their understandings in the domain of number and geometrical magnitude, and bridge the gap between students' initial conceptions and the intended mathematical meanings.

Acknowledgement

This work is supported by the project ANALOGY: Humans – The Analogy-Making Species, financed by the FP6 NEST Programme of the European Commission (STREP Contr. 029088).

Note

1. The aleph numbers are a sequence of numbers used in set theory to denote the cardinality of infinite sets. Aleph null denotes the cardinality of the set of natural numbers and is the first in the sequence of the aleph numbers.

References

Bell, J. L. (2005). Continuity and infinitesimals. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy.* Retrieved October, 2008, from http://plato.stanford.edu/archives/win2008/entries/continuity.

Boyer, C. B. (1959). *The history of calculus and its conceptual development.* New York: Dover Publications.

Brown, D. E., & Clement, J. (1989). Overcoming misconceptions via analogical reasoning: Abstract transfer versus explanatory model construction. *Instructional Science, 18*, 237–261.

Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus, 133*, 59–68.

Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching, 30*, 1241–1257.

Dantzig, T. (2005). *Number: The language of science*. New York: Pi Press. (Original work published 1930).

Dörfler, W. (1995). Mathematical objects, representations and imagery. In R. Sutherland & J. Mason (Eds.), *Exploiting mental imagery with computers in mathematics education* (pp. 82–94). NATO ASI Series F: Advanced Educational Technology. Berlin, Germany: Springer-Verlag.

Doritou, M., & Gray, E. (2007). The number line as metaphor of the number system: A case study of a primary school. In D. Pitta-Pantazi & G. Philippou (Eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education* (pp. 111–120). Cyprus: ERME, Department of Education, University of Cyprus.

Dufour-Janvier, B., Bednarz, N., & Belanger, M. (1987). Pedagogical considerations concerning the problem of representation. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 159–195). Hillsdale, NJ: Erlbaum.

Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics, 61*, 103–131.

English, L. (1993). Reasoning by analogy in constructing mathematical ideas. In G. Bell, B. Wright, N. Leeson, & J. Geake (Eds.), *Challenges in mathematics education: Constraints on construction* (pp. 213–222). Lismore, Australia: Mathematics Education Group of Australasia.

Fischbein, E. (1987). *Intuition in science and mathematics.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Gelman, R. (2000). The epigenesis of mathematical thinking. *Journal of Applied Developmental Psychology, 21*, 27–37.

Gentner, D., Brem, S., Ferguson, R. W., Markman, A. B., Levidow, B. B., Wolff, P., & Forbus, K. D. (1997). Analogical reasoning and conceptual change: A case study of Johannes Kepler. *Journal of the Learning Sciences, 6*, 3–40.

Giannakoulias, E., Souyoul, A., & Zachariades, T. (2007). Students' thinking about fundamental real numbers properties. In D. Pitta-Pantazi & G. Philippou (Eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education* (pp. 416–425). Cyprus: ERME, Department of Education, University of Cyprus.

Hartnett, P. M., & Gelman, R. (1998). Early understandings of number: Paths or barriers to the construction of new understandings? *Learning and Instruction, 8*, 341–374.

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics.* Washington, DC: National Academy Press.

Klein, J. (1992). *Greek mathematical thought and the origin of algebra.* New York: Dover Publications. (Original work published 1968).

Lakoff, G., & Núñez, R. (2000). *Where mathematics comes from. How the embodied mind brings mathematics into being.* New York: Basic Books.

Markovits, Z., & Sowder, J. (1991). Students' understanding of the relationship between fractions and decimals. *Focus on Learning Problems in Mathematics, 13*(1), 3–11.

Merenluoto, K., & Lehtinen, E. (2002). Conceptual change in mathematics: Understanding the real numbers. In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 233–258). Dordrecht, The Netherlands: Kluwer.

Moss, J., & Case, R. (1999). Developing children's understanding of rational numbers: A new model and an experimental curriculum. *Journal for Research in Mathematics and Education, 30*, 122–147.

Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. In R. N. Giere (Ed.), *Cognitive models of science* (pp. 3–45). Minneapolis, MN: University of Minnesota Press.

Ni, Y., & Zhou, Y.-D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist, 40*, 27–52.

Núñez, R., & Lakoff, G. (1998). What did Weirstrass really define? The cognitive structure of natural and ε-δ continuity. *Mathematical Cognition, 4*, 85–101.

Sbaragli, S. (2006). Primary school teachers' beliefs and change of beliefs on mathematical infinity. *Mediterranean Journal for Research in Mathematics Education, 5*(2), 49–76.

Schechter, E. (2006). What are the real numbers, really? Retrieved February, 2008, from http://www.math.vanderbilt.edu/~schectex/courses/thereals/.

Singer, F. M., & Voica, C. (2008). Between perception and intuition: Learning about infinity. *Journal of Mathematical Behavior, 27*, 188–205.

Sirotic, N., & Zazkis, R. (2007). Irrational numbers on the number line – where are they? *International Journal of Mathematical Education in Science and Technology, 38*, 477–488.

Smith, C. L., Solomon, G. E. A., & Carey, S. (2005). Never getting to zero: Elementary school students' understanding of the infinite divisibility of number and matter. *Cognitive Psychology, 51*, 101–140.

Tirosh, D., Fischbein, E., Graeber, A. O., & Wilson, J. W. (1999). *Prospective elementary teachers' conceptions of rational numbers.* Retrieved May 5, 2005, from http://jwilson.coe.uga.edu/Texts.Folder/Tirosh/Pros.El.Tchrs.html.

Tirosh, D., & Stavy, R. (1996). Intuitive rules in science and mathematics: The case of "everything can be divided by two". *International Journal of Science Education, 18*, 669–683.

Tsamir, P., & Tirosh, D. (1999). Consistency and representations: The case of actual infinity. *Journal for Research in Mathematics Education, 30*, 213–219.

Vamvakoussi, X., & Chatzimanolis, N. (2008, August). *Are there as many numbers as points on a line? The effect of the number line on students' thinking about density.* Poster presented at the 6th International Conference on Conceptual Change, Turku: Finland.

Vamvakoussi, X., Christou, K., & Vosniadou, S. (in preparation). A bridging analogy to foster secondary students' understanding of the density property of rational numbers: Evidence from a microgenetic study.

Vamvakoussi, X., Katsigiannis, K., & Vosniadou, S. (2009). Bridging the gap between discreteness and density. In M. Tzekaki, M. Kaldrimidou, & C. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 5, pp. 273–280). Thessaloniki, Greece: PME.

Vamvakoussi, X., & Vosniadou, S. (2004). Understanding the structure of the set of rational numbers: A conceptual change approach. *Learning and Instruction, 14*, 453–467.

Vamvakoussi, X., & Vosniadou, S. (2007). How many numbers are there in an interval? Presuppositions, synthetic models and the effect of the number line. In S. Vosniadou, A. Baltas, & X. Vamvakoussi (Eds.), *Reframing the conceptual change approach in learning and instruction* (pp. 267–283). Oxford, UK: Elsevier.

Vamvakoussi, X., & Vosniadou, S. (2010). How many decimals are there between two fractions? Aspects of secondary school students' understanding of rational numbers and their notation. *Cognition and Instruction, 28*, 181–209.

Vosniadou, S. (1989). Analogical reasoning as a mechanism in knowledge acquisition: A developmental perspective. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 413–436). Cambridge, MA: Cambridge University Press.

Chapter 12

Use of external representations in science

Prompting and reinforcing prior knowledge activation

Sandra A. J. Wetzels, Liesbeth Kester, and Jeroen J. G. van Merriënboer

Introduction

Prior knowledge activation has strong facilitative effects on learning. De Grave, Schmidt, and Boshuizen (2001), for example, prompted students to activate their prior knowledge by means of problem-based discussion. Before studying a text that described the process of blood pressure regulation, medical students collaboratively analysed either a problem of blood pressure regulation or a problem of vision. When formulating hypotheses regarding a specific problem, students relied on their prior knowledge to account for it in terms of an underlying process. Students who activated text-relevant prior knowledge about blood pressure regulation recalled more information from the text than students who activated text-irrelevant prior knowledge about vision. Prior knowledge activation functioned as a bridge between prior knowledge and knowledge still to be acquired. More specifically, problem-based discussion facilitated the integration of new information into the existing knowledge base, resulting in higher recall.

This chapter will focus on the use of external representations of low sophistication (i.e., simple pictures and animations, or brief notes with few interrelations) during prior knowledge activation in the science domain. Research on the use of external representations in prior knowledge activation is still quite limited and therefore, a theoretical framework that provides more insights into the effects of external representations on the process of prior knowledge activation is described. More specifically, it is assumed that external representations can be used to *prompt* (i.e., initiate) prior knowledge activation as well as *reinforce* (i.e., facilitate) the activation process. In addition, these prompting and reinforcing effects of external representations are hypothesised to be mediated by learners' level of prior knowledge (see Figure 12.1).

The structure of this chapter is as follows. First, the facilitative effects of prior knowledge activation on learning are described. What is prior knowledge activation and how does it facilitate learning? While answering this question, one prior knowledge activation strategy (i.e., mobilisation) is outlined. Second, the use

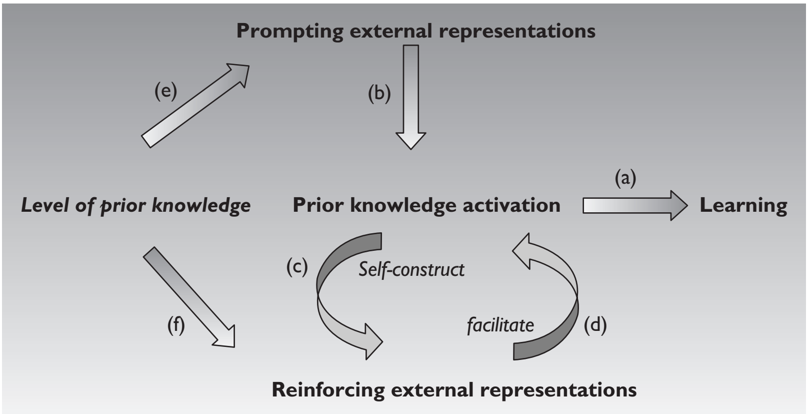


Figure 12.1 Theoretical framework illustrating the use of low-sophistication external representations in prior knowledge activation.

of external representations in prior knowledge activation is explored, addressing the question of how prior knowledge activation can be optimised through the use of external representations. Here, the different functions of external representations in prior knowledge activation are outlined. Third, the role of learners' level of prior knowledge on the effects of external representations in prior knowledge activation is explored. Finally, an empirical study is presented that provides support for specific parts of the theoretical framework.

Prior knowledge activation

In line with De Grave et al. (2001), many studies have provided evidence for a strong positive impact of prior knowledge activation on learning (see arrow (a) in Figure 12.1) (e.g., Goetz et al., 1983; Ozgungor & Guthrie, 2004; Verkoeijen et al., 2005). According to Mayer (1979, p. 134), learning involves 'relating new, potentially meaningful material to an assimilative context of existing knowledge'. This implies that it is not sufficient merely to possess prior knowledge. In order to reach higher learning outcomes, the available knowledge should be used actively during information processing in order to establish relationships between the already available knowledge and new information provided to learners (Mayer, 1979).

The accuracy and efficiency with which knowledge can be activated and used as a framework for integrating new information is influenced by the way knowledge is represented in memory. Existing knowledge is represented by an associative network of nodes and links (Kintsch, 1988). The nodes represent concepts, which are important units of knowledge. A concept is an idea about a phenomenon or object (e.g., cat, burglar) that is related to other concepts (e.g., animal, crime). The relations between different concepts are represented by the links that connect different nodes. This interconnected pattern of nodes (i.e., network) enables learners to meaningfully organise knowledge contained in these connections. If prior knowledge is activated, specific nodes in the network are activated. Because of the links between nodes, activation can easily spread from a specific node (e.g., heart) to other connected nodes (e.g., blood flow, love). The more often particular links between nodes are used, the stronger these links become. As a result of frequent use, learning takes place through strengthening of connections. In addition, the network provides a framework in which new information can be integrated, resulting in new links between nodes. This framework facilitates learning because it offers the opportunity to establish connections between new information and the existing knowledge contained in the pattern of nodes (Anderson, 1983). This bridges the gap between the existing knowledge base and new information that is provided to the learner.

A well-known technique for activating prior knowledge is mobilisation, where learners are encouraged to bring to mind all knowledge they have in a certain domain (Peeck, 1982). Machiels-Bongaerts, Schmidt, and Boshuizen (1993) asked students in two experimental groups to mobilise either names of US states or names of US presidents. A control group mobilised names of composers. Subsequently, all students studied a list containing the names of 32 US states and presidents. Time to study the list and individual items on the list were fixed. The experimental groups showed higher recall scores than the control group. This higher recall was caused entirely by enhanced recall of items of the mobilised category (i.e., states or presidents). Especially, items of the mobilised category that were not explicitly mobilised (e.g., less well-known president names, such as Coolidge or Polk) benefitted from mobilisation. So, the beneficial effects of mobilisation seemed to spill over to items that were not previously mobilised. During mobilisation, activation from mobilised items spread to items that were not retrieved but were nevertheless processed to some extent. Because of this spreading activation, non-mobilised items of the mobilised category also benefitted from mobilisation.

In another study, Machiels-Bongaerts, Schmidt, and Boshuizen (1995) encouraged students to mobilise all knowledge they had about the fishery policy of the European Union and its consequences. A control group activated prior knowledge about a neutral topic (i.e., tennis). Subsequently, all students studied a text about the consequences of the EU fishery policy for a fictitious fishery village. The text contained information that matched the activated prior knowledge of the experimental group (e.g., a rise in unemployment) and additional, new information (e.g., an alternative income source) that became important in light of the activated prior knowledge. The experimental group outperformed the control group in recall of information from the text. This higher recall was caused by enhanced recall of information that was explicitly activated and of the new information. By relating the activated prior knowledge to the new information, new links are established which facilitate the integration of this information into the existing knowledge base.

Until now, researchers have mainly used verbal instructions (e.g., 'bring to mind . . .') to activate learners' prior knowledge. External representations, such as pictures, animations, and notes, are rarely used for this purpose. The next section will explore this type of prior knowledge activation.

The use of external representations in prior knowledge activation

Before exploring the use and effectiveness of external representations in prior knowledge activation, several dimensions of external representations (i.e., verbal/pictorial, provided/self-constructed) and their effects on learning are outlined. Then, the different functions of external representations in prior knowledge activation are explored.

Dimensions of external representations

Verbal and pictorial external representations

Although external representations can come in many variants, there are only two basic forms: *verbal* (descriptive) and *pictorial* (depictive) representations. Verbal representations consist of symbols and are powerful in expressing abstract knowledge. Pictorial representations consist of icons and have the advantage of being 'informationally complete'. Because information can be directly inferred from them, pictorial representations are more useful for drawing inferences (Schnotz, 2005). This implies that the processing of pictorial representations may require less mental effort than the processing of verbal representations (Cox, 1999; Mayer, 2001). Larkin and Simon (1987) explain this by making a distinction between the informational and computational equivalence of external representations. Two representations are informationally equivalent if information that can be inferred from one representation can also be inferred from the other. For example, the manual of a DVD recorder may contain a text and a sequence of pictures that provide users with equivalent information on how to program the recorder. Informational equivalence is a precondition for computational equivalence. Representations are computationally equivalent if inferences that can be easily and quickly drawn from information given in one representation can also be easily and quickly drawn from the information that is explicitly provided in the other. Two representations that are informationally equivalent may, however, differ in their computational equivalence. For example, many users may have experienced that it is easier and quicker to use the pictures when programming the DVD recorder as compared with using the text. In this case, the pictures are more computationally efficient.

Although pictorial representations are often considered to be more computationally efficient than verbal representations, this may depend on the type of information (e.g., conceptual, spatial) that is contained in the representation. Pictorial representations that correspond on a one-to-one basis (i.e., are analogue) to the subject may indeed be more efficient when conveying spatial and temporal relations. Verbal representations that use symbols to represent the subject may be more efficient when conveying information about conceptual relations and logical sequences (Larkin & Simon, 1987; Schnotz, 2005).

Provided and self-constructed external representations

In addition to the verbal–pictorial dimension, external representations can be *provided* to learners or they can be *self-constructed* by the learner. Provided external representations have to be interpreted by learners (Cox, 1999). If learning materials are enriched with familiar external representations, this might facilitate learning because information can be coded both verbally and visually (Mayer, 2001). However, if learners are provided with a representation they are unfamiliar with, they might experience cognitive overload as a result of having to integrate, verbally and visually, this unfamiliar representation. This may enhance cognitive load which hampers learning. In these situations, it might be more beneficial for learners to self-construct a representation because learners can use the type of representation they prefer and are familiar with. De Westelinck and Valcke (2005) showed that students who were actively engaged in constructing external representations while studying learning materials scored higher on retention and transfer tests than students who studied the learning materials with provided external representations they were not familiar with.

Self-constructed external representations reveal learners' knowledge and the structure of that knowledge (i.e., its internal representation) by externalising this knowledge through the use of symbols and objects (Lee & Nelson, 2005). In addition, they can be used for clarification and elaboration of learners' own conceptual understanding. The process of constructing an external representation and interacting with it may foster learners' understanding, especially if the representation has a high level of sophistication (i.e., many interrelations). Therefore, self-constructing external representations can be an important component of learning. This is in line with the active processing assumption (Mayer, 2001) and the focused processing stance (Renkl & Atkinson, 2007), according to which, actively building external representations might promote organisation and integration processes that foster the development of mental models. This implies that constructing external representations may enhance cognitive load, which is beneficial for learning.

A well-known example of self-constructing external representations is taking notes. The overall effects of note taking on learning are positive (cf., Kiewra, 1985; Kobayashi, 2005). Note-taking research has primarily focused on learning from taking notes while attending a lecture (e.g., Austin et al., 2002; Kiewra et al., 1991) or reading a text (e.g., Kobayashi, 2009; Slotte & Lonka, 1999). Most studies have shown that learners who take notes reach higher learning outcomes than learners who do not take notes (e.g., Barnett, Di Vesta & Rogozinski, 1981). Externally representing information by means of note taking might support the organisation of information and the establishment of idiosyncratic relations between prior knowledge and the information provided in the lecture or text. This facilitates the comprehension of a lecture or text with beneficial effects on learning (Castelló & Monereo, 2005).

Research on the use of external representations for activating prior knowledge in the science domain is rather limited. However, external representations might serve important functions in the process of prior knowledge activation, *prompting* prior knowledge activation and *reinforcing* the activation process. It is important to emphasise here that external representations can differ in their level of sophistication on a continuum from low to high sophistication. Representations of low sophistication consist of simple pictures or brief notes with *few interrelations* that primarily help to *activate* possibly relevant knowledge by *offloading* memory. High-sophistication external representations are more elaborate pictures and notes that support learners to *activate* their prior knowledge *and construct* this knowledge by establishing *many interrelations.* These representations help learners to *elaborate* on their prior knowledge. In this chapter, the focus is on low-sophistication external representations for two reasons. First, external representations that are used to prompt prior knowledge activation should activate learners' prior knowledge and not provide information to learners. Second, low-sophistication external representations can be constructed by learners regardless of their level of prior knowledge. In contrast, learners need a considerable amount of prior knowledge to construct a high-sophistication external representation.

Prompting prior knowledge activation

Low-sophistication external representations could be used to *prompt* prior knowledge activation (see arrow (b) in Figure 12.1). Learners could be provided with an external representation of low sophistication and asked to activate their prior knowledge about a specific topic using this representation. Learners' understanding of the organisation and functioning of objects, events or activities (e.g., the structure of the circulatory system and the functioning of the heart) is an important part of science learning (Chi et al., 1994). Structural models are internal, pictorial models that describe how objects, events or activities are spatially or temporally related to each other. These models support learners' understanding of how a particular domain is organised. Causal models are internal, pictorial models that focus on how objects, events, or activities affect each other and help to interpret processes, give explanations, and make predictions. In these models, cause and effect relations play an important role which enables learners to see how a particular domain functions and how changes in one component are related to changes in other components (Van Merriënboer & Kirschner, 2007).

Structural and causal models are important for elaborating and refining knowledge in the science domain. Because pictures and animations correspond on a one-to-one basis to the subject they represent and are informationally complete (Schnotz, 2005), they might be better able to represent these kinds of models than verbal representations. Therefore, low-sophistication pictorial representations are expected to be more suitable to prompt prior knowledge activation. More specifically, pictures may be very useful to illustrate how a domain is organised in space, whereas animations may be very useful for illustrating how changes in one component affect changes in other components. This would imply that pictures might be most suitable for activating structural models, and animations for activating causal models. It is important to emphasise here that only pictorial representations of low sophistication are considered suitable for prompting prior knowledge activation. Although these representations contain more information and are thus more sophisticated than verbal representations, they do not contain any labels or additional explanative text. However, more sophisticated pictorial representations do contain accompanying text and thus convey more information. This implies that these representations are more susceptible for deducing information from, which may interfere with prior knowledge activation.

Reinforcing prior knowledge activation

Low-sophistication external representations might not only prompt prior knowledge activation, but might also *reinforce* the activation process. The reinforcing effect of external representations arises if learners are given an opportunity to self-construct a low-sophistication representation of their prior knowledge (see arrow (c) in Figure 12.1). When considering the beneficial effects of prior knowledge activation on learning, working memory is an important factor. Learners can hold about seven elements at a time in working memory (Baddeley, 1992; Miller, 1956). When required to process elements simultaneously, the capacity of working memory is even more severely limited; about two to three elements can be related or manipulated at any given time in working memory (Sweller, van Merriënboer, & Paas, 1998). If learners activate their prior knowledge, information is brought from long-term memory to working memory. As a result of the limited capacity of working memory, there are limits to the amount of information (i.e., the number of elements) that can simultaneously be held and processed in working memory (Baddeley, 1992; Miller, 1956). This implies that learners might be overwhelmed by the activation process, leading them to experience cognitive overload. If learners are overloaded, there is not enough capacity to activate all elements in the existing knowledge base, which will hamper the activation process (Van Merriënboer & Sweller, 2005).

Cognitive overload might be prevented if learners are given an opportunity to represent their prior knowledge externally by means of taking notes. Note taking enables learners to activate many concepts and to relate these concepts to one another without having to keep all the concepts active in working memory. This will facilitate the activation process by reducing the load imposed on working memory during prior knowledge activation (see arrow (d) in Figure 12.1). In addition, learners might be enabled to easily retrieve and hold these concepts in working memory when confronted with new information. If relations are built between the concepts activated during prior knowledge activation and new information provided to the learner, new links between nodes can be established. This facilitates the integration of new information into the existing knowledge base, with beneficial effects on learning.

Although externally representing prior knowledge by means of taking notes is primarily expected to have a reinforcing effect on the activation process, it may also serve as a prompt for additional prior knowledge activation. By taking notes, new ideas might be triggered in long-term memory because of the spreading of activation to interconnected nodes in the knowledge base (Anderson, 1983). If these ideas are subsequently written down, this may again reinforce the activation process. This implies that the prompting and reinforcing effects of low-sophistication external representations might be closely intertwined.

External representations, prior knowledge activation, and level of prior knowledge

In prior knowledge activation, low-sophistication external representations might serve as a prompt to activate prior knowledge and reinforce the activation process. However, the effects of external representations in prior knowledge activation might be mediated by learners' level of prior domain knowledge. If learners' prior domain knowledge is limited, low-sophistication external representations may be less effective in prompting correct and relevant prior knowledge than if learners possess more prior knowledge or more elaborate prior knowledge (see arrow (e) in Figure 12.1). Although pictorial representations might be more suitable to prompt prior knowledge activation than verbal representations for all learners, the effectiveness of pictures and animations as prompts might also depend on learners' level of prior knowledge. Pictures may be very useful for activating structural models, and animations for activating causal models. Before learners are able to build causal models, they need to possess some knowledge about how the domain is organised. Learners with relatively limited prior knowledge might possess knowledge about how the domain is structured but do not yet know how changes in one component result in changes in other components. For example, they know that the heart consists of atria, ventricles and valves, but they do not yet know that if the ventricles contract, the valves between atria and ventricles close. Animations might therefore be less beneficial for learners with limited prior knowledge, because such learners do not yet possess the knowledge that is triggered by the animations. For learners with high levels of prior knowledge who possess sophisticated structural *and* causal models, animations might be more effective than pictures, because they prompt both structural and causal knowledge.

The reinforcing effect of low-sophistication external representations may also be mediated by learners' level of prior knowledge (see arrow (f) in Figure 12.1). If learners have limited prior domain knowledge, it may be more difficult to self-construct a low-sophistication external representation which, in turn, might influence the beneficial offloading effect of note taking. For these learners, prior knowledge is not meaningfully organised because their knowledge is not yet represented in an interconnected pattern of nodes (Anderson, 1983). This makes it difficult for them to build an external representation by taking notes, because they cannot distinguish relevant from irrelevant concepts or draw relations between concepts (Anderson, 1977). Therefore, self-constructing low-sophistication external representations is not expected to have beneficial offloading effects on working memory for learners who have only limited prior domain knowledge.

In sum, as the framework presented in Figure 12.1 illustrates, low-sophistication external representations are assumed to play different roles in prior knowledge activation. They can be used to prompt prior knowledge activation and they can be used to reinforce the activation process. If low-sophistication external representations are used to *prompt* prior knowledge activation, the representation is *provided* to learners and preferably *pictorial*, because these representations may be more suitable to represent and activate structural and causal models that are important for learning in the science domain. Low-sophistication external representations are also assumed to *reinforce* the activation process. If learners *self-construct* a low-sophistication external representation of their prior knowledge by means of taking notes, the activation process is *facilitated* by reducing the load imposed on working memory. In addition, the effects of low-sophistication external representations are expected to be mediated by learners' *level of prior knowledge*. The prompting and reinforcing effects of external representations in prior knowledge activation are assumed to be stronger for learners who already possess sufficient prior domain knowledge.

A study by Wetzels, Kester, and van Merriënboer (2009) provided support for the mechanism of reinforcing prior knowledge activation (see the bottom part of Figure 12.1). This study investigated the effects of note taking on learning during prior knowledge activation, depending on learners' level of prior knowledge. High school students completed a prior knowledge test about the circulatory system (i.e., the structure of the circulatory system and the functioning of the heart). Students were assigned to a low prior knowledge or a high prior knowledge group based on the median score of the prior knowledge test.

About one week later, the experimental session took place. Before working on tasks about the circulatory system, students activated their prior knowledge prompted by two prior knowledge activation pictures that represented (a) the structure of the circulatory system and (b) the functioning of the heart. Because the same prompts were used for all students, the prompting effect was not investigated in this study. First, students were provided with the picture representing the structure of the circulatory system and encouraged to bring to mind (i.e., mobilise) all knowledge they had about how blood flows through the body. Think-aloud protocols were recorded to check what knowledge was being activated and whether this knowledge was correct. Subsequently, students worked on learning tasks about this topic. Students were, for example, given the following problem that had to be solved: '*A child cuts itself in its finger with a piece of glass resulting in bacteria entering the blood stream. What way do the bacteria travel through the circulatory system before they reach the kidneys?*'

After activating prior knowledge about the structure of the circulatory system and working on tasks about this topic, students activated their prior knowledge about the functioning of the heart. The picture illustrated in Figure 12.2 was used to activate students' causal model of heart functioning. Students were provided with this picture and encouraged to bring to mind all knowledge they had about the electrical system and the functioning of the heart. Again, think-aloud protocols were recorded. Subsequently, students worked on learning tasks about this more refined aspect of the circulatory system. An example of a learning task in this context was: '*How does the electrical system of the heart work?*' Half of the participants were allowed to represent their prior knowledge externally by

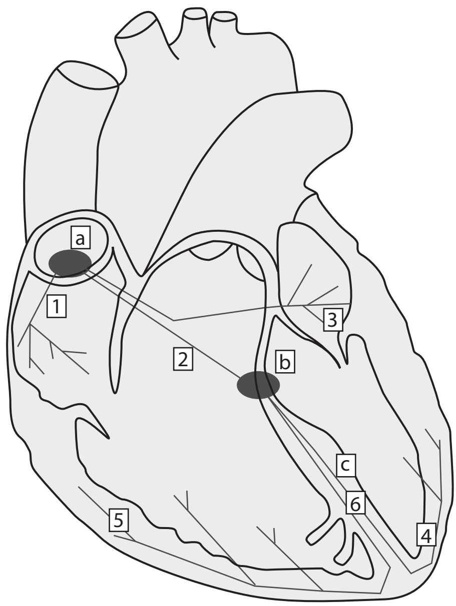


Figure 12.2 Picture used to activate prior knowledge about the functioning of the heart.

means of taking notes while activating their prior knowledge, whereas the other half were not allowed to take notes.

Finally, students worked on a number of transfer tasks concerning the structure of the circulatory system and the functioning of the heart. These tasks provided an indication of how well students understood what they had learned during the learning phase. More specifically, transfer tasks assessed whether students were able to transfer the principles (e.g., blood flows from the atria to the ventricles) they had learned while working on the learning tasks. Students had to apply these principles both in familiar situations (e.g., blood flow in a healthy individual) as well as in unfamiliar situations (e.g., blood flow in a child with a congenital heart defect). Students who took notes while activating their prior knowledge were not allowed to review or elaborate on their notes while working on the learning and transfer tasks. Students were also not allowed to take notes while working on the tasks.

Learning effectiveness and efficiency were measured by means of performance, mental effort, and mental efficiency. Mental effort represented the amount of effort students had to invest to solve a task as rated on a subjective rating scale. Mental efficiency was a combination of transfer test performance and invested mental effort during transfer. A high efficiency indicated a transfer test performance that was higher than expected based on the amount of invested mental effort during the transfer phase, whereas a low efficiency indicated a transfer test performance that was lower than expected based on invested mental effort (Paas & van Merriënboer, 1993). So, mental efficiency is a learning measure that provides information that goes beyond information provided by performance and mental effort measures alone.

Results showed that the efficiency of note taking (i.e., the reinforcing effect of external representations) during prior knowledge activation was influenced by the amount of prior domain knowledge learners already possessed. For learners with higher levels of prior knowledge about the circulatory system, note taking lowered mental effort while working on transfer tasks and increased mental efficiency during transfer. For learners with lower levels of prior knowledge, note taking yielded the opposite effect on mental effort and efficiency during the transfer phase. Figure 12.3 illustrates the interaction effect between level of prior knowledge and note taking on mental effort (A) and mental efficiency (B) during the transfer phase.

By representing their prior knowledge externally, learners with higher levels of prior domain knowledge are enabled to activate concepts and relate these concepts to one another without having to keep them active in working memory. The resulting low-sophistication external representation reduces the load imposed on working memory, while activating prior knowledge. This offloading effect of taking notes facilitates the activation process which enhances learning for high prior knowledge learners. However, if prior knowledge is very limited, learners might not be able to distinguish relevant from irrelevant concepts or draw relations between activated concepts. This makes it difficult for them to

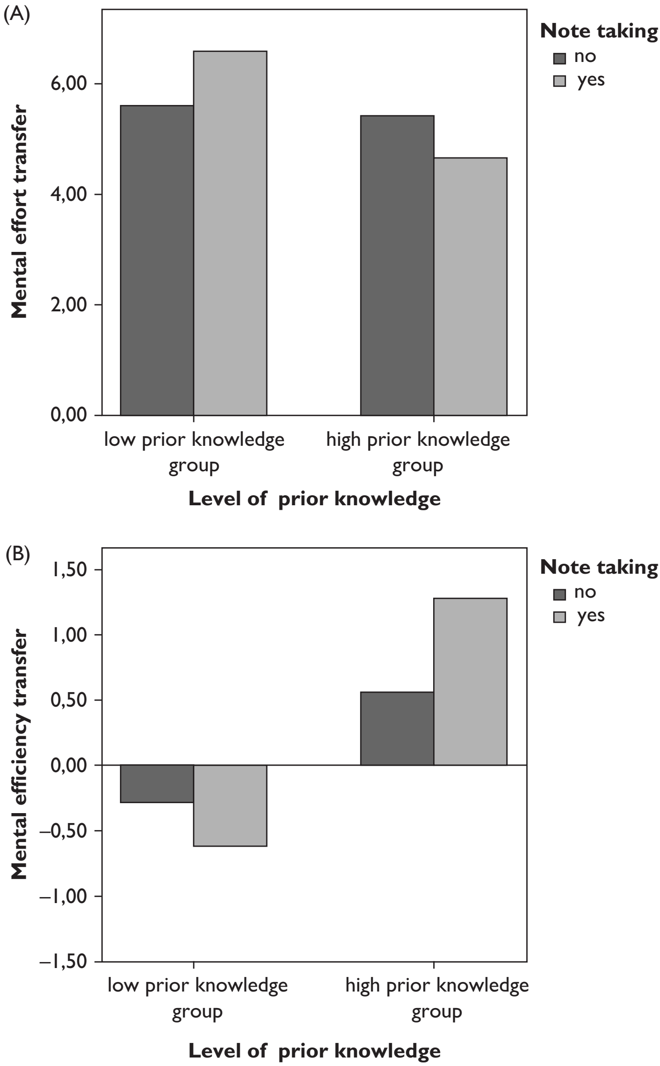


Figure 12.3 Interaction effect between level of prior knowledge and note taking for mental effort (A) and mental efficiency (B) during transfer.

build a low-sophistication external representation of their prior knowledge by taking notes. Therefore, note taking might not have had any offloading effects on working memory for low prior knowledge learners.

Surprisingly, high prior knowledge learners did not construct a more sophisticated external representation of their prior knowledge than low prior knowledge learners. The number of (correct) relations described in the notes was very low, regardless of learners' level of prior knowledge. This is consistent with Kiewra's (1985) observation that relational note taking is difficult for learners.

There was some support for the assumption that the prompting effect of external representations is mediated by learners' level of prior knowledge. Learners with higher levels of prior knowledge generated more concepts, more relations between activated concepts, and more correct relations in the think-aloud protocols than learners with lower levels of prior knowledge. High prior knowledge learners also generated more concepts in their notes compared with low prior knowledge learners. These results seem to suggest that pictorial representations prompt more elaborate and more organised knowledge in learners with more prior domain knowledge. In sum, both the prompting and the reinforcing effects of external representations seem to be influenced by how much prior domain knowledge learners already possess.

General discussion

In this chapter, a theoretical framework was outlined that described the effects of low-sophistication external representations during prior knowledge activation in the science domain. First, it was suggested that low-sophistication external representations can be used to prompt prior knowledge activation. External representations that are used as prompts to activate prior knowledge can be provided to learners. In addition, these representations are preferably pictorial; pictorial representations are assumed to be more suitable for representing and prompting structural and causal models that are important for science learning. Second, low-sophistication external representations were considered to reinforce the activation process. By self-constructing an external representation of learners' prior knowledge, the load imposed on working memory during prior knowledge activation is reduced. This was expected to facilitate the activation process and consequently learning. Third, it was outlined that the prompting and reinforcing effects of external representations might be mediated by learners' level of prior knowledge. More specifically, these effects were assumed to be more pronounced for learners with relatively higher levels of prior domain knowledge. Finally, the mechanism for reinforcing prior knowledge activation and the influence of learners' prior knowledge on the reinforcing effect of low-sophistication external representations was supported by the results of an empirical study.

The theoretical framework described in this chapter is based on prior knowledge activation in the science domain in which the activation of structural and causal models are important for learning. This implies that the framework, and especially the prompting part of it, might be less applicable for more conceptually oriented domains in which the organisation and the functioning of objects, events, or activities are not essential for learning. Another limitation of the framework is that it does not consider any other learner characteristics than prior knowledge. For example, learners with different levels of prior knowledge may also differ in intelligence, motivation, or interest, which may influence the activation process and learning.

Despite these limitations, the framework provides interesting insights into the various variables that may be involved in prior knowledge activation. In addition, the theoretical framework broadens the note-taking research. The traditional note-taking research focused on the encoding and the external storage effect of note taking (Di Vesta & Gray, 1972). The encoding function of note taking signifies that the process of taking notes while attending a lecture or reading a text is beneficial for learning. So, the encoding effect represents the effects of note taking *during* learning. The external storage function signifies that having notes available for review after attending a lecture or reading a text is beneficial for learning. So, the external storage function represents the effects of note taking *after* learning. However, in this chapter, the effects of note taking *before* learning are investigated. Learners take notes while activating their prior knowledge before they are provided with learning materials.

In this chapter, it is argued that self-constructing external representations by means of note taking externalises the internal representations of knowledge. Note taking enables learners to represent their prior knowledge externally, which reduces the load imposed on working memory. However, for learners with relatively high levels of prior knowledge, note taking may result in an active, constructive process. They build a high-sophistication external representation that not only represents prior knowledge externally, but also reconstructs this knowledge. If this happens, cognitive load may increase as a result of effortful learning.

Future research should focus on several aspects of the framework. The first line of research could focus on the prompting effect of external representations and how this is mediated by learners' level of prior knowledge. It would be interesting to explore whether, and under which circumstances, pictures and animations are more efficient in prompting prior knowledge activation. Are pictures, indeed, more suitable for activating structural models, and animations more suitable for activating causal models? And how is the effectiveness of pictures and animations influenced by learners' prior knowledge? When investigating the prompting effect of pictures and animations, the possibility that learners learn from an external representation should be considered. Even if low-sophistication pictorial representations are used to prompt prior knowledge activation, learners may deduce information from it. This implies that the prompt might provide learners with new knowledge, which may result in learning even though this probably will not exceed the recognition level. Therefore, it should be investigated how genuine prior knowledge activation can be discerned from information that is deduced from pictures and animations.

A second line of research is related to self-constructing external representations by taking notes. More specifically, self-constructing external representations might not only have reinforcing effects but also serve as a prompt for prior knowledge activation. Prior knowledge might initially be prompted by the provided pictorial representation, with further prompts resulting both from the provided pictorial representation and the self-constructed representation. The extent to which self-constructing an external representation may serve as a prompt for further prior knowledge activation may again depend on learners' level of prior knowledge.

Although this seems a very plausible and interesting idea, it might be quite difficult to disentangle the prompting effect resulting from the provided pictorial representation from the prompt that results from the self-constructed representation. This implies that it is also important to explore if and how these prompting effects could be differentiated.

A third line of research might further investigate the influence of learners' level of prior knowledge on the reinforcing effect of self-constructed external representations. The study of Wetzels et al. (2009) showed that externally representing prior knowledge by means of taking notes was more beneficial for learners with sufficient prior knowledge. However, all participants in this study were high school students. So, all participants might have been on the low end of the expertise continuum. The question is how increasing and stronger differentiated levels of prior knowledge affect the reinforcing effect of external representations. Does this effect get stronger for learners who are on the higher end of the expertise continuum (e.g., medical students)? Or perhaps self-constructing external representations has no beneficial offloading effects on working memory for learners with higher levels of prior knowledge because these learners may easily hold a representation of their prior knowledge in working memory without overloading it. This might imply that these learners will not benefit from self-constructing external representations and that the reinforcing effect is not as strong as for learners with intermediate levels of prior knowledge. These issues could be tackled in future research.

A practical implication that follows from the presented framework is related to teaching practices. Encouraging learners to represent their prior knowledge externally might facilitate the activation process and learning, but only for learners who already have sufficient prior domain knowledge. For learners with too little prior knowledge, self-constructing external representations might not have any beneficial offloading effects on working memory. Therefore, teachers should take their students' level of prior knowledge into account when asking them to self-construct a low-sophistication external representation of their prior knowledge.

In sum, the presented framework provides more insights into how low-sophistication external representations can be used to support the process of prior knowledge activation and how this is influenced by learners' level of prior knowledge. However, future research is necessary to elaborate and refine the framework further.

References

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, R. C. (1977). The notion of schemata and the educational enterprise: General discussion of the conference. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 415–431). Hillsdale, NJ: Erlbaum.

Austin, J. L., Gilbert Lee, M., Thibeault, M. D., Carr, J. E., & Bailey, J. S. (2002). Effects of guided notes on university students' responding and recall of information. *Journal of Behavioral Education, 11*, 243–254.

Baddeley, A. D. (1992). Working memory. *Science, 255*, 556–559.

Barnett, J. E., Di Vesta, F. J., & Rogozinski, J. T. (1981). What is learned in note taking? *Journal of Educational Psychology, 73*, 181–192.

Castelló, M., & Monereo, C. (2005). Students' note-taking as a knowledge construction tool. *Educational Studies in Language and Literature, 5*, 265–285.

Chi, M. T. H., de Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439–477.

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

De Grave, W. S., Schmidt, H. G., & Boshuizen, H. P. A. (2001). Effects of problem-based discussion on studying a subsequent text: A randomized trial among first year medical students. *Instructional Science, 29*, 33–44.

De Westelinck, K., & Valcke, M. (2005). The impact of external graphical representations in different knowledge domains: Is there a domain effect? In L. Verschaffel, E. De Corte, G. Kanselaar, & M. Valcke (Eds.), *Powerful environments for promoting deep conceptual and strategic learning* (pp. 213–232). Leuven, Belgium: Leuven University Press.

Di Vesta, F. J., & Gray, G. S. (1972). Listening and notetaking. *Journal of Educational Psychology, 63*, 8–14.

Goetz, E. T., Schallert, D. L., Reynolds, R. E., & Radin, D. I. (1983). Reading in perspective: What real cops and pretend burglars look for in a story. *Journal of Educational Psychology, 75*, 500–510.

Kiewra, K. A. (1985). Investigating notetaking and review: A depth of processing alternative. *Educational Psychologist, 20*, 23–32.

Kiewra, K. A., DuBois, N. F., Christian, D., McShane, A., Meyerhoffer, M., & Roskelley, D. (1991). Note-taking functions and techniques. *Journal of Educational Psychology, 83*, 240–245.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163–182.

Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination. *Contemporary Educational Psychology, 30*, 242–262.

Kobayashi, K. (2009). Comprehension of relations among controversial texts: Effects of external strategy use. *Instructional Science, 37*, 311–324.

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

Lee, Y., & Nelson, D. W. (2005). Design of a cognitive tool to enhance problem-solving performance. *Educational Media International, 42*, 3–18.

Machiels-Bongaerts, M., Schmidt, H. G., & Boshuizen, H. P. A. (1993). Effects of mobilizing prior knowledge on information processing: Studies of free recall and allocation of study time. *British Journal of Psychology, 84*, 481–498.

Machiels-Bongaerts, M., Schmidt, H. G., & Boshuizen, H. P. A. (1995). The effect of prior knowledge activation on text recall: An investigation of two conflicting hypotheses. *British Journal of Educational Psychology, 65*, 409–423.

Mayer, R. E. (1979). Twenty years of research on advance organizers: Assimilation theory is still the best predictor of results. *Instructional Science, 8*, 133–167.

Mayer, R. E. (2001). *Multimedia learning.* Cambridge, UK: Cambridge University Press.

Miller, G. A. (1956). The magical number seven, plus or minus two. Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Ozgungor, S., & Guthrie, J. T. (2004). Interactions among elaborative interrogation, knowledge, and interest in the process of constructing knowledge from text. *Journal of Educational Psychology, 96*, 437–443.

Paas, F., & van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors, 35*, 737–743.

Peeck, J. (1982). Effects of mobilization of prior knowledge on free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 608–612.

Renkl, A., & Atkinson, R. K. (2007). Interactive learning environments: Contemporary issues and trends. An introduction to the special issue. *Educational Psychology Review, 19*, 235–238.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49–69). New York: Cambridge University Press.

Slotte, V., & Lonka, K. (1999). Review and process effects of spontaneous note-taking on text comprehension. *Contemporary Educational Psychology, 24*, 1–20.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

Van Merriënboer, J. J. G., & Kirschner, P. A. (2007). *Ten steps to complex learning. A systematic approach to four-component instructional design*. Mahwah, NJ: Erlbaum.

Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*, 147–177.

Verkoeijen, P. P. J. L., Rikers, R. M. J. P., Augustus, M. I., & Schmidt, H. G. (2005). A critical look at the discrepancy reduction mechanism of study time allocation. *European Journal of Cognitive Psychology, 17*, 371–387.

Wetzels, S. A. J., Kester, L., & van Merriënboer, J. J. G. (2009). *The influence of prior knowledge on the retrieval-directed function of note taking in prior knowledge activation.* Manuscript submitted for publication.

Chapter 13

Visualisation of argumentation as shared activity

Gijsbert Erkens, Jeroen Janssen, and Paul A. Kirschner

Introduction

Computer Supported Collaborative Learning (CSCL) systems are assumed to have the potential to enhance the effectiveness of peer learning interactions (Andriessen et al., 1996; Dillenbourg, 1999). Groupware programmes are used for CSCL as they generally support and integrate three functions: task support, communicative support, and group support. Computer tools in groupware programmes are either task-oriented (information sharing, cooperation and coordination), communication-oriented (interpersonal exchange) or group-oriented (Andriessen, 2003). They are meant to support collaborative group work by sharing tools and resources between group members, by supporting group dynamics, and by giving communication opportunities within the group and to the external world.

Shared argumentation maps are task-related tools that are often used in CSCL. They are constructed by the collaborating students and are designed to be helpful in completing the inquiry task at hand (e.g., CSILE: Scardamalia, Bereiter & Lamon, 1994; *Belvédère*: Suthers et al., 1995). The maps visually represent the argumentative structures the students agree upon. The aim of this chapter is to investigate the effects of argumentative maps or diagrams on students working on collaborative writing tasks. Often, inconsistent or disappointing results are found in the way argumentative diagrams support reasoning and discussion in CSCL (Van Drie et al., 2005). We will present two research projects that investigated the effects of two types of argumentative diagrams for supporting collaborative writing and inquiry. In both the COSAR (Computer Support for Collaborative and Argumentative Writing) and the CRoCiCL (Computerized Representation of Coordination in Collaborative Learning) projects, argumentative diagrams were used to support collaboration and argumentation on inquiry tasks. In the COSAR project the effectiveness of the argumentative diagram for the quality of the students' group products was disappointing. For the CRoCiCL project we redesigned the representational features of the argumentative diagram. Although the purpose of the two tools was similar, the effectiveness of the diagram for stimulating the quality of the students' group products improved. For the explanation of these differences in effectiveness, we offer some ideas but no definitive answers, as the two projects differed in more aspects and the two tools were not compared directly. However, we hypothesise that the representational guidance (Suthers, 2003; Suthers & Hundhausen, 2003) the two tools offer to collaborating students differs substantially. The differences in guidance may have resulted in differences in the effectiveness of the argumentative tools on the quality of students' group products.

Representational guidance of two argumentative maps

Representational guidance refers to the fact that different representations are capable of expressing different information, making different information salient or stimulating different cognitive processes than others (Suthers & Hundhausen, 2003). Several studies investigating the effects of different argumentation tools showed that representational guidance can influence students' behaviour and learning process (e.g., Schwarz et al., 2003; Suthers, 2001; Van Amelsvoort, Andriessen & Kanselaar, 2007). Van Bruggen, Boshuizen and Kirschner (2003) distinguish five characteristics of representation tools that affect representational guidance: ontology (i.e., the type of representing elements), perspective (i.e., the view on the subject matter the representation allows), specificity (i.e., the categorical choice the representation forces, see also Suthers, 2001), precision (i.e., the accuracy of representation) and modality (i.e., form of expression: graph, text, list, matrix etc.).

In the COSAR project, we examined the effects of the Diagram, a tool used for constructing argumentation maps (see Figure 13.1). The Diagram is a shared tool for generating, organising and relating arguments in a graphical knowledge structure comparable to Belvédère (Suthers & Hundhausen, 2003; Suthers et al., 1995). The tool was conceptualised to the students as a graphical summary of the arguments in an essay. Students were instructed that the information contained in the diagram had to faithfully represent the information in the final version of their essay. This requirement was meant to help students notice inconsistencies, gaps, and other imperfections in their texts, and encourage them to review and revise. In the Diagram, several types of text boxes can be used: information ('informatie'), position ('standpunt'), argument pro ('voorargument'), support ('onderbouwing'), argument contra ('tegenargument'), refutation ('weerlegging'), and conclusion ('conclusie').

In the CRoCiCL project, we investigated the effects of the Graphical Debate tool (GD tool, see Figure 13.2). In this project, students were required to co-construct a representation of a historical debate. Comparable to the COSAR project, this activity precedes a writing task where students have to co-author an essay. The GD tool was designed after our experiences with the Diagram in the COSAR project.

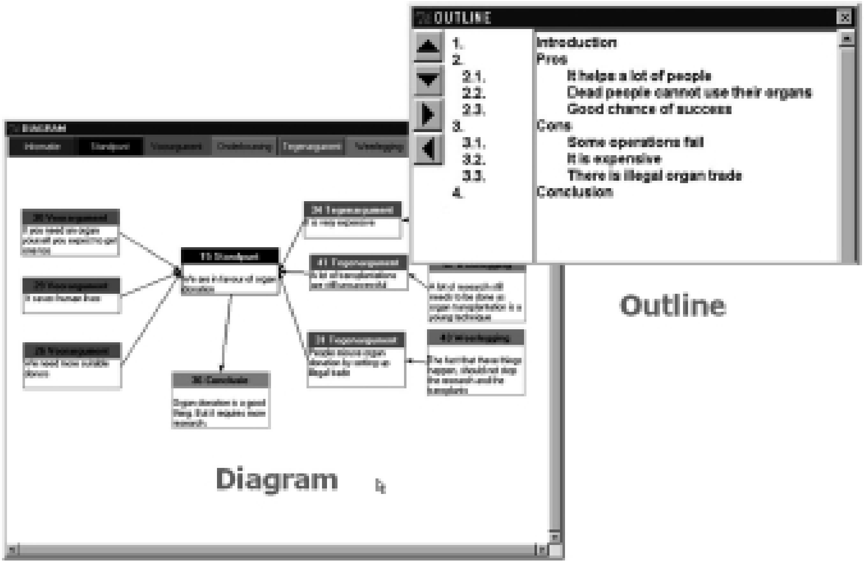


Figure 13.1 The Diagram and outline in the TC3 programme (translated from Dutch).

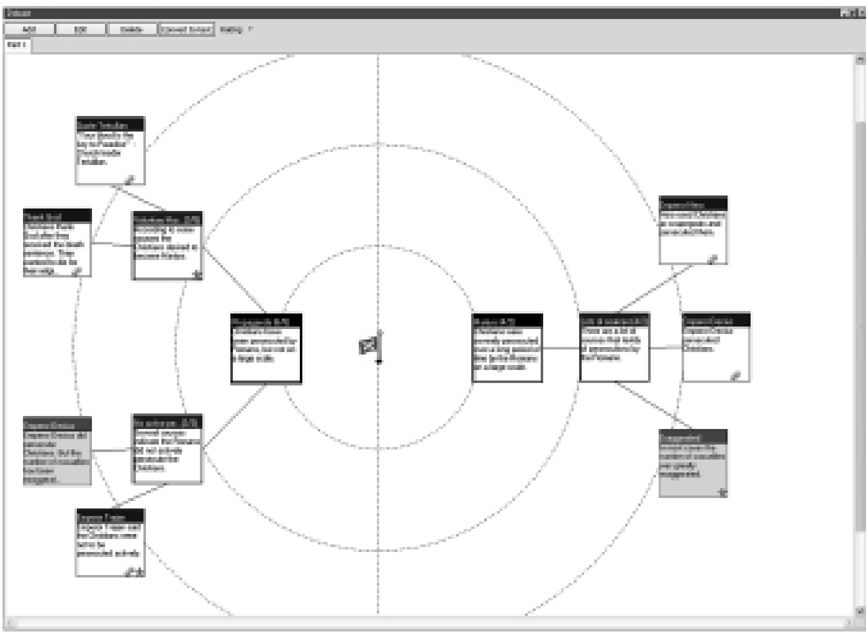


Figure 13.2 Screenshot of the graphical debate-tool (translated from Dutch).

The boxes labelled Martyrs and Propaganda represent both positions of the debate. While working with the GD tool, students can add arguments to either of the positions. These arguments can be found in the given sources. The sources also contain information that supports or refutes the arguments students add to the tool. Elements that represent supporting information have a white background, while elements that represent refuting information have a grey background.

The GD tool visualises how well positions are supported by arguments and supporting information. Each time an argument or a supporting piece of evidence is added to a position, it moves closer to the central flag. Conversely, when a refutation is added, the position moves away from the flag. Thus, when a position is located closer to the flag, it is better supported by arguments: the argumentation is more strongly in favour of the position. The embedded representational guidance of the GD tool may help students draw a conclusion about the debate and thus may contribute to computational offloading (Ainsworth, 2006). The GD tool also visualises students' progress through the problem (Cox, 1999). For example, the boldness of the lines around the position and argument boxes serves as an indication for their elaborateness and complexity. Finally, in the GD tool students have the option to rate the quality of arguments, supports, and refutations. Students can express this by giving ratings to arguments, positions, and refutations (indicated by the star in the corresponding boxes). A rating influences the distance of the position from the flag. When a rating is given to an argument, support or refutation, its corresponding position moves closer to or away from the flag. The rating functionality of the GD tool stimulates students to think about and discuss the importance of arguments and may help them to see which arguments are more important than others.

Table 13.1 contains a comparison between the Diagram used in the COSAR project and the GD tool used in the CRoCiCL project. From this Table important differences in the representational guidance offered by both tools become apparent. The most important differences concern the perspective, specificity and precision of both tools. The Diagram offers a perspective on argumentation comparable to a concept map: students can construct a map containing their own arguments and arguments found in the sources. The GD tool offers a battlefield perspective on argumentation: arguments advance or retract based on how well they are supported. It can also be argued that the GD tool gives more specific guidance than the Diagram does, because it gives feedback about the strength of argumentation and because it draws attention to the relative weight of arguments. On the other hand, it can be argued that the specificity of the Diagram is greater because students can use a larger number of different elements to construct their argumentation map.

It can be hypothesised that the differences in representational guidance offered by the Diagram and the GD tool will affect students' collaborative process (Suthers, 2006; Van Drie et al., 2005). In the remainder of this chapter we will describe two studies that – separately – investigated the effects of the

Table 13.1 Features of the argumentation maps compared between the two studies.

Features of the argumentation map	*COSAR-project: Diagram*	*CRoCiCL-project: Graphical Debate tool*
• Positioning of elements	Free – anywhere on screen	Constrained by tool
• Guidance – ontology	Elements of an argumentation (position, argument pro, argument contra, support, refutation, conclusion) and relations between them	Elements of a debate (position, argument, support, refutation)
• Guidance – perspective	Argumentation as concept map: Construction of own arguments and arguments found in sources	Argumentation as battle field: Reconstructing a debate from arguments found in sources
• Guidance – specificity	No feedback about strength of argumentation	Feedback about strength of argumentation, attention for weight of arguments
• Guidance – precision	Larger number of elements	Smaller number of elements
• Guidance – modality	Graphical and textual	Graphical and textual

Diagram and GD tool. Although both studies were similar in certain aspects, other features of the studies differed (see Table 13.2). In spite of these differences, we offer tentative suggestions in the general discussion for why the effects of Diagram and GD tool differed.

The COSAR project

The COSAR project investigated the effects of using argumentative diagrams for argument generation and organisation, compared with outlining tools for argument linearisation in collaborative writing of source-based argumentative texts. In argumentative writing (in contrast to narrative writing) the generation and organisation of arguments and ideas and the linearisation of the collected arguments in a linear text are the biggest problems for novice writers (Andriessen et al., 1996). A Diagram tool and an Outline tool were developed to support these specific writing processes.

A groupware environment called TC3 (Text Composer, Computer supported & Collaborative) was developed with which pairs of students collaboratively write argumentative essays (Erkens et al., 2005). This environment combines a shared word processor, a chat facility, and access to a private notepad and online information sources. Each partner works at his/her own computer, and wherever possible partners were assigned to different classrooms. The basic TC3 environment, shown in Figure 13.3, contains four main windows, of which the upper two windows are private and the lower two are shared:

- INFORMATION (upper right window): This private window contains tabs for the assignment ('i'), sources ('bron') and TC3 operating instructions. Sources are divided evenly between students. Each partner has three or five different sources plus one – fairly factual – common source.
- NOTES (upper left window, 'AANTEKENINGEN'): A private notepad where students can make non-shared notes.
- CHAT (lower left, three small windows): The student adds his/her chat message in the bottom box. Every letter typed is immediately sent to the partner via the network, so that both boxes are WYSIWIS: What You See Is What I See. The middle box shows the incoming messages from the partner. The scrollable upper chat box contains the discussion history.
- SHARED TEXT (lower right window, 'GEMEENSCHAPPELIJKE TEKST'): A simple word processor (also WYSIWIS) in which the shared text is written while taking turns.

In addition, two representational tools and a supporting facility were developed for the experimental conditions: the Diagram (described above), the Outline, and the Advisor. The Outline (see Figure 13.1) is a shared tool for generating and organising information units as an outline of consecutive arguments in the text. This tool was conceptualised to the students as producing a meaningful outline of the paper, and as is the case for the Diagram, the participants were

Table 13.2 Comparison of the features of the COSAR- and CRoCiCL-study.

Features of the study	*COSAR-study: Diagram*	*CRoCiCL-study: Graphical Debate tool*
• Assignment to groups	Random	Random
• Group size	Dyads	Mostly groups of three
• Task	Writing task based on sources	Writing task based on sources
• Duration	4–6 lessons	8 lessons
• Subject	Humanities: Social studies	Humanities: History
• Control condition(s)	Basic environment augmented with Outline and/or Advisor or basic environment only	Basic environment augmented with Textual Debate-tool
• Dependent variables	Quality of written texts, collaborative process focused on task-related and meta-cognitive activities	Quality of written texts, collaborative process focused on task-related, meta-cognitive, social, and meta-social activities, quality of representation, post-test performance
• Control variables	Pre-test of writing and argumentation skills	Pre-test on subject matter knowledge

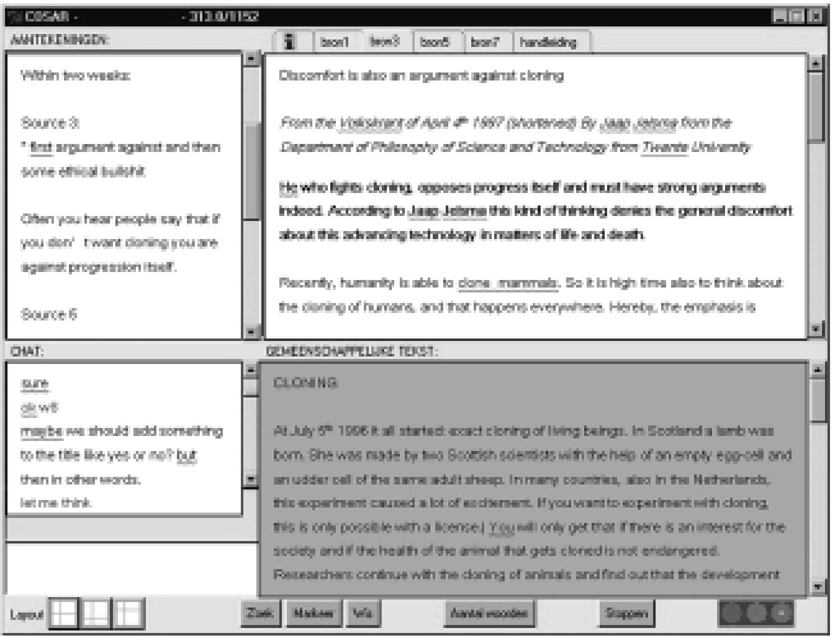


Figure 13.3 The Interface of the basic TC3-environment (translated from Dutch).

required to have the information in the Outline faithfully represent the information in the final text. The Outline tool was designed to support planning and organisation of the linear structure of the texts. In addition, the Outline tool has the pedagogic function of making the user aware of characteristics of good textual structure, thus allowing the user to learn to write better structured texts. The Outline has a maximum of four automatically indented, numbered levels. Both planning windows are WYSIWIS.

The Advisor is an extra help facility that provides advice on how to use the Diagram and/or Outline before and during task fulfilment. The Advisor consists of a tab sheet added to the information window, with tips and instructions for optimum use of the representational tools: the Diagram or the Outline.

Method

Design

The experiment was executed in two phases. In the first year, a control group (39 dyads) fulfilled the collaborative writing task in the basic TC3 environment without Diagram, Outline, or Advisor. In the second year, six experimental

Table 13.3 Experimental design.

Abbreviation	*Condition*	*Tools & facilities*	*No. dyads*	*Year*
C	Control	Basic TC3	39	1
D	Diagram	Basic TC3 + Diagram	17	2
DA	Diagram Advisor	Basic TC3 + Diagram + Advisor	26	2
DO	Diagram Outline	Basic TC3 + Diagram + Outline	23	2
DOA	Diagram Outline Advisor	Basic TC3 + Diagram + Outline + Advisor	11	2
O	Outline	Basic TC3 + Outline	18	2
OA	Outline Advisor	Basic TC3 + Outline + Advisor	11	2

groups (106 dyads in total) fulfilled the same task in the basic TC3 environment in which the planning tools and/or advising facility added were varied (see Table 13.3).

To control for school effects, classes from different schools were assigned to each condition. To control for differences in writing and argumentation skills, two pre-tests were administered individually before students worked on collaborative writing tasks. No systematic differences between students from different school classes were found in writing or argumentative competencies.

Participants

Participants were 290 students, aged 16–18, from six secondary schools in the Netherlands. The assignment was completed during four to six lessons. The analysed sample included 151 girls and 139 boys. All students from a class were randomly assigned to pairs by the experimenter on the basis of the list of names provided by the teacher. As the writing task for the students was part of the school curriculum (the essays were graded), it was not possible to stratify the group formation on argumentative competence. Mixed gender dyads comprised 58 pairs of the total sample; 46 dyads were all female and 41 were all male.

Task

The collaborative writing task was to write an argumentative essay of 600–1,000 words in Dutch on cloning or organ donation. The assignment was to convince the Minister of Health, Welfare, and Sport of the position the students chose to defend. The arguments for or against the position had to be based upon facts and discussions about the issue presented in external information sources. The sources were taken from the Internet sites of Dutch newspapers. Each student had access to one common source and half of the remaining sources. By dividing the sources between the students, they were stimulated to discuss the relevance of the information for their common text. In all dyads, partners were seated in separate computer rooms to encourage them to only communicate through TC3. The students received grades for their texts from their teachers as part of their normal school work. These grades were separate from the scoring of the essays by two of the experimenters.

Analyses

Each of the 145 essays was coded on several dimensions. Before coding, the experimenters manually divided the texts into segments, largely based on the existing paragraph structure. The texts were scored on four variables on a scale of 1–10:

- Textual structure: formally defined by introduction, body, and conclusion;
- Segment argumentation: argumentative quality of the paragraphs;
- Overall argumentation: quality of the main line of argumentation in the text, and
- Audience focus: presentation towards the reader and level of formality of the text.

The interrater reliability for these measures was very high, with correlations between two independent raters for the four text scores on five texts ranging from .71 to 1.00 ($p < .01$).

Results

Quality of essays

First, the tool conditions in relation to the quality of the essays will be discussed. Table 13.4 shows the means and standard deviations of quality scores of the argumentative texts for all conditions separately and for the sample as a whole.

The table shows that the scores were quite similar for all groups. Independent samples t-tests showed no differences between the two topics – organ donation and cloning – and there were no significant gender differences between female, male or mixed groups. The mean quality of the texts was 6.2 on a scale of 1–10. We only found a few differences in a multiple comparison analysis (Bonferroni) on the conditions: the Diagram-Advisor group had significantly lower scores on textual structure of the essays in comparison with the Control, the Diagram and the Diagram-Outline-Advisor conditions (mean differences: −.73, −.68 and −1.12, all $p < .05$) and had a significantly lower score on segment argumentation in comparison with the Control condition (mean difference: .70, $p < .05$).

Table 13.4 Descriptive statistics for text quality per condition.

Condition	*n*	*Textual structure*		*Segment argumentation*		*Overall argumentation*		*Audience focus*	
		Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Control	39	6.76	1.13	6.19	1.36	5.75	2.37	6.20	2.10
Diagram	17	6.71	.97	5.63	1.34	6.81	2.29	5.81	1.84
Diagram + Advisor	26	6.03	.82	5.49	1.34	6.41	2.07	6.01	1.64
Diagram + Outline	23	6.44	.83	5.64	1.32	6.16	2.25	6.20	1.60
Diagram + Outline + Advisor	11	7.15	.88	5.42	.84	5.76	1.69	5.57	1.00
Outline	18	6.59	1.00	5.90	1.06	5.74	1.80	6.04	1.95
Outline + Advisor	11	6.49	.83	6.34	.94	5.76	1.52	6.59	1.90
Total	145	6.56	1.00	5.83	1.28	6.06	2.13	6.08	1.81

In general, we can say that the representation tool conditions in themselves did not have a positive effect on the quality of the resulting texts.

However, the availability of a tool is no guarantee of adequate use. Comparing the frequency of use of the Diagram with the frequency of use of the Outline, the Outline tool was more successful. Use of the Outline tool was weakly positively related with text quality ($r = .13$, $p < .05$). The use of the Diagram was even negatively correlated with text quality ($r = -.25$, $p < .01$), except for the last phase of writing ($r = .17$, $p > .05$). These relations were even stronger when the use of the Advisor was correlated to text quality (Diagram-Advisor, $r = -.21$, $p < .05$ and Outline-Advisor, $r = .44$, $p < .01$). Correlation analyses showed that the frequency of using the Diagram to specify supports and refutations of positions tended to be weakly positively related to segment argumentation ($r = .21$, $p < .05$ and $r = .13$, $p < .07$). Furthermore, the more the Diagram was used for specifying arguments from the sources instead of self-generated arguments, the less the overall argumentative quality of the texts proved to be ($r = -.21$, $p < .05$), ending up in an enumeration of arguments in the text. As for the Outline tool, a positive effect was found of the proper use of the Outline (especially in outline-text congruence) and its Advisor on segment argumentation in the resulting argumentative text ($r = .26$, $p < .01$ and $r = .23$, $p < .01$).

Contrary to these findings, however, an evaluation survey of students and teachers showed that both groups evaluated both tools, but especially the Diagram, as useful and helpful.

Online collaboration

All utterances in the chat discussions were coded with a coding system consisting of three main levels: metacognitive (planning and monitoring), cognitive (executive) and non-task or social. In this Task Act coding system 34 different categories were distinguished in total. Reliability analyses showed Cohen's κ's of .57 and 64. Both Diagram and Outline affected the collaborative chat discussion of the students in a substantial way: 70 per cent of the utterances versus only 47 per cent in the control condition were on a metacognitive level. That is, both the Diagram and Outline increased by how often students deliberately planned and monitored the task completion.

In order to find an explanation of these findings, a qualitative analysis of the discussion of the students while using the tools was undertaken. Analyses of the chat protocols showed that the Diagram often functioned as a visual representation for arguments mentioned, but not as a basis for discussion or as a tool for idea generation. Thus, the Diagram only functioned as a visual summary, and not as a basis for discussion of the argumentative structure or as a tool for generating and organising new ideas and arguments. When a diagram stimulates and reflects the discussion itself, it can be a valuable starting point for writing the text, and can benefit the textual structure. A more guiding function of the representation tool might encourage the students to use it as it was intended, and thus lead to different results.

Conclusions for COSAR project

In the COSAR project, a complex relationship between the use of representation tools, like the Diagram and Outline, and the argumentative quality of the texts was found. Inconsistent, small and even negative relations existed between using the argumentative Diagram and the final argumentative text. Positive relations, although small, were found between the use of the Outline linearisation tool and text quality. However, both representation tools seemed to stimulate discussion and coordination on a planning level in the collaborative chat of the students. Furthermore, both students and teachers evaluated the Diagram tool as very valuable and useful (more useful than the 'more effective' Outline tool). The question therefore remains, why the argumentative Diagram did not help students write better-grounded texts. We assumed that the tool offered too little representational guidance to students. The free, unrestricted manner in which the arguments can be displayed in the Diagram by the students, prevents them from getting a systematic insight in the argumentative structure and organisation of the debate they study. Furthermore, the loose graphical structure of an argumentative map gives no indication of the relative strength of the positions depicted. In the CRoCiCL project, we tried to develop an argumentative diagram tool that offers more representational guidance with regard to the argumentative structure of the debate and the argumentative strength of the positions.

The CRoCiCL project

The aim of CRoCiCL was to examine how the representational guidance offered by an argumentative diagramming tool influenced the collaborative process. To answer this question, the Graphical Debate tool (described above) was compared with a Textual Debate tool (TD tool, see Figure 13.4). In this version of the tool, students also add arguments to the corresponding positions. No distinction is made, however, between arguments, supports and refutations. Instead, information is added to the TD tool in a list-wise manner (cf., Erkens et al., 2005; Van Drie et al., 2005). On the other hand, this makes the TD tool somewhat comparable to the Outline tool from the COSAR project, because both tools stimulate students to organise arguments in a list. However, the TD tool also differs from the Outline tool because it uses – like the GD tool – given positions. The process of co-constructing representations (i.e., reading and processing historical sources, extracting relevant information, placing this information in the appropriate position in the representation) is almost the same for both versions of the Debate tool.

The main difference between the GD tool and the TD tool concerns the representational guidance they offer (Suthers, 2001, 2003; Suthers & Hundhausen, 2003; Suthers, Hundhausen, & Girardeau, 2003). Compared

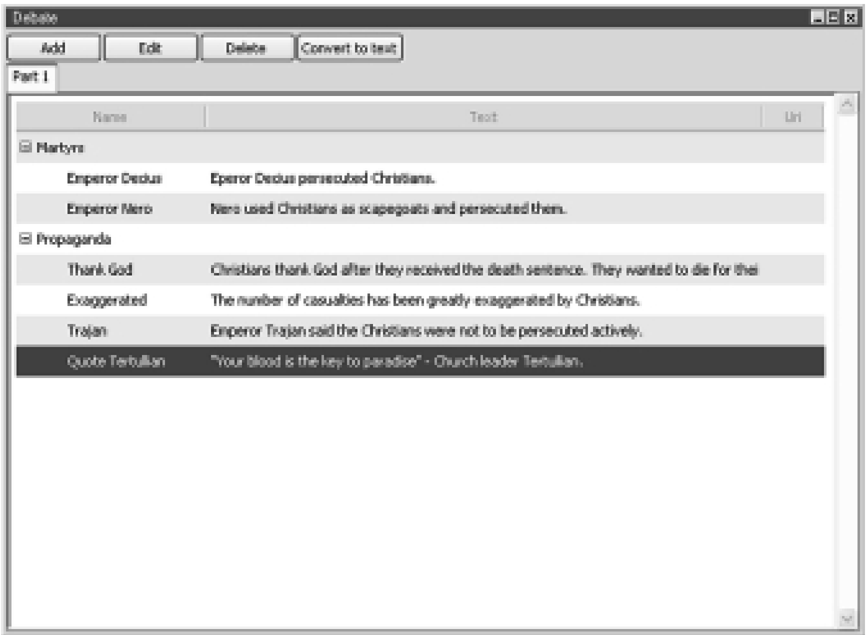


Figure 13.4 Screenshot of the textual debate-tool (translated from Dutch).

with the TD tool, the GD tool uses several visualisation techniques to make information salient and help students complete the representation more effectively and efficiently. For example, the GD tool discerns between arguments, supports and refutations. This feature may stimulate and guide students to find supporting and refuting information, and to formulate arguments, since it is immediately clear to them whether this information is present or not. Furthermore, the GD tool visualises how well positions are supported by arguments and supporting information. It is more difficult to infer this from the TD tool because no distinction is made between arguments, supports and refutations. Finally, the option to rate the quality of arguments, supports and refutations available in the GD tool may stimulate students to think about and discuss the importance of arguments and may help them to see which arguments are more important than others.

Method

Design

We used a single-factor, between-subjects design, with two different groups defined by the type of representation used: GD or TD tool. We randomly assigned three classes to the GD condition, and two classes to the TD condition. In total, 79 students in 24 groups worked in the GD condition, and 45 students in 15 groups formed the TD condition. Before the start of the study, students completed a 15-item knowledge pre-test. No differences were found between the two conditions with respect to their subject matter knowledge.

Participants

The participants were students from five different history classes from two secondary schools. The total sample consisted of 124 eleventh-grade students (55 male, 69 female), with an average age of 16.24 years ($SD = 0.57$). Their teachers randomly assigned them to different groups. Due to uneven class sizes and student drop-out, this resulted in 1 two-person group, 30 three-person groups and 8 four-person groups.

CSCL environment: VCRI

Students worked in a CSCL environment named Virtual Collaborative Research Institute (VCRI). VCRI is the successor to the TC3 used in the COSAR project. Students use the Chat tool to communicate synchronously with other group members. To read the description of their group task or to search and read relevant information, students can use the Sources tool. This tool lists a number of sources that can be opened and read from the screen. Group members use the Cowriter as a shared word processor. Using the Cowriter, group members can work simultaneously on different parts of their texts. VCRI contains several other tools designed to support the inquiry process.

Task

Students collaborated on an inquiry group task in the domain of history. Students were given 14 historical and contemporary information sources and were asked to explore and discuss the different sources with respect to the debate. Students were required to co-construct a representation of this debate in either the GD or TD tool. After they had completed their representation, they had to co-author an argumentative essay based on their findings.

Analyses

To determine whether groups in the GD condition constructed argumentative diagrams of higher quality than groups in the TD condition, we rated all the items placed in the tool on a 5-point scale (ranging from 0 to 4). Interrater reliability of the rating process was assessed by two independent coders. Cohen's κ was .69.

To determine whether groups in the GD condition wrote better essays than groups in the TD condition, we analysed the quality of these essays with respect to quality of grounds used for argumentation, and conceptual quality of the argumentation. The evidence provided by students to back up the claims and opinions in their texts formed the starting point for the analyses of grounds quality. Each text segment was judged on a 4-point scale, ranging from 0 to 3, in terms of how well and how elaborately it was supported by evidence or explanations (Clark et al., 2007). The conceptual adequacy of the arguments given by the students constituted the basis for the analyses of conceptual quality (Clark et al., 2007). Each segment was judged in terms of its conceptual correctness; thus segments containing, for example, flawed conclusions, misinterpretations or incorrect statements received lower scores for conceptual quality than segments containing no errors. Conceptual quality was also rated on a 4-point scale (0–3). Two independent judges assessed the quality of seven essays to establish the interrater reliability. Cohen's κ was .85 for grounds quality and .88 for conceptual quality.

To investigate whether students in the GD condition learned more than those in the TD condition, knowledge pre- and post-tests were developed. Both tests consisted of the same 15 multiple-choice items addressing topics covered in the inquiry group task.

Finally, to investigate the impact of representational guidance on the collaborative process, we used a coding scheme to analyse the online collaboration between group members (see Janssen, Erkens, & Kanselaar, 2007; Janssen et al., 2007). The online collaboration process was captured in log files, containing all actions performed by the students. This coding scheme consists of four main categories: task-related activities, regulation of task-related activities, social activities and regulation of social activities. Like the Task Act coding scheme for the COSAR study, this coding scheme distinguished between metacognitive and task-related activities, but also focused on the social aspect of collaboration. Interrater reliability of this coding scheme was determined by two independent coders. Cohen's κ was found to be .90.

Results

Quality of constructed argumentative diagrams

Groups in the GD condition made argumentative diagrams of significantly higher quality than groups in the TD condition, $t(37) = 3.90$, $p < .01$, $d = 1.28$.

Additionally, we correlated the number of items produced with the average quality of these items and found a significantly negative correlation, $r = -.66$, $p = .00$, meaning that when groups attempted to include a large number of items in their representations this had a negative effect on the quality of their representations.

Quality of essays

Analyses show that GD groups wrote significantly better essays than TD groups in terms of grounds quality and conceptual quality, $F(1,39) = 6.15$, $p < .01$, $\eta^2 = .15$ and $F(1,39) = 8.30$, $p < .01$, $\eta^2 = .19$, respectively.

Additional analyses showed that groups that received high scores for grounds quality also received high scores for conceptual quality ($r = .89$, $p < .01$).

Post-test performance

To determine the effect of condition, while controlling for prior knowledge, condition and pre-test score were added to a multilevel model. The first step in this analysis was to examine the results of a model without any independent variables, the so-called null model. This model contained two levels. Because students were nested in groups, the individual student constituted the lowest level, while the group constituted the highest level. Next, condition and pre-test score were added to the multilevel model. This model explained significantly more variance compared with the null model, $\chi^2 = 11.07$, $p < .01$.

Both pre-test performance and condition had a significant effect on students' post-test performance. As expected, a higher pre-test score contributed to a better post-test performance, $\beta = 0.28$, $p < .01$.

Furthermore, condition contributed significantly to post-test performance, indicating a positive effect of working with the GD tool, $\beta = 0.42$, $p < .05$.

Online collaboration

When we examined the collaboration protocols, we expected to find that GD groups would be less busy coordinating, regulating, and monitoring their task performance. This, however, was not the case. Students who worked with the GD tool were engaged in planning, monitoring and evaluating their task progress as much as students who worked with the TD tool. In sum, we did not find evidence that the GD tool facilitated the coordination of collaboration.

Conclusions for the CRoCiCL project

Based on the results, we conclude that the representational guidance offered by the GD tool has a positive effect on the quality of shared products students construct. First, the GD tool helps students construct better argumentative diagrams. Furthermore, the GD tool also helps group members write better essays. Finally, students who worked with the GD tool performed better on a knowledge post-test. These findings contrast with other studies that found limited effects of representational guidance (e.g., Suthers & Hundhausen, 2003; Toth, Suthers, & Lesgold, 2002; Van Drie et al., 2005). An explanation may lie in the representational guidance offered by the GD tool compared with the guidance offered by the tools in the work of other researchers. Our tool directs students' attention to the distinction between arguments, supports, and refutations, and this may stimulate students to incorporate these elements in their diagrams and essays. It has been argued that tools that support *linearisation*, that is the ordering of content and arguments into an essay, may be better supported by tools specifically designed to support the planning of the linear structure of essays (e.g., the Outline tool used in the COSAR project). Although the GD-tool was not specifically designed to support the process of linearisation, it may be the case that stimulating students to systematically address all arguments, supports and refutations of a position also facilitates the process of converting a representation into an essay.

Interestingly, representational guidance has been found to affect students' collaborative process in previous research. In this study, this result was not replicated. Our study offers no support for the expectation that representational guidance decreases group members' need to coordinate and regulate their task performance in the online discussions. Students could use the representations in both the GD tool and the TD tool to exchange information (Van Drie et al., 2005). Because both tools were shared, adding an element to the representation equates to exchanging information with group members. Thus, there might be less need to engage in extensive information exchange in the chat discussions, and the need to coordinate this process may also be diminished.

Although the GD tool seems to help students write better texts, it is noteworthy that the students evaluated the GD tool somewhat less positively than the TD tool.

Conclusions and discussion

In this chapter we compared two tools meant to help collaborating students understand and represent arguments and positions from different external sources within a societal or scientific debate. The Diagram in the COSAR project provided students with a graphical mapping tool in which they could collaboratively specify positions, pro-arguments, con-arguments, supports, refutations and conclusions, and could draw links or arrows between these elements. No restrictions on spatial structure and representation were made. Although the Diagram tool was highly valued by the students and did effect their collaborative deliberation, no effects or even negative effects were found on the quality of the argumentative essays they wrote. The Graphical Debate tool in the CRoCiCL project is also a tool for argument mapping by which students could collaboratively specify pro-arguments, con-arguments, supports and refutations with regard to two (given) positions. However, the spatial structure and representation of the connections between the elements were fixed and the relative argumentative strength of the positions was visualised. Furthermore, students could differentiate between the relative weights of supporting or refuting arguments. The Graphical Debate tool resulted in better-grounded and conceptually correct argumentative essays, and in learning effects on a knowledge post-test, but did not significantly affect the collaborative deliberation between the students and was not valued very highly.

We assume that these differences can be – at least partly – explained by the representational guidance both tools offer. The specificity of the weighting of the arguments available in the GD tool directs the students towards the relative strength of the arguments. The feedback given by the GD tool about the relative strength of positions and arguments and the complexity of the representation further heightens the representational guidance. Furthermore, the perspective of the debate as a sort of battlefield with advancing and retracting units supports the view of a debate as competing positions with justifications and supports for each side. In our view, the greater representational guidance offered by the GD tool may partly explain why students in the CRoCiCL project performed better than students working with the Diagram in the COSAR project. Further research on whether these changes in specificity and perspective actually can be observed in the understanding and thinking of students working with the Graphical Debate tool could support the representational guidance hypothesis.

It should be kept in mind that the Diagram and Graphical Debate tools were not compared directly in an experiment. As Table 13.2 shows, there were differences and similarities between both studies. The question as to whether the differences between both studies can account for the difference in effectiveness is difficult to answer. Looking at Table 13.2, the most important differences concerned the size of the groups, the duration of the project, the subject of the task and the operationalisation of the dependent variables. It is, of course, possible that the difference in, for example, group size (dyads for COSAR, mostly triads for CRoCiCL) influenced the effectiveness of the tools. However, it is regularly found that smaller groups perform better than larger groups (e.g., Schellens & Valcke, 2006), possibly due to the fact that in larger groups coordination is more difficult. This would mean that the working condition in the COSAR project would have been better – not worse.

In the COSAR project the students had to use the Diagram tool to organise positions and arguments found in internet and newspaper sources with regard to societal debates (organ donation and cloning). In contrast, in the CRoCiCL project the students had to use the GD tool to organise positions and arguments found in historical sources about a debate on early Christianity. Although the subject differed (social sciences and history), the task was similar (writing an argumentative essay) and meant for the same class level in secondary education. So it is unlikely that differences in subject can explain the differences in effect.

Further research is needed, however, to ascertain the precise impact of representational guidance on collaborative construction of argumentation maps.

References

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183–198.

Andriessen, J. E. B., Erkens, G., Overeem, E., & Jaspers, J. (1996). *Using complex information in argumentation for collaborative text production*. Paper presented at the Using complex information systems, Poitiers, France.

Andriessen, J. H. E. (2003). *Working with Groupware: Understanding and evaluating collaboration technology*. London: Springer.

Clark, D. B., Sampson, V., Weinberger, A., & Erkens, G. (2007). Analytic frameworks for assessing dialogic argumentation in online learning environments. *Educational Psychology Review, 19*, 343–374.

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

Dillenbourg, P. (1999). Introduction: What do you mean by 'Collaborative Learning'? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1–19). Amsterdam: Pergamon.

Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior, 21*, 463–486.

Janssen, J., Erkens, G., & Kanselaar, G. (2007). Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in Human Behavior, 23*, 1105–1125.

Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education, 49*, 1037–1065.

Scardamalia, M., Bereiter, C., & Lamon, M. (1994). The CSILE project: Trying to bring the classroom into world 3. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 201–229). Cambridge, MA: MIT Press.

Schellens, T., & Valcke, M. (2006). Fostering knowledge construction in university students through asynchronous discussion groups. *Computers & Education, 46*, 349–370.

Schwarz, B. B., Neuman, Y., Gil, J., & Ilya, M. (2003). Construction of collective and individual knowledge in argumentative activity. *Journal of the Learning Sciences, 12*, 219–256.

Suthers, D. D. (2001). Towards a systematic study of representational guidance for collaborative learning discourse. *Journal of Universal Computer Science, 7*, 254–277.

Suthers, D. D. (2003). Representational guidance for collaborative inquiry. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 27–46). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Suthers, D. D. (2006). Technology affordances for intersubjective meaning making. *International Journal of Computer Supported Collaborative Learning, 1*, 315–337.

Suthers, D. D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *Journal of the Learning Sciences, 12*, 183–218.

Suthers, D. D., Hundhausen, C. D., & Girardeau, L. E. (2003). Comparing the roles of representations in face-to-face and online computer supported collaborative learning. *Computers & Education, 41*, 335–351.

Suthers, D. D., Weiner, A., Connelly, J., & Paolucci, M. (1995, August). *Belvédère: Engaging students in critical discussion of science and public policy issues.* Paper presented at the AI-ED 95, The 7th World Conference on Artificial Intelligence and Education, Washington, DC.

Toth, E. E., Suthers, D. D., & Lesgold, A. M. (2002). Mapping to know: The effects of representational guidance and reflective assessment on scientific inquiry. *Science Education, 86*, 264–286.

Van Amelsvoort, M., Andriessen, J., & Kanselaar, G. (2007). Representational tools in computer-supported collaborative argumentation-based learning: How dyads work with constructed and inspected argumentative diagrams. *Journal of the Learning Sciences, 16*, 485–521.

Van Bruggen, J. M., Boshuizen, H. P., & Kirschner, P. A. (2003). A cognitive framework for cooperative problem solving with argument visualization. In P. A. Kirschner, S. J. Buckingham-Shum, & C. S. Carr (Eds.), *Visualizing argumentation: Software tools for collaborative and educational sense-making* (pp. 25–47). London: Springer.

Van Drie, J., Van Boxtel, C., Jaspers, J., & Kanselaar, G. (2005). Effects of representational guidance on domain specific reasoning in CSCL. *Computers in Human Behavior, 21*, 575–602.