Dirk Ifenthaler
Pablo Pirnay-Dummer
J. Michael Spector
Editors

Understanding Models for Learning and Instruction

Essays in Honor of Norbert M. Seel

Springer

Understanding Models for Learning and Instruction

Dirk Ifenthaler
Pablo Pirnay-Dummer
J. Michael Spector
(Editors)

Understanding Models for Learning and Instruction

Essays in Honor of Norbert M. Seel

Springer

Dirk Ifenthaler
Albert-Ludwigs University
Freiburg 79085
Germany
ifenthaler@ezw.uni-freiburg.de

Pablo Pirnay-Dummer
Albert-Ludwigs University
Freiburg 79085
Germany
Pablo.dummer@ezw.uni-freiburg.de

J. Michael Spector
Florida State University
Learning Systems Institute
Tallahassee, FL
USA
mspector@lsi.fsu.edu

Library of Congress Control Number: 2007940361

ISBN-13: 978-0-387-76897-7 e-ISBN-13: 978-0-387-76898-4

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

Table of Contents

Contributors vii

Preface xi

Prologue xv

Foundations and Methodologies of Mental Model Research 1
Pablo Pirnay-Dummer and Dirk Ifenthaler

1. The Fragmented Nature of Learning and Instruction 3
Michael J. Spector

2. Using an Evidence-Based Approach to Assess Mental Models ...23
Valerie J. Shute and Diego Zapata-Rivera

3. Practical Solutions for the Diagnosis of Progressing Mental Models 43
Dirk Ifenthaler

4. Constructing Mental Models during Learning from Science Text 63
M. Mikkilä-Erdmann, M. Penttinen, E. Anto and E. Olkinuora

5. Progress in the Diagnostics of Mental Models 81
Sabine Al-Diban

Model-Based Theories of Instruction 103
Dirk Ifenthaler and Pablo Pirnay-Dummer

6. Rendezvous with a Quantum of Learning 105
Pablo Pirnay-Dummer

7. Externalizing Mental Models with Mindtools 145
David Jonassen and Young Hoan Cho

8. Model-Centered Instruction, the Design and the Designer 161
Andrew S. Gibbons

9. Realizing Model-Based Instruction ………………………..…….. 175
Ulrike Hanke

Engineering the Model Environment …………………………... 187
Pablo Pirnay-Dummer and Dirk Ifenthaler

10. Developments in the Design of Instruction …………………… 189
Sanne Dijkstra and Henny Leemkuil

11. Bridging a Gap between Psychology and Instructional Practice ……………………………………... 211
Andrei I. Podolskij

12. Model-Based Learning Environments …………………………225
Gerhard Weber

13. Model-Centered Learning and Instructional Design …………..247
Patrick Blumschein

14. From Cognitive Theory to Instructional Practice: Technology and the Evolution of Anchored Instruction ………277
James W. Pellegrino and Sean Brophy

15. Complex Skills Development for Today's Workforce ………...305
Tristan E. Johnson and Wen-hai D. Huang

Index ………………………………………………………… 327

Contributors

Sabine Al-Diban is a qualified psychologist and Assistant Lecturer at the Institute of Educational Psychology, Department of Educational and Developmental Psychology at Dresden University of Technology, Germany. Her fields of research include the diagnosis of mental models and model-based intervention studies.

Erkki Anto is a Researcher at the Department of Teacher Education and in the Centre for Learning Research, University of Turku, Finland. His research interests include development of learning materials, eye-tracking, conceptual change during reading, and refutational text effect.

Patrick Blumschein is Assistant Professor at the Department of Educational Science at the Albert-Ludwigs-University of Freiburg, Germany. His research interests lie in the field of model-based learning and instruction and learning with new media. Actually he works on a theory of model-centered instructional design.

Sean Brophy is Assistant Professor of Engineering Education at Purdue University. He leads the Learning Sciences group for the VaNTH Engineering Research Center in Bioengineering Educational Technologies. Dr. Brophy is a learning scientist and engineer; part of his research focuses on the development of learners' ability to solve complex problems in engineering, mathematics and science contexts. Recently he has been exploring the interaction between innovative skills and rich content knowledge as an important component of developing *adaptive expertise.* This research aligns with Dr. Brophy's work to develop effective learning environments using challenge-based instruction.

Young Hoan Cho is a doctoral student in Learning Technologies at the University of Missouri. His research interests focus on problem solving in mathematics education.

Sanne Dijkstra is emeritus Professor of education at the Faculty of Behavioral Sciences of the University of Twente, The Netherlands. His areas of interest include industrial and educational psychology. Many of his publications are on instructional design and the effectiveness and efficiency of the models used. He has a strong interest in the relationship between the science of psychology and its application for instruction, especially in the acquisition of knowledge and skills by solving problems.

Andrew S. Gibbons is Chair and Professor of Instructional Psychology and Technology Department at Brigham Young University, Prove, UT. Dr. Gibbons research interests focus on the design of model-centered instruction and instruction simulation. Much of his work is focused on dimension of technological practice vs. scientific practice, tool principles for design and development, the architecture of instructional design, and layer theory of design.

Ulrike Hanke is Assistant Professor at the Department of Educational Science at the Albert-Ludwigs-University of Freiburg, Germany. Her research interests include model-based learning and instruction and further education.

Wenhao David Huang is Assistant Professor position in the Department of Human Resource Education at University of Illinois at Urbana-Champaign. His research interests include the manipulation of cognitive load in complex learning environments (e.g., games) and the refinement of serious gaming design process.

Dirk Ifenthaler is Assistant Professor at the Department of Educational Science at the Albert-Ludwigs-University of Freiburg, Germany. Dr. Ifenthaler's research interests focus on the learning-dependent progression of mental models, problem solving, decision making, situational awareness, and instructional science. He developed an automated and computer-based methodology for the analysis of graphical and natural language representations (SMD Technology). Additionally, he developed components of course management software and an educational simulation game (DIVOSA).

Tristan E. Johnson is Assistant Professor of Instructional Systems Design and Associate Director of Research at the Learning Systems Institute at Florida State University in Tallahassee, Florida. His research interests include assessment of team cognition, team-based learning, group learning processes measurements, and shared mental models measures.

David Jonassen is Distinguished Professor of Education at the University of Missouri where he teaches in the areas of Learning Technologies and Educational Psychology. He has published 30 books and numerous articles, papers, and reports on text design, task analysis, instructional design, computer-based learning, hypermedia, constructivist learning, cognitive tools, and problem solving. His current research focuses on the cognitive processes engaged by problem solving and models and methods for supporting those processes during learning.

Henny Leemkuil is Assistant Professor at the Department of Instructional Technology at the University of Twente, The Netherlands. His research focuses on the use of computer based applications in education, teacher training, and telematics. He worked on several EC sponsored projects like MODEM: Multi Media Optimisation and Demonstration for Education in Microelectronics, and KITS: Knowledge management Interactive Training System. The last 4 years his research focused on learner support in educational games. Currently he participates in the European APOSDLE project which develops an Advanced Process Oriented Self-Directed Learning Environment.

Joost Lowyck is emeritus Professor at the Centre for Instructional psychology and Technology, Catholic University of Leuven (K.U.Leuven), Belgium. His research topics were situated in the fields of educational psychology, educational technology and cognitive instructional design. More recent research focused on student's conceptions of learning environments.

Mirjamaija Mikkilä-Erdmann is Assistant Professor in the Department of Teacher Education and Centre for Learning Research in the University of Turku, Finland. Her research interests are conceptual change, text comprehension, development of learning materials and learning in higher education.

Erkki Olkinuora is Professor of Educational Sciences at the Department of Education and in the Centre for Learning Research, University of Turku, Finland. His current work focuses on learning and studying in higher education, learning material research, motivation and assessment.

James W. Pellegrino is Liberal Arts and Sciences Distinguished Professor of Cognitive Psychology and Distinguished Professor of Education at the University of Illinois at Chicago. He also serves as Co-director of UIC's interdisciplinary *Learning Sciences Research Institute*. Dr. Pellegrino's research and development interests focus on children's and adult's thinking and learning and the implications of cognitive research and theory for assessment and instructional practice. Much of his current work is focused on analyses of complex learning and instructional environments, including those incorporating powerful information technology tools, with the goal of better understanding the nature of student learning and the conditions that enhance deep understanding. A special concern of his research is the incorporation of effective formative assessment practices, assisted by technology, to maximize student learning and understanding.

Pablo Pirnay-Dummer is Assistant Professor at the Department of Educational Science at the Albert-Ludwigs-University of Freiburg, Germany. His research and publications are located in the area of cognition, learning and technology. He developed, implemented and validated the language oriented model assessment methodology MITOCAR (Model Inspection Trace of Concepts and Relations) which is built to assess, analyze and compare individual and group models of expertise. Pirnay-Dummer also developed the web based training software Empirix, including new approaches of automated evaluation and automated tasks synthesis algorithms.

Marjaana Penttinen is a Researcher at the Department of Teacher Education and in the Centre for Learning Research, at the University of Turku, Finland. Her research interests lie in conceptual change during reading, refutational text effect, eye-tracking, and learning strategies.

Andrei I. Podolskij is Professor of Developmental Psychology at the Lomonosov Moscow State University, Russia. His research interests include process of internalization, formation of mental activity, and investigation of developmental dimensions of learning and instruction.

Valerie J. Shute is an Associate Professor of Educational Psychology and Instructional/Learning Systems at Florida State University, Tallahassee, Florida, USA. Her research interests include cognitive modeling/diagnosis, adaptive instructional system design, formative and stealth assessments to support learning, and innovative evaluation methodologies.

J. Michael Spector is Professor of Educational Psychology and Learning Systems and Associate Director of the Learning Systems Institute at Florida State University. His research pertains to learning in complex domains and has focused in recent years on how to assess progress of learning with regard to ill-structured problem solving; he also conducts research on the design of technology facilitated learning environments.

Gerhard Weber is Professor of Pedagogical Psychology at the University of Education in Freiburg, Germany. His research interests include modeling of problem solving and case-based reasoning in learning situations, the development of adaptive learning environments, and the investigation of learning processes in computer-aided learning environments.

Diego Zapata-Rivera is a Research scientist in the Research & Development Division at Educational Testing Service in Princeton, NJ. His current research interests include the design, development, and evaluation of innovative assessment-based learning environments (ABLE), Evidence-Centered Design (ECD), Bayesian student modelling and the use of open student models as communication tools and as mechanisms to gather and share assessment information with students, teachers, and parents.

Preface

Layers and Transfers for Learning Theories and Applications

Mental Models continued to be a key subject in different fields of research for almost half a century. For good reason. Foundations from cognitive science, computer science, philosophy and cognitive psychology describe the workings of the human mind in tasks of deductive and inductive reasoning, especially for reasoning under uncertainty. They lead to theories of problem solving and to theories of learning and instruction which are both highly interdependent. Stepping into the design of both computer-based and face to face learning environments is obviously not very far since well founded theories on learning and instruction are ready for transfer into implementation and applications. By following these layers, we will always find two processes of transfer.

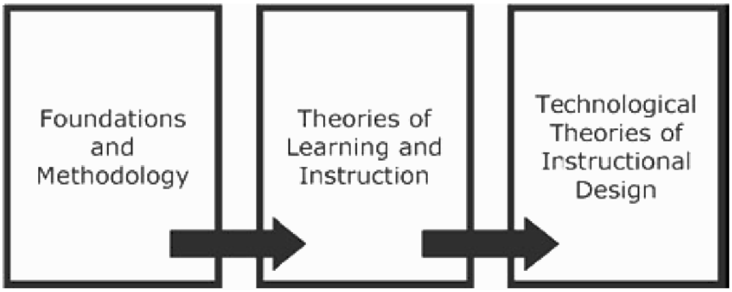

Fig. 1. Layers and transfers for learning theories and applications.

The first transfer is to be made from the theoretical foundations and methodology towards the theories of learning, instruction and problem solving, by taking into account the insights about reasoning and mental model building. This transfer is not unique for the field of mental models and should be applied for all theories of learning and instruction. In our case it leads to *Model-based Learning and Instruction* which is sometimes also referred to as *Model-centered* or *Model-oriented.*

Because learning environments are too complex to be described directly by the theories of foundation, the layer of learning theories is necessarily needed in between. In most cases insights on the foundation layer can not directly be translated and properly transferred into applications. Consequently the second transfer aims at the construction of well founded learning environments on the bases of the theories of learning and instruction. This transfer is the integral part of *Instructional Design*. However it is nearly impossible to conduct the second transfer without understanding the foundations and the first transfer process. And of course insights on learning environments can lead to new hypothesis for the foundations. Hence only the consolidated investigation of all three layers leads to a deep understanding of the field.

Norbert M. Seel's Lifework

Norbert M. Seel dedicated his lifework to all of the three fields and consequently contributed to them with great reputation on the levels of theory, psychometrics, empirical studies and instructional design. We invite international researchers to participate in an integral work on all the three domains of expertise and the corresponding transfers within the field of mental models. This work will contain the actual state of research, methodology and technology. The three parts for the outline of the work are:

- Foundations and Methodologies of Mental Model Research
- Model-based Theories of Instruction
- Engineering the Model Environment

Foundations and Methodologies of Mental Model Research

The first part of the work will focus on the foundations of mental models and on methodologies which allow to measure constructs of mental models and how to track changes to them over time. Backgrounds and interdisciplinary interdependences between cognitive science, computer science, philosophy and psychology will be thoroughly presented and discussed.

Model-based Theories of Instruction

The second part will consequently be about the transfers into theories of model-centered learning and instruction on the basis of the foundations and methodologies. It will show how the foundations can be generalized into larger settings of learning and instructions from a perspective of educational science and instructional design. This part will also show how the corresponding findings can be specified again for the referring theories.

Engineering the Model Environment

The third part will lead us to technological theories on applications for instructional design and educational engineering. Selected examples and empirical findings on learning environments based on theories of model-centered learning and instruction will show how state-of-the-art technologies can be build and evaluated.

The Book Project

The editorial committee has selected a wide range of internationally know distinguished researchers who present innovative work in the areas of educational psychology, instructional design, and the learning sciences. The audience for this volume includes professors, students and professional practitioners in the general area of educational psychology and instructional technology. Without the assistance of several specialists the editors would have been unable to prepare this volume for publication. They wish to thank Joost Lowyk and Sanne Dijkstra for their tremendous help with both reviewing the chapters and linguistic editing. Their thanks also go to David Heyde for proofread of several chapters and Andreas Lachner for preparing the chapters to meet the editorial style prescriptions.

Dirk Ifenthaler
Pablo Pirnay-Dummer
J. Michael Spector

Prologue

"Knowledge is no longer an immobile solid; it has been liquefied. It is actively moving in all the currents of society itself"
(John Dewey, The school and society)

The Quest to Understand Mental Life

Norbert M. Seel's focus on both understanding mental functioning of students and supporting the construction of mental models through instructional design, is highly relevant for the educational field in a knowledge society. No doubt that he has been influenced by his German roots in philosophy and more specifically in epistemology. It is an interesting issue to look back in history and to appreciate early efforts to gain insight in psychological phenomena. A shallow look at the work of Herbart (1776-1841) already reveals a systematic approach of scrutinizing mental life. In Herbart's view, initially chaotic presentations ('ideas') in the human mind can be modelled through the conscious process of apperception that links new ideas to former experiences. Herbart postulates that interactions between ideas can be expressed in mathematical formulas, which is perfectly in line with the scientific ambitions of his time. In this way, psychology is conceived of as a real 'science' with a specific object and a strict methodology. Education, then, aims at systematically directing the process of apperception using predefined, sequenced materials as sources to trigger expected experiences of children. Since the mind is filled with ideas, the kind of ideas and their sequence are important design components, which gave rise to the so-called 'formal steps' in Herbart's didactics.

Though times drastically changed, away from a mechanical view of the mind, there remains some similarity with Herbart regarding the ambitions of modern cognitive psychology. It is a premium merit of Norbert M. Seel to have undertaken a systematic search to understand mental models from an interdisciplinary angle. This is not evident given the Zeitgeist of behaviourism at times he started his academic career. Indeed, after a period of almost exclusive attention to behavioural aspects of human functioning in education and training during the sixties and seventies of the former century, renewed interest in what happens in the human mind became predominant in successive waves of cognitivism, constructivism and socio-constructivism. Norbert M. Seel acknowledged already in an early stage the importance of mental representations and mental models as parts of student's individual knowledge. Indeed, a central issue for education in a knowledge-intensive society is how humans represent information in order to use it in interaction with the world they are living in. Therefore, 'mental model' is a key construct to search for the way knowledge is organised, developed, and changed under pressure of a steadily developing self and world.

The Need for Strong Research

Once a research topic has more or less been stabilized -which is the case with 'mental models'- effort is invested in further clarification of the concept, development of valid and reliable tools, design, and implementation strategies. As has been acknowledged by the research community in educational psychology, a theoretical concept is only valid if open to observation and testing. Indeed, sciences develop along with the quality of the instrumentation. A recent example is the multi-level statistical approach that allows for measuring and interpreting complex, layered phenomena. In this book, vantage points for strong interdisciplinary research on mental models are presented. At first, the interrelationship between different knowledge domains needs clarification of the conceptual framework and the methodological toolset of each discipline in order to link that domains. Pirnay-Dummer reflects on the interdisciplinary links between research in cognitive psychology and theories of learning and instruction. He thoroughly analyzes methodological and logical traps in this complex interrelationship between domains. Spector rightly considers any interdisciplinary domain as a puzzle with missing pieces, pieces that often do not fit together or pieces from different puzzles that are quickly thrown together. More specifically, incompleteness of knowledge on learning brings about uncertainty in instructional design. In his opinion, Seel's research endeavours are important steps toward a more comprehensive theory of instruction.

Mental models are not so much viewed as static entities but they refer to knowledge development in individuals. Consequently, knowledge diagnostics of mental models are vital endeavours that call for the use and construction of standardized tests (Al-Diban). Since 'mental model' is not a 'static' but a dynamic construct, changes in mental model construction can only be captured through valid methods and techniques to measure these progressing mental models of change (Ifenthaler). In a similar vein, a possible powerful tool for measuring mental model construction and change is the use of eye-tracking methodology (Mikkilä-Erdman, Penttinen, Anto & Olkinuora). Mental models are complex, multifaceted and multi-layered, and no single form of assessment can represent all different kinds of knowledge. This is exemplified with Mindtools as a cognitive tool for externalization of mental models (Jonassen & Hoan Cho). Many authors contend that mental models are not an aim but a means that contributes to the quality of cognitive performance. A question in this regard is how measurement of knowledge can be related to measures of cognitive ability in order to identify individual's level of proficiency (Shute & Zapara-Rivera).

Designing Powerful Learning Environments

If one knows better than before the basic aspects and structure of mental models, this knowledge-base needs transformation into design principles for learning environments that are suited to help learners build knowledge or learn to know. This is an intensive and complex endeavour since descriptive knowledge ('knowing that') needs to be transformed into prescriptive knowledge ('knowing how'). This transformation is not an automatic or routine process given the many epistemological and empirical differences between descriptive and prescriptive knowledge. Consequently, there is a huge need for so-called transformation knowledge and skills. In terms of Podolskij, an utmost important aspect of scientific knowledge is its practical application. More specifically, this author refers to the Planned Stage-by-Stage Formation of Mental Actions (PSFMA) elaborated by Galperin in order to bridge the gulf between knowing that and knowing how.

One of the outcomes from a vast amount of research on mental models is that construction of mental models is a developmental activity. This clearly refers to learning processes that need explicit guidance through instruction. Instructional design with a focus on mental models necessarily has to adapt a model-centred perspective. Essential elements of that model are the learner's conditions, the domain-specific knowledge and the reflexive nature of constructing mental models. This calls for flexible and adaptive designs at the micro-, meso- and macro-level (Blumschein). Hanke links the constructive nature of schema building with instructional design. A mental model passes through several steps before it is learnt. She depicts different learning sub-processes that can be supported by five specific teaching interventions. Whereas some instructional designers clearly emphasize instructional design in terms of the representation and organisation of subject-matter content to enhance learning others opt for (complex) problem solving at the core of instructional design. This problem solving activity in designing mental models is multi-layered and depends on the mental model of the instructor, the experience used to communicate the mental model, and the evolving mental model of the learner in order to connect teaching and learning. In this multi-layered architecture, the content layer, being the form in which subject-matter is stored, plays an important role (Gibbons).

Powerful learning environments obviously also encompass rich technologies to support the construction of mental models. More specifically the relationship between theories of cognition and learning on the one hand and the design of instruction on the other can be furthered through the use of information and communication technologies, like simulations and games (Dijkstra & Leemkuil). Technology, however, is not limited to educational software products but, in line with the intelligent tutoring systems (ITS), adaptive web-based learning environments show

how user modelling is apt to facilitate learning (Weber). However, building mental models is not restricted to regular educational settings. The use of games can be conceived of as model-centred learning environments that not only serve traditional education, but the workforce as well in supporting complex skill development (Johnson & Huang).

In Honour of Norbert M. Seel

It may wonder that a topic like 'mental models' is well and alive despite its long history, its often different meaning, and the fact that it is a construct and not a reality. It stems from research that needs to define and confine possibly relevant aspects of human cognitive functioning in real settings. The reason for its success could be that this concept is basic to knowledge building and understanding how people construct knowledge in a complex society. Even a quick scan on Google reveals the existence of many specific websites that illustrate the penetration of the concept in research, consultancy, and daily life contexts.

In this book, different topics were passed in review. They mostly are at the spearhead of evolutions in this domain, contributing to a better conceptual understanding of the interdisciplinary phenomenon at hand, constructing new but valid research tools, and refining principles for advanced instructional design. An interesting reading activity could consist of scrutinizing the list of references after each chapter in order to grasp the real impact of Norbert Seel on researching mental models and designing suitable (technological) environments. The reader will recognize Seel's leading role in building a valid knowledge-base and will value as well the many research projects on mental models he launched in Germany and abroad. It is nice to observe how the different authors of the chapters are tackling each another topic and by doing so, produce a clear portrait of Norbert M. Seel.

Joost Lowyck

Foundations and Methodologies of Mental Model Research

Pablo Pirnay-Dummer and Dirk Ifenthaler

University of Freiburg, Germany

Filling the Granary of Fundamental Research

Strong theoretical foundations and precise methodology are always the one and only starting point for good research. Without sound foundations nothing follows, and thus a deep understanding of the theory of mental models and the methodology involved is mandatory for research on cognition and learning as well as for instructional design. This part contains contributions from J. Michael Spector, Val Shute & Diego Zapata-Rivera, Dirk Ifenthaler, M. Mikkilä-Erdmann, M. Penttinen, E. Anto & E. Olkinuora, and Sabine Al-Diban.

Spector begins with an overview of the field of learning sciences and instructional design from the critical and constructive perspective of the philosophy of science. This leads the focus directly to the inner relations between research questions and topic areas, between interest and methodology. After discussing the difficulties of the interaction, Spector presents Bob Dylan's Dream and other songs as a parable for the workings of theory and research, of design and application. He carefully shows possible ways out of the many possible illusions which can be created by fast plausibility and early consensus in the domain of cognition and learning and its applications, taking the reader to a path from pattern recognition all the way up to a profound revision of constructivism.

A new approach for tracking the flexible belief networks in mental models is shown by Shute and Zapata-Rivera. Following the understanding of mental models and their different forms of external representations, the authors show the differences between summative and formative assessment and between knowledge and belief. In addition to discussing concrete research, Shute and Zapata-Rivera show examples of how the differences apply to mental model assessment.

On the basis of classical model assessment, Ifenthaler shows new approaches which enable researchers to track change over time within individuals and groups. This is of particular interest in the field of learning and instruction because learning always evokes systematic change. Therefore, learning progress has to be evaluated through observation of the functions of change. For a complete understanding of the tasks involved, a methodological synopsis is also provided.

Ifenthaler shows and systematically compares selected measures from graph theory and analyses how they can be applied to the measurement of change. He also shows that these procedures can be implemented within fully automated computer programs which serve as assessment and analysis tools.

Mikkilä-Ermann, Penttinen, Anto, and Olkinoura focus on conceptual change within tasks of text understanding, thus building a bridge to the theories of learning and instruction. New methods which use eye-tracking within studies of immediate text understanding are mapped to cognitive conflicts to gain knowledge about the processes involved in conceptual change and understanding.

Al-Diban gives an overview of the classical approaches to mental model assessment. Consequently, she discusses several of the methodological problems of morphism. With an emphasis on causal models, Al-Diban shows best practice examples from research on mental models and discusses and evaluates the common strategies of data analysis and testing.

1. The Fragmented Nature of Learning and Instruction

Remarks on the Philosophy of Science, the Psychology of Learning and the Design of Instruction

Michael J. Spector

Florida State University, Tallahassee, Florida

Abstract: This chapter falls roughly into the intersection formed by the philosophy of science, cognitive psychology, and instructional design. I have drawn heavily on notes written for my students that have been read and commented on by Professor Seel in their earlier incarnations. Seel's work on the progressive development of mental models and the implications for the design of instruction have inspired many of my remarks. I regard this general domain of discourse as somewhat like a puzzle with missing pieces and pieces that should fit together well but often do not. It is almost as if the building blocks of instructional systems research were pieces from different puzzles thrown together hastily. The general thrust of my argument is that we do not yet have comprehensive and completely coherent accounts of how people learn and, as a consequence, we lack a complete theory of how best to design instruction and assess its effectiveness. Seel's research over the years represents important steps towards such a comprehensive theory of instruction.

Keywords: Constructionism; instructional design; instructional science; learning theory; mental model.

Remarks on Scientific Inquiry and Instructional Design Research

In a recent doctoral seminar in the Instructional Systems program at Florida State University, Professor Seel asked participants to indicate what each regarded as the single most important research question to be addressed in the next five years in the domain of instructional systems, broadly and loosely defined to include instructional analysis, design, development, evaluation, management and technology. Answers reflected topic areas rather than research questions. He then asked each student to indicate an appropriate research methodology to address [part or all of] the indicated question. This second request turned out to be problematic since topic areas rather than research questions had been provided by students.

I was struck by two things. First, the notions of *science* and *research* seemed to vary considerably from one person to another. Second, specific responses indicated a strong tendency to only consider those aspects of instructional systems with which a particular individual was engaged, with the implicit assumption being that what each was doing represented the most critical research issue in instructional systems.

What is science? What is the nature of scientific inquiry? What distinguishes scientific research from other forms of research? What do scientists do? There are many answers to such questions. They can be found in books on the philosophy of science and in nearly every introductory text to a particular scientific discipline. I found myself generating such questions during Professor Seel's seminar as various doctoral students provided their responses. I settled on a rough and ready representation of inquiry in physics as a starting point. For centuries, physicists have been asking such questions as these: (a) what kinds of things are there in the universe? and, (b) how do these different kinds of things affect each other? My first thought was that the basic questions within that discipline had remained fairly stable over the years; what have changed are the instruments and tools used to investigate various phenomena, which have led to new answers to basic questions and to improved understanding of the phenomena being investigated. Of course research methods and perspectives have also evolved, partly based on new answers to the basic questions. The basic research questions are basically unchanging. What changes are the tools used to investigate possible answers and the answers themselves. Moreover, interpretations of the basic questions may change considerably over the years; new interpretations of the basic questions might be regarded as representing a new approach, or possibly even a paradigm shift.

For example, Empedocles, (a pre-Socratic physicist who lived circa 492-432 BCE) believed that there were only four basic things – earth, air, fire and water – and that the physical world and our experiences could be completely accounted for in terms of these four elements. Aristotle further elaborated this view of matter and argued that all earthly substances contained mixtures of these four elements, with the particular distribution of the basic elements determining the nature and appearance of a particular object. For example, a rock contained much more earth than air, fire or water, according to Aristotle, which is presumably why rocks are hard, not readily combustible, and not easily transformed into liquid or gaseous forms. Aristotle then identified four kinds of causes: (a) material cause – the basic composition of an object; (b) formal cause – the inherent or underlying structure of a thing; (c) efficient cause – how the thing came to be in its current state; and (d) final cause – the purpose of an object.

We do not think about the world in the same way as did Empedocles or Aristotle. Physicists no longer accept their accounts of the physical world. In spite of dramatic advances in physics in the last two thousand years, much has not changed. What has not changed are the basic questions: What kinds of things exist and how do they interact? Scientists are still attempting to elaborate adequate answers to these basic questions. Modern answers are that there are some 118 or so elements – a few more than four – and these elements are comprised of more basic building blocks – with leptons, quarks, and bosons being the basic categories for these sub-atomic building blocks. Furthermore, a small number of forces have been proposed to explain interactions among these basic building blocks of the universe – gravity, electromagnetism, weak nuclear force and strong nuclear force.

Okay – I did not recall all of those details late at night after the seminar. I had to look up a few things. My basic line of thought, however, was that this framework might be applicable to Seel's questions. Imagine a door that has this question posted outside: What do instructional design researchers regard as the basic elements and what do they propose as the critical interactions among these elements? Shall I open this door? What might I find inside?

There is someone pulling on my elbow telling me not to waste my time opening that door. This person says that such an account applies only to the *hard* sciences – the physical sciences, such as astronomy, biology, chemistry, and physics. This person says that the *soft* sciences, which include the social sciences and what Herbert Simon (1996) called the sciences of the artificial, are fundamentally different. I understand those distinctions, I think, but there are some common concerns across all the sciences. Basically, what nearly all scientists want to know and understand is what exists – the building blocks – and how these things interact to bring about the things we observe, want to observe or would like to create. While causal interactions might be more difficult to establish in the social sciences, there is still strong interest in understanding, explaining, and predicting critical interactions. While the things that social scientists investigate might not be as precisely defined as those investigated by physical scientists, there is still strong interest in identifying the basic elements that explain what we have observed and are likely to observe in the future.

Perhaps this is a biased or naïve interpretation of science. Perhaps not. Nonetheless, I am going to push that door open and go looking for the basic elements and their interactions in the domain of instructional systems. What will I find?

What are the basic building blocks of an instructional system? What comes to mind immediately are students, instructors, things to be learned, and instructional resources. This might be an earth-air-fire-and-water kind of answer, though. Each of these elements might be further elaborated in terms of more discrete components which are more informative with regard to explaining interactions that are observed or desired.

What are the essential interactions or causal influences in an instructional system? Outcomes common to most instructional systems include improved understanding and performance with regard to some body of knowledge or set of skills. This implies that there should be reliable ways to assess relative levels of understanding and performance (relative to past performance or understanding or relative to a desired standard or goal). Other outcomes might be identified, and these might be further elaborated in terms of types of outcomes (e.g., affective, cognitive, psycho-motor or … there are many ways to cluster outcomes) and their relationship to other knowledge and skills.

Regardless of the sophistication and granularity of the components and interactions, we want to understand the various things that comprise an instructional system and how they are related, especially with regard to efficacy in achieving desired outcomes. Maybe. Well, I seem to recall Robert Gagné saying that our job as instructional designers was to help people learn better. What can we do at a systems level to fulfill that responsibility? How can we measure success? These and related questions represent the overarching areas of inquiry and scholarship in the general domain of instructional design and technology.

Lastly, there is the notion of research issues central to progress in a domain. The students who responded to Professor Seel each had a favorite area of inquiry. Why believe that one's favorite area of inquiry is critical to progress in instructional systems research, however? What evidence can one bring to bear to defend such a view? How might one identify critical areas of research inquiry?

One might think beyond oneself and beyond one's own training and set of predispositions. One might look at what distinguished researchers have said. The *Book of Problems* (see the 2002 events archive at www.learndev.org) would be a good starting point, I would think. That collection includes the contributions of 22 scholars and researchers who were asked by Jan Visser to describe what we do not know about human learning and to identify key unresolved problems. Contributors included a nobel prize winning physicist, a renowned biochemist, a neuroscientist, several sociologists and psychologists, an anthropologist, a number of educational researchers, and the odd philosopher. This collection is well worth a visit – what do such distinguished scholars believe is lacking in our knowledge of human learning? I leave the answer to this question as an exercise for the reader – an eminently worthwhile exercise, I believe.

I recall the advice I was given when searching for a dissertation topic by Professor Ed Allaire: Pick the central domain of inquiry within a discipline and then pick a central unresolved issue within that domain of inquiry that can sustain your interest. Of course there is much subjectivity in this – there will be different views about the centrality of domains and issues. I suspect, however, that a small number of alternatives can be identified. What might these alternatives be for instructional systems?

Addressing that last question is where I thought the discussion might have gone in Professsor Seel's seminar; at least that is where it was going in my mind. What are the central research issues in instructional design? I will suggest a few such issues later in this chapter.

What values are at the core of scientific inquiry? Are values relevant in our work? Of course they are. The starting point of scientific research is a desire to understand a phenomenon or situation or sequence of events. This implies that one admits to a state of relative ignorance: "I do not understand this." One might then say that humility ("I know that I do not know") is the starting point of every scientific inquiry or investigation. Humility would then be one of those core values in scientific inquiry. What do leading instructional systems researchers admit to not knowing or not understanding? That was the focus of Visser's *Book of Problems*. It would be interesting and revealing to find out to what extent academics and practitioners agreed with the things identified in the *Book of Problems*. I have not conducted a survey and am not positioned to answer, but I would propose such an exploratory investigation as relevant for our discipline.

By way of encouragement for others to explore this question, I shall provide a small sampling of contributions to the *Book of Problems*. John Shotter asked this in his contribution:

"To what extent is our living involvement in a whole situation necessary for us to get an *evaluative* grasp of the meaning for action of a small part of it – as when a music teacher points out a subtle matter of timing, or a painter a subtle change of hew, or a philosopher a subtle conceptual distinction, such as that between, say, a *mistake* and an *accident*?"

Vera John-Steiner asked: "How would our understanding of learning be transformed if its purpose were joint discovery and shared knowledge rather than competition and achievement?" Gavriel Salomon noted that "what we'd need to study is what makes socialization and acculturation so effective and how their 'active ingredients' could be incorporated into instruction." Leon Lederman suggested that we should figure out "how to construct a dossier of misconceptions, of 'natural' assumptions that must be viewed with suspicion."

Basarab Nicolescue posed this question: "If we distinguish three types of learning, the mental (cognitive), the feeling (affective) and the body (instinctive), how important are, for a given type of learning, the other two types?" Federico Mayor observed that we do not know much about "learning *to be*, to transform information into personal knowledge" even though we know a lot about learning *to know* and learning *to do*.

David Perkins posed four general questions about learning:

1. The Question of Mechanism - When we learn, in what form is that learning captured in us and our physical, social, and symbolic surround? – in the form of mental representations, the weightings of neural networks, conditioned reflexes, runnable mental models, priming or expectancy and different degrees of primability, distributed cognition, etc.? …
2. The Question of Difficulty - When learning is hard, what makes it hard? When learning is easy, what makes it easy? Answers would have to deal with the match between mechanism and the things to be learned. …

3. The Question of Design - What can we do to make learning something easier? This is the problem of instructional design taken broadly, not just for schools but for groups, teams, families, societies, even for immune systems and genetic codes. …

4. The Question of Worth - What's worth learning, for whom, for what purposes practical or ideological, at what cost? Do we find the guide to what's worth learning … in Adler's great books, in Dewey's pragmatism, in Socrates' insistence that we know our own ignorance, in more humble crafts and skills of the kitchen, the tailor's shop, the chemist's laboratory, the accountant's spreadsheet, in the ancient human modes of love, parenting, friendship, ownership, command, peace, war? …"

In order to conduct sustained scientific investigation, one must be open to alternative explanations – this (what I or someone else has proposed) is one possible explanation; perhaps there are other explanations. Open-mindedness is then a second important value. The inability to imagine alternative explanations does not mean that alternative explanations do not exist. It would be a remarkable coincidence if the limits of reality happened to coincide with the limits of one's imagination. Alternative explanations always exist (this is a remark about the logic of scientific explanations). Humility is the starting point, and openness to alternative explanations is required for sustained inquiry.

Perhaps none of Seel's doctoral students mentioned such things because they are so obvious. I find myself requiring such reminders, though. In answer to Seel's question about important research questions in instructional systems for the near future, I offer this: How can we reliably determine which interventions intended to help improve understanding of complex and dynamic systems are effective, to what extent they are effective, with whom they are effective and why? By way of clarification, I offer a few additional remarks. Complex and dynamic systems, be they natural or artificial, create challenging problem-solving and decision-making situations for humans. In such systems, many problems arise that are not especially well-structured; there may be incomplete information about critical aspects or goals, there may not be one standard or correct solution, there might be multiple approaches to the problem, and there might even be alternative interpretations of the problem situation itself. Complex and dynamic systems are pervasive. Examples include economic policy development, engineering design, environmental planning, instructional design, medical diagnosis, and many more. There are university curricula built around such problem solving areas. How might one go about determining whether and to what extent various curricula that support development of knowledge and skill in solving complex problems are effective? How might the findings of such an investigation be used to improve human understanding and performance in complex problem solving domains? I admit to not knowing the answers to these questions, but I am engaged in trying to find reasonable answers, as are Professor Seel and others.

The Fragmentary Nature of Psychology

From Bob Dylan's Dream:

As easy it was to tell black from white
It was all that easy to tell wrong from right
An' our choices they was few so the thoughts never hit
That the one road we traveled would ever shatter or split.

One of my graduate students asked about the fragmentary nature of psychology. This question implies that psychology is incomplete or disconnected. Perhaps the request is for an explanation why psychology is incomplete or disconnected. I am not sure. Perhaps I am incomplete and disconnected; I am sure this is often the case.

Suppose we agree for the sake of this discussion that psychology is the disciplined investigation of human thought and behavior. The aim of psychology is to provide a general account of the processes that underlie observable behavior and reported thought processes. As is true in other scientific enterprises, the desired general account will consist of causal factors and underlying mechanisms that explain what has been observed and reported and that predict what is likely to be observed and reported.

Where shall we start? Perhaps we should begin with a familiar phenomenon, such as confusing the names of two people. Suppose this phenomenon is common to nearly everyone – it represents the one road we are now all traveling together. Suppose further that we are all able to easily recognize this phenomenon and know when we have in fact confused one person's name with that of another person – it is easy to tell black from white – at least at the outset. Someone said that in the beginning there was chaos and confusion. Or was it chaos and the void? In any event, this beginning is not like that other one that happened a long time ago.

Now, we are underway. The journey has begun. Let us begin by collecting explanations for this phenomenon. One person says that he confuses X's name with that of Y because X resembles Y. Another person says that she confuses the names of X and Y because she met them both at the same time. Still another claims that the cause for mixing up the names is that the situations in which each person was first encountered were remarkably similar, although X and Y were met separately at different times and in different parts of the world by that person.

We already have three different accounts and we have barely begun. One explanation focuses on physical resemblance and implies a recall mechanism that assumes a search for an association between two kinds of mental representations – one textual (the name) and one visual (the person). Another explanation focuses on storage and retrieval cues, implying that the circumstances in which a person's name is first learned are stored along with that person's name and then used at least sometimes in recalling that person's name. The third explanation also focuses on storage and retrieval mechanisms and also implies links between a retrieval cue (one kind of mental object) and a name (another kind of mental object).

Just as these three different explanations were beginning to coalesce, along comes another person who says that he confuses two people because he likes them both a lot and both remind him of another person. Emotions may also play a role in cognition, at least on some occasions for some people – namely those with emotions. I am heartless and unable to understand this person, so I continue on my way.

Just as I realized that fact about myself, along came yet another person who said that she confuses X and Y because their names are syntactically similar. Let us not overlook the mediating influences of language. Language pervades so much of what we do and learn. How can we properly account for the role that language plays in learning? What tales these twisted tongues will tell. One might even sense a change of language use and tone within this very chapter – even our use of language is fragmented. Back to our investigation of a *simple* phenomenon.

Then a person driving a convertible drove up, stopped and told me that the reason that I confused those two names was that I was in love with them both. Or perhaps I was in love with my deceased mother. Or perhaps with someone else's deceased father. Or just lusting after them for no particular reason. Beware people driving convertibles.

Next there came along a large truck – a moving van, in fact. The driver stopped next to me, unloaded a couch, and invited me to sit down and tell her my troubles. I began to cry realizing that there were just too many possible explanations for this apparently simple phenomenon. She consoled me and gave me a lollipop.

What a strange beginning, I thought. When I looked around after the convertible and truck had driven off, I found myself all alone. Those with whom I had begun this quest were no longer in sight. I suppose they had followed a different bend in the road. Perhaps they escaped to Canada. Then I began to think about that other beginning, the one involving chaos and confusion. I concluded that not much had changed in all the intervening years.

Is it no wonder that psychology is incomplete and disconnected? Humans are complex creatures. Consciousness is especially complex. In the *Tractatus Logico-Philosophicus* Wittgenstein said that we picture facts to ourselves. Is it not remarkable that we are able to do that and then to talk about those pictures with others? We picture facts to ourselves. We also are able to picture to ourselves things that are not facts. Misperceptions, misconceptions, and misinformation account for many of these misleading internal pictures. We picture facts to ourselves. We cannot stop picturing facts to ourselves. This is a natural and ongoing process. Some are apparently able to improve the quality of these internal pictures, but such improvements are difficult to assess because these internal pictures are not directly available for public scrutiny; we cannot even examine our own internal pictures – mental models and schema, if you like. We construct internal representations to make sense of our experiences; these internal representations are hidden from view but affect what we come to believe and how effective and efficient we are able to learn. Is not a critical issue for instructional research the investigation of these internal representations and their role in learning? What interactions might exist between external representations provided by an instructor or a co-learner or oneself and these internal representations? What kinds of internal representations do we create? When? Why? What kinds of external representations are likely to engender more effective internal representations and result in improved learning (and, of course, with whom, when, and why)? Professor Seel is exploring this question, as are others, including myself. Given what we know and do not know about the human mind and its development, answers are quite fragmented at this point in time.

Recognizing Patterns

One of the most remarkable statements I have encountered is also one of the simplest. It is this: "Wir machen uns Bilder der Tatsachen," which has been translated from the German as "We picture facts to ourselves" (Wittgenstein, 1961). I mentioned this in the previous section and implied that these internal pictures are one of the basic building blocks of learning and instruction. Is this not a remarkable statement – we picture facts to ourselves? Is it not a remarkable ability? To highlight why this is so remarkable, perhaps a short philosophical sojourn is in order. Such a sojourn is consistent with my conception of *philosophy as thought in slow motion*.

We picture facts to ourselves. Or, more literally, we make for ourselves pictures of actualities. Making such internal pictures is not at all like drawing a sketch of something. We can observe a person making a sketch. Many people have drawn sketches while sitting in philosophy classes listening to boring lectures on epistemology. Some sketches are made more or less thoughtlessly, but many are constructed intentionally to represent something. Sketches are typically created on flat surfaces in one or more colors. They may or may not bear some resemblance to an object in the surroundings of the person making the sketch. So, we make sketches – that is not so remarkable, although drawing is a skill that can take years to master. This other ability, though, is something else. We make pictures to ourselves. How long did it take to learn to draw internal pictures? No time at all, although perhaps one can improve with practice – a serious matter well worth investigating, this last claim. Where is the hand that draws the picture? Oh, no hand is involved with internal pictures. Where is the flat surface on which the image is drawn? Oh, there is no such surface. Where is the crowd that gathers to watch the image being drawn? Oh, no one can observe this process, not even the person making the internal picture. Oh.

Is this statement – *we picture facts to ourselves* – a metaphorical remark, then? What can it mean? It seems to be a critical claim in Wittgenstein's *Tractatus Logico-Philosophicus*, and it is closely related to concepts fundamental to cognitive science. It does seem worth exploring a bit more, especially since most people seem to accept it as obvious and non-problematic. Only the odd beast – the philosopher – draws attention to such apparently simple claims. Somehow or other we seem to build up an understanding of our surroundings, including other people and unusual phenomena we encounter. Moreover, we have a nearly irresistible urge to talk about our experiences and our surroundings, especially odd people and unusual phenomena. How is this possible?

Apparently, humans excel at recognizing particular images and patterns. Infants seem to naturally realize the significance of faces and quickly learn to recognize the faces of their mothers, for example. Applying a simple pattern matching algorithm to this ability only makes the ability seem more mysterious. The appearance of a face changes often and for many reasons, including changes in mood, hair styling, lighting and more. Moreover, the person may be moving, and the angle at which a face is viewed is rarely the same. How does an infant come to recognize a particular face as the same one viewed on previous occasions? Indeed, how do we come to recognize things at all? Something more than a physical perceptual mechanism must be involved. Memory is surely involved. There must also be something – some kind of process – that fills in missing parts of an image or suggests that an image is sufficiently similar to one that is recalled to regard it as representing the same thing. This process suggests a kind of pattern matching logic. *We picture facts to ourselves*. Babies create internal representations of their mothers.

The logic of this process can quickly escape our control. We start with one external reality (a mother's face) and one internal reality (a baby's internally constructed image of that face). When considering how the infant recognizes that face as its mother, a third reality intrudes – a recalled image. When making the judgment that this external reality is one's mother, one produces an internal image, recalls prior images one associates with one's mother, decides that the internally constructed image is sufficiently similar to the accepted mother-images to be part of that collection, and finally concludes that the external reality is indeed one's mother – presumably it is also sufficiently similar to the internally constructed image. Whew. All that pattern matching is enough to make one cry.

It is a good thing that babies are not logical – they would never recognize their mothers if they were. Consider this. For X (the internally constructed image of mother) to be judged as *truly representative* of Y (the mother), there must exist a third thing Z (a previously constructed and stored internal representation of mother) that is *truly representative* of Y. But how did the infant come to the conclusion that Z was truly representative of mother? Hmmm. Presumably, on a prior encounter with the mother Y (which we are allowing to be the same Y for the sake of simplicity), the infant constructed an internal image Z (which we are allowing to be the same as the one recalled earlier for the sake of simplicity), compared it with another stored image (let's call it W – why ever not?) of mother that had been previously accepted as truly representative, realized that this one was sufficiently similar to belong to that collection, and thereby concluded that it was also truly representative of mother. Well, this *third mother regress* cannot be infinite since the baby was born at some point and only then started constructing internal pictures and collecting images.

Infinite regress arguments seem to lead nowhere, which is where we were headed with that analysis. It is somewhat reminiscent of Aristotle's unmoved mover problem. The problem with infinite causal sequences is getting them started (or stopped, depending on your point of view). *That* problem. All events have causes. Event Z is caused by Y which was caused by X which was caused by … and so on back to a much prior event situation and on to an endless sequence of prior events and causes. The imagination handles that logic about as well as the pattern matching regress and usually concludes that there was a big bang or some other bursting out party to get things going. Happy birthday.

So, how does the baby come to recognize mother? The ability to recognize faces takes some *reflection* – or at least may at one time have taken some time for reflection and comparison and recall. In any case, understanding how we recognize faces requires some serious reflection on various human abilities, characteristics and tendencies. To emphasize the significance of reflection, I pause for this moment of reflection:

> And indeed there will be time
> For the yellow smoke that slides along the street,
> Rubbing its back upon the window-panes;
> There will be time, there will be time
> To prepare a face to meet the faces that you meet;
> There will be time to murder and create,
> And time for all the works and days of hands
> That lift and drop a question on your plate;
> Time for you and time for me,
> And time yet for a hundred indecisions,
> And for a hundred visions and revisions,
> Before the taking of a toast and tea.
>
> From T. S. Eliot's "The Love Song of J. Alfred Prufrock"

What was that question dropped on our plate? What was it that we were trying to understand? How we are able to recognize faces. Well, visual cues typically come bundled with other perceptual cues. Perhaps babies also smell their mothers; the reverse is certainly the case. There is the famous biblical story of Jacob and Rebecca deceiving the blind Isaac using the *smell* of Esau's clothes and the *feeling* of an arm disguised using goat skin to feel hairy and rough, like Esau's. While we might use several senses together to identify objects, we can nonetheless be misled. Descartes makes much of this possibility in his *Mediations on First Philosophy* – never trust a source that has misled you even once. If we followed such advice, we would suffer the same fate as those babies who never learn to recognize their mothers due to the third mother regress. Perhaps fewer politicians would get re-elected, though, if we decided not to trust a source that had even once misled us. There is almost always a silver lining to inquiry and investigation.

Well, it seems there is a need for some kind of explanation with regard to how we manage to build up an understanding of our surroundings and experiences out of these internal representations we construct. *We picture facts to ourselves*. Is it not remarkable that we are able to make sense of these pictures? Even more remarkable is how quickly an infant is able to automatically recognize a face, even one of those *early-morning-after-a-very-late-night* faces. Perhaps it takes an infant two or three times to develop an association of a face, and perhaps also a smell, a sound and a touch, with mother, milk and such. Quickly the recognition process becomes highly automated and only a momentary glimpse is required. Once the pattern recognition process is established, it is highly resilient. What might this suggest about human reasoning?

The notion that we create internal representations to make sense of our experiences and surroundings is a fundamental tenet of a naturalistic epistemology that has roots in the philosophical works of David Hume and Immanuel Kant. In the 20th century, naturalistic epistemology became the basis for socio-constructivist approaches to education, although there is much confused discourse pertaining to constructivism in learning and instruction. As a tenet within naturalistic epistemology, the claim that we picture facts to ourselves is not a prescriptive claim – it does not tell the baby that it *should* create an internal picture of that face to see if it is mother and likely to bring comfort. We simply do, it seems, create internal representations of things. We do so naturally and without any prompting or guidance. We cannot stop creating these internal representations, and many will argue that the process continues even while we are sleeping. Ah, "to sleep, perchance to dream" … pass the rubbing alcohol … all this talk about epistemology is making my brain muscles ache.

> Half of the people can be part right all of the time,
> Some of the people can be all right part of the time.
> But all the people can't be all right all the time
> I think Abraham Lincoln said that.
> "I'll let you be in my dreams if I can be in yours,"
> I said that.
>
> From Bob Dylan's "Talking World War III Blues"

We create internal representations to make sense of our experiences, and then we use these representations to guide our actions and to structure our discussions with others. We realize that on occasion others may be viewing the same situation and engaging in a similar process of creating representations and sense making. What about these others? Might some of them be constructing internal representations that are sufficiently similar to mine to guide them to similar actions and conclusions? Half of them might say and do things similar to those that I would say or do. Hmmm. I suppose we need another distinction – that between internal and external representations, to which I have alluded already on several occasions. *We picture facts to ourselves*. These internal pictures are private and cannot be directly inspected or shared. We also create external representations of these internal pictures that are public and can be shared. These artifacts become part of the observable world and might also be worthy of investigation and consideration.

Ouch. Occam's razor just got stuck in my beard. Am I multiplying entities beyond necessity? We began with external realities (mothers) and internal representations (constructed internal images of mothers). We added more internal things – things stored in and recalled from memory. Now we are adding more external things – human constructed representations to be used in talking about other external things as well as about those non-shareable, unobservable internal representations. Okay, it is a lot to keep up with, but perhaps it will explain how it is that we are able to build up an understanding of our experiences and surroundings and share that understanding with others. Perhaps. Perhaps these things are more of those basic building blocks of learning and instruction mentioned earlier.

Deconstructing Dylan and Reconstructing Constructivism

The title of this section is intended to get you to read this first introductory sentence, which is intended to get you to read the next. The previous sentence is not true. One can be paradoxical without contradicting oneself.

The thread that ties Gödel, Escher and Bach together for Hofstadter (1979) is self-reference. Escher creates paradoxical drawings by having one element of the picture connected with another which is connected to another that eventually connects back to the original – a visual form of self-reference. Bach's musical compositions often follow a pattern of self-reference such that a specific operation is performed on the previous part of the composition to create the next; as this pattern of modifying the previous part to get the next in a particular way is repeated, the composition returns surprisingly to its original form. Paradoxical outcomes may result from self-reference. Hofstadter uses self-reference to elaborate and explain Gödel's incompleteness theorem. Stated in non-technical terms, the incompleteness theorem postulates that in a formal system sufficiently powerful to generate all of the truths of integer arithmetic there are well-formed statements that are known to be true but which cannot be proven or refuted in that system. Even our most solid mathematical knowledge is inherently incomplete.

One such example of an *unprovable* truth is the claim that 'MU' is not a word in the formal language I call 'M-I-U'. The next section contains Hofstadter's (1979) MU puzzle along with an external representation of a problem-solving sequence followed by a similar puzzle that might be used to test transfer of learning.

The MU Puzzle

Given:

A) An artificial language with three letters in its alphabet: 'M', 'U', and 'I'

B) All strings or words in this language begin with and contain exactly one 'M'

C) Four rules to apply to existing strings of letters to generate new strings:

1) A 'U' can be added to any string that ends in 'I' (MxI → MxIU)
2) The entire string after the initial 'M' and be replicated and appended to generate a new string (Mx → Mxx)
3) Two consecutive 'U's can be dropped (MxUUy → Mxy)

4) A 'U' can be substituted for three consecutive 'I's (MxIIIy → MxUy)

D) One initial string – namely 'MI'

Derive: 'MU'

Notes: The variables 'x' and 'y' represent any sequence of letters in M-I-U including the null set. In the derivation, show the source string, the rule, and the resultant string. The search for a derivation may be represented as a tree structure as shown in Figure 1.

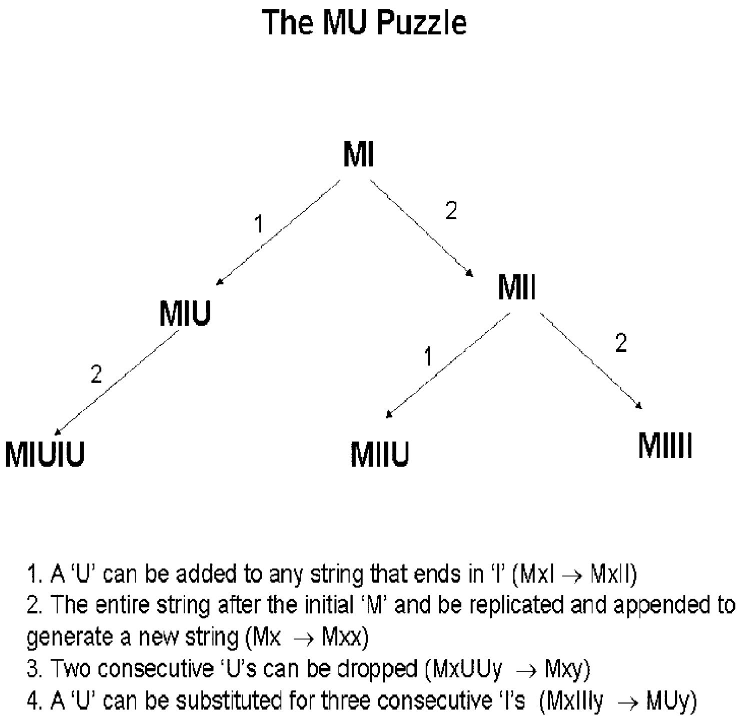

Fig. 1. The beginning of a derivation of MU.

A Different Puzzle?

Given:

- An artificial language with exactly five letters: M, E, Y, O, U
- The initial sequence: M E

- And six transformation rules (x and y are variables that stand for any sequence of letters including the null set)

1. If the sequence ends in E, a U can be added (M x E → M x E U)
2. The entire sequence following the initial letter can be replicated (M x → M x x; Y x → Y x x)
3. Two consecutive U's can be dropped (M x U U y → M x y ; Y x U U → Y x y)
4. Three consecutive E's can be changed into an U (M x E E E y → M x U y; Y x E E E → Y x U y)
5. An M can be changed into a Y (M x → Y x)
6. An O can be added after a Y (Y x → Y O x)

Derive: Y O U

You should know that 'MU' is not a word in the M-I-U language, but there is nothing in that language itself which can demonstrate that 'MU' is not among the infinite words that are in the language. A computer programmed to generate words in the 'M-I-U' language according to the four MIU rules and check at each step to see if 'MU' was generated would never terminate – that is an example of the halting problem (it might be better named the non-halting problem). It would be perverse to engage computers in such meaningless tasks – we have no such qualms with regard to people unfortunately. Is 'YOU' a word in the 'M-E-Y-O-U' language?

Language, Learning and Constructivism

I introduced those problems to get you thinking about reasoning that was difficult to explain in terms of prior experience and to help you appreciate the value of collaborative problem solving. I have discovered over the years that students reach the correct solution to those puzzles more often and more quickly when working in small groups of two or three. I am reminded of what Bob Dylan said in "Talking World War III Blues" and cited earlier:

> Half of the people can be part right all of the time,
> Some of the people can be all right part of the time.
> But all the people can't be all right all the time …

Of course I object to 'all' in the claim that "all the people can't be all right all the time," but this is a nice transition from the value of collaborative problem solving to constructivism. Beware, I am beginning to apply some tension to the rope around your neck. This may hurt.

As Vygotsky (1978) and many others have noted, we do and can learn *with* as well as *from* others. Much learning involves language, although one can cite instances of learning that do not involve language. Language is socially constructed. Using language requires some general agreement on rules, recognition that certain statements are acceptable or not, and much more. Wittgenstein (1963) argues that there can be no such thing as a private language. The nature of language requires recognizable rules and shareable interpretations of utterances. Language is learned through interaction with others. Language mediates much of our thinking, and language is inherently a socially-based enterprise. For Wittgenstein, then, the meanings of words are dependent on their use in various social situations. One may use the phrase 'it's raining' for example to convey to a second person that it is time to pack the car and leave town due to an impending hurricane – this is not a weather update as the words stripped from the situation might suggest. It is conceivable that one could use that same phrase in many other ways. What is critical is the shared interpretation – the use within a particular community of language users – a language game. You may resist this thought so give it some time to settle in before proceeding to the next paragraph. There cannot be a completely private language. Repeat this 30 times and go back to the beginning of this note, being careful to follow all directions to the reader, of course.

A constructivist may claim that each individual perceives a reality and constructs a private, internal interpretation of that reality (Glasersfeld, 1987). This idea is not new. Immanuel Kant's *Critique of Pure Reason* was published in 1781. In that book Kant claimed that individuals construct interpretations of the world based on their experience. Kant went on to claim that language was in fact possible because people were inclined to interpret things, and then talk about them, in terms of commonly accepted and recognizable categories, such as space, time and causality. Where these *a priori* categories came from was not clear in Kant's account. While individuals might differ with regard to what item was in a particular place at a particular time and how it may have influenced other things and events, the discourse about those differences is possible just because we all recognize the basic categories (space, time, causality) and we construct similar kinds of interpretations. This is an important point. What we say about our experiences is mediated by language and language is a social enterprise.

Wittgenstein (1961) put it this way: *we picture facts to ourselves*. Is this not a marvelous ability – to picture facts to oneself – and to then represent one's interpretation of those facts to others – is it not amazing that we have such capabilities? (One is allowed to be repetitive when in the presence of such marvelous wonders; the repetition in this chapter is intentional; perhaps it is the third repetition that produces the desired effect.) Wittgenstein, like Kant, did not go on to conclude that the meaning was also internal and private. For Wittgenstein, meaning occurs within the context of language and use within a community of speakers.

Some constructivists, however, go on to say that meaning-making is a completely private enterprise (Glasersfeld, 1987). Each person interprets reality and no interpretation is better or more accurate or more acceptable than any other. Moreover, one person cannot ever see or experience or judge or evaluate what another person has constructed. Beware. The rope has now been sufficiently tightened.

Radical constructivism ends in *epistemological solipsism* [a new term you can use to recover the good favor of those you got to try to solve the MU puzzle for or with you]: the only world that I can know about is my world – that is what the radical constructivist concludes. This is a form of disguised nonsense. The epistemological solipsist may go on to conclude, even more radically, that the only world that exists is his or her world [actually, there could be no 'his or her' on this view as there is only one world – my world – ruling out the other alternative and much more]. Ontological solipsism ("I am the world" or "Only my world exists") may be more obviously confused than the epistemological variety, but neither is defensible. Why? Because to defend or even make such claims one is required to make use of language, and language can only be learned from others and used in ways that make sense, at least on many occasions, to others. The radical constructivist has taken away the possibility of language and meaningful communication. My sense is that it is best to allow such persons to continue in their silent worlds; time is better spent talking with the other half of the people who might be almost right part of the time.

Okay, there are two points to be made at this interlude. First, constructivism refers to a particular kind of epistemology – it falls into the general category of naturalistic epistemologies in contrast with deductive epistemologies (a philosophical distinction you need not remember except for the test to be administered to all those who successfully complete this chapter following all the directions provided to readers). There are two versions of constructivism – radical constructivism and social constructivism, as others have mentioned; the latter is the kind that Vygotsky and Wittgenstein and many others, including Immanuel Kant and Robert Gagné would recognize as legitimate and meaningful. Yes, you heard me correctly. Gagné regarded himself as a constructivist in the sense just explained; he recognized constructivism as a reasonable epistemological perspective and accepted that position readily. He detested the discourse with regard to constructivism in the instructional design and educational technology communities, however, and refused to dignify the positions of many so-called constructivists with comments. Gagné believed that the so-called instructivist-constructivist distinction was confused and illegitimate.

The second point I wish to make is more general – one ought to pay careful attention to the implications of what one says. In deliberations about anchored instruction, situated cognition, and cognitive apprenticeship. there are often stimulating and animated discussions about the design of instruction and transfer of learning. Consider the following statements for example:

1. Learning activities ought to be designed so as to be meaningful.
2. Learning activities ought to be designed so as to be authentic.

3. The first can be linked in a reasonably clear chain of reasoning, from a naturalistic epistemology such as social constructivism (i.e., a philosophical foundation on which there is very broad general agreement) to a cognitive psychology of learning and finally to principles to guide the design of instruction. Design activities that will engage learners in meaningful ways such that learners will be interested, able to activate relevant prior knowledge, interpret new phenomena within the context of things already understood, explain and anticipate likely outcomes of these phenomena, and so on. Of course there can be measurable outcomes that will help instructors and designers determine the efficacy of efforts to support learning. In short, the first statement can be interpreted in a way that makes it a testable hypothesis. It would not be surprising to find empirical evidence to support the first claim since it is now quite well established that the more time that learners spend on tasks the more likely it is that they will acquire intended competencies. Meaningful learning activities tend to keep learners engaged, resulting in more time spent on task. Perhaps more importantly, what learners regard as meaningful are things that can in some way be related (a) to things they already understand, and (b) to things they want to understand.
4. The second statement goes beyond this claim by suggesting that activities that are meaningful will be authentic in the sense that they represent actual tasks to be encountered outside the learning situation. While this may be a testable hypothesis, it is rarely tested. If it were tested, it might prove quite limiting. Moreover, it does not address learning hierarchies – the important notion of mastering simpler tasks and problem solving activities and then going on to more complex and challenging tasks and problems. The second claim buries the notion of being related to something already understood. Rather, it becomes a mantra for those who advocate a particular approach to learning and instruction – namely use only actual tasks in learning activities. One ought not confuse advocacy with scientific inquiry. I regard this as a fundamental point of departure, but I know that many advocates of post-modernism disagree with this position. I know that I am old fashioned – older than dirt, so to speak. I regard it as an essential part of my job as a teacher to train learners' ears so that they will become insightful members of a community of speakers of a language that can also be called M-I-U – Mastering Instruction for Understanding. Train your ears to hear the difference between 'meaningful' and 'authentic' – ask authors and interlocuters what is meant by such terms, what positions associated with these terms imply, what has been assumed, and what evidence exists or might exist to confirm or refute such claims. If no evidence will convince a person that a claim is wrong, then you can conclude that this person is not open to scientific inquiry and is advocating a position; you may or may not agree with the position being advocated, but you ought not confuse advocacy with inquiry.

   We picture facts to ourselves (Wittgenstein, 1961). Try picturing these (Spector, n. d.):

1. It would be a remarkable coincidence if the limits of my imagination happened to coincide with the limits of reality.
2. We can say more than we can know.
3. We can know more than we can say.

What we cannot speak about we must pass over in silence (Wittgenstein, 1961).

Success in this enterprise requires having mastered conservation. I know that I have mastered conservation because when I take a glass half full of water and pour it into a similar glass, the new glass into which the water is poured will be half empty.

References

Glasersfeld, E. V. (1987). *The construction of knowledge.* Seaside: Intersystems Publications.

Hofstadter, D. R. (1979). *Gödel, Escher and Bach: An eternal golden braid*. New York: Basic Books.

Popper, K. (1963). *Conjectures and refutations: The growth of scientific knowledge*. London: Routledge and Kegan Paul.

Simon, H. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, MA: MIT Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. (M. Cole, V. John-Steiner, S. Scribner & E. Souberman, Editors and Translators). Cambridge, MA: Harvard University Press.

Wittgenstein, L. (1961). *Tractatus logico-philosophicus* (translated by D. F. Pears and B. F. McGuiness). London: Routledge and Kegan Paul.

Wittgenstein, L. (1963). *Philosophical investigations* (translated by G.E.M. Anscombe). Oxford: Basil Blackwell.

2. Using an Evidence-Based Approach to Assess Mental Models

Valerie J. Shute[1] and Diego Zapata-Rivera[2]

[1] Florida State University, Tallahassee, FL, USA

[2] Educational Testing Service, Princeton, NJ, USA

Abstract: This chapter describes a new idea for the design and development of assessments for mental models using "flexible belief networks" (FBNs). The idea involves joining and extending two assessment approaches—evidence-centered design (ECD) and concept mapping (CM). ECD will be extended beyond single, static proficiency models to dynamic models of learning over time. CM will be extended to include belief networks, which may be accomplished by overlaying concept maps with Bayesian networks. Our goal is to derive a methodology to better assess mental models as they evolve over time, with valid inferences regarding both syntactic (structural) and semantic (conceptual) similarities to reference models. This work leverages the seminal research conducted in the area of assessing mental models by Norbert M. Seel.

Keywords: Belief patterns; concept maps; mental models; formative assessment; evidence-centered design.

Introduction

One rich and enduring area of research in educational and cognitive psychology focuses on learners' construction and use of symbolic (or mental) models of knowledge. Mental models have been implicated in many phenomena that are fundamental parts of human cognition, such as the ability to reason—inductively and deductively—about complex physical and social systems, to generate predictions about the world, and to realize causal explanations for what happens around us (e.g., Gentner & Stevens, 1983).

In an increasingly technological society, understanding the nature of mental models for complex systems, and figuring out how to help people develop and hone good mental models are important goals with potentially large educational and economic benefits (e.g., Seel, 1999; Spector, Dennen, & Koszalka, 2006). In addition to knowledge and systems understanding, such constructed representations can also represent and communicate subjective experiences, ideas, thoughts, and feelings (e.g., Mayer et al., 1999; Seel, 2003).

Learners with access to good mental models demonstrate greater learning—outcomes and efficiency—compared to those with less adequate models in various domains (e.g., Mayer, 1989; DeKleer & Brown, 1981; White & Frederiksen, 1987), particularly mathematics and science. However, *assessing* these internal (hence invisible) mental models is a difficult task. Currently, to assess mental models, researchers often rely on learners' construction of external representations (e.g., concept maps) as a proxy for what resides inside the learner's head. And when the externalized maps are compared with experts' or other reference maps, structural similarities may be computed. But what about assessment of the quality or semantics of the underlying map? New methodologies in educational psychology and artificial intelligence are emerging which may help in this type of assessment effort. We will discuss this in more detail later in the chapter.

Besides difficulties associated with assessing mental models, instructing (or fostering) mental model construction is another large challenge. According to Seel (2003), there are three main instructional paradigms that have been used to promote model building: discovery learning, guided discovery learning, and the more common receptive learning that ensues from a teacher's explanation or an expert's demonstration. The basic premise underlying model-based instructional interventions (that are not purely discovery learning) is that providing learners with models—of tasks and/or representations of causal relations—facilitates knowledge and skill acquisition in the content area, particularly if the models are provided sufficiently early during the course of learning. But this premise is still largely unsubstantiated (see Johnson-Laird, 1989; and Seel, 2003 for more).

The glue that binds these ideas together is called evidence-centered design (ECD; e.g., Mislevy, Steinberg, & Almond, 2003) for assessment, which provides (a) a way of reasoning about assessment design, and (b) a way of reasoning about student understanding. For our purposes, ECD allows the assessment pieces to be joined together to form an informative profile of the learner, and provides the mechanism for specifying and linking concepts and propositions with appropriate evidence needed to demonstrate particular levels of proficiency (or belief). This will be discussed in the next section.

The organization of this chapter is as follows. We begin with some simple definitions to ground the ensuing discussion. This includes: (a) clarifying the distinction between mental models and concept maps; (b) specifying the underlying models and functionality of ECD (e.g., proficiency, evidence, and task models); (c) distinguishing between summative and formative assessment (see Black & Wliam, 1998a; 1998b; Shute, in press; Stiggins, 2002); and (d) distinguishing beliefs from knowledge. After defining key terms and concepts, we will summarize the important contributions of Norbert Seel to the field, and show how we plan to leverage this research for the purpose of assessing the structure and content of mental models using externalized representations. This will then pave the way for deriving innovative instructional interventions—using a formative assessment approach to assist learners in building better mental models.

Definitions

This section operationalizes and compares various terms and concepts including: (a) concept maps vs. mental models, (b) evidence-centered design models, (c) summative vs. formative assessment, and (d) beliefs vs. knowledge.

Concept Maps vs. Mental Models

Concept maps are *external* representations. They comprise the output or product emanating from the process of "concept mapping," which is a popular technique used for visualizing the relationships among different concepts. A concept map (or "causal influence diagram;" see Sterman, 1994; Spector, Dennen, & Koszalka, 2006) is usually a diagram depicting relationships among concepts. Concepts are connected to each other via labeled arrows, typically in a hierarchical structure. Some common links include: "is part of," "causes", "is required by," or "contributes to." Concept mapping began in the 1970s by Novak and colleagues to represent students' emerging knowledge of science (e.g., Novak, 1995; Novak & Gowin, 1984). It has subsequently been used as a tool to increase meaningful learning in the sciences and other subjects as well as to represent the expert knowledge of individuals and teams in education, government, and business.

Mental models are the *internal* representations of reality that people use to understand specific phenomena. Gentner and Stevens (1983) note that these internal models provide predictive and explanatory power for understanding interactions with the world around us. Mental models have also played a prominent role in cognitive processing theories. For instance, Johnson-Laird (1983) proposed that mental models are the basic structure of cognition, "It is now plausible to suppose that mental models play a central and unifying role in representing objects, states of affairs, sequences of events, the way the world is, and the social and psychological actions of daily life." (p. 397). Some characteristics of mental models include: (a) they are incomplete and constantly evolving; (b) they may contain errors, misconceptions, and contradictions; (c) they may provide simplified explanations of complex phenomena; and (d) they often contain implicit measures of uncertainty about their validity that allow them to used even if incorrect.

Evidence-centered Design — Models and Framework

Evidence-centered assessment design (ECD; Mislevy, Steinberg, & Almond, 2003) is a methodology for designing assessments based around the central question of how to gather evidence about a student's knowledge, skills, and abilities. ECD is a knowledge elicitation and management process whereby the goal is a detailed blueprint of the assessment called the conceptual assessment framework (CAF). The CAF is comprised of five different types of models, and a typical CAF contains multiples of each type:

- *Proficiency Model*—Describes students' knowledge, skills, and abilities about which we want to make claims.
- *Evidence Model*—Describes the relationship between observable outcomes from tasks and the relevant proficiency variables.
- *Task Model*—Describes the kinds of situations in which we can observe evidence of proficiencies.
- *Assembly Model*—Describes the collection of proficiency, evidence, and task models that will constitute a given assessment. It contains the rules used to assemble the form of the assessment seen by a learner from a pool of potential tasks.
- *Presentation Model and Delivery System Model*—Describes characteristics of a particular delivery environment, including format, platform and security considerations.

Almond and Mislevy (1999) describe how to use this framework to track the state of an individual learner as more and more observations arrive. The proficiency model, often represented by a Bayesian network (Mislevy, 1994; Almond et al., in press; Shute, Hansen, & Almond, 2007), is instantiated with the prior distribution over the proficiencies for a particular learner. When a set of observations from a task arrives, the appropriate evidence model is attached to the proficiency model and the evidence is absorbed. The evidence model fragment is then discarded and the proficiency model remains, tracking our beliefs about the knowledge, skills, and abilities of the student posterior to the observations.

Summative vs. Formative Assessment

> If we think of our children as plants... summative assessment of the plants is the process of simply measuring them. The measurements might be interesting to compare and analyze, but, in themselves, they do not affect the growth of the plants. On the other hand, formative assessment is the garden equivalent of feeding and watering the plants - directly affecting their growth. Clarke (2001, p. 2).

Summative assessment reflects the traditional approach used to assess educational outcomes. This involves using assessment information for high-stakes, cumulative purposes, such as promotion, certification, and so on. It is usually administered after some major event, like the end of the school year or marking period. Benefits of this approach include the following: (a) it allows for comparing student performances across diverse populations on clearly defined educational objectives and standards; (b) it provides reliable data (e.g., scores) that can be used for accountability purposes at various levels (e.g., classroom, school, district, state, and national) and for various stakeholders (e.g., students, teachers, and administrators); and (c) it can inform educational policy (e.g., curriculum or funding decisions).

Formative assessment reflects a more progressive approach to education. This involves using assessments to support teaching and learning. Formative assessment is tied directly into the fabric of the classroom and uses results from students' activities as the basis on which to adjust instruction to promote learning in a timely manner. This type of assessment is administered much more frequently than summative assessment, and has shown great potential for harnessing the power of assessments to support learning in different content areas and for diverse audiences. When teachers or computer-based instructional systems know how students are progressing and where they are having problems, they can use that information to make real-time instructional adjustments such as re-teaching, trying alternative instructional approaches, altering the difficulty level of tasks or assignments, or offering more opportunities for practice. Such events are, broadly speaking, formative assessment (Black & Wiliam, 1998a). Formative assessment has been shown to improve student achievement (Black & Wiliam, 1998b; Shute, Hansen & Almond, 2007).

In addition to providing teachers with evidence about how their students are learning so that they can revise instruction appropriately, formative assessments (FAs) may directly involve students in the learning process, such as by providing feedback that will help students gain insight about how to improve. Feedback in FA should generally guide students toward obtaining their goal(s). The most helpful feedback provides specific comments to students about errors and suggestions for improvement. It also encourages students to focus their attention thoughtfully on the task rather than on simply getting the right answer (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Shute, 2007). This type of feedback may be particularly helpful to lower-achieving students because it emphasizes that students can improve as a result of effort rather than be doomed to low achievement due to some presumed lack of innate ability (e.g., Hoska, 1993).

An indirect way of helping students learn via FA includes instructional adjustments that are based on assessment results (Stiggins, 2002). Different types of FA data can be used by the teacher or instructional environment to support learning, such as diagnostic information relating to levels of student understanding, and readiness information indicating who is ready or not to begin a new lesson or unit. FAs can also provide teachers or computer-based learning environments with instructional support based on individual student (or classroom) data. Examples of instructional support include: (a) recommendations about how to use FA information to alter instruction (e.g., speed up, slow down, give concrete examples), and (b) prescriptions for what to do next, links to web-based lessons and other resources, and so on.

Black and Wiliam (1998a; 1998b) very clearly established the importance of formative assessment to both teaching and learning. They also originated the widely-used distinction between (a) assessment *for* learning, and (b) assessment *of* learning, which maps to formative and summative assessment, respectively.

Knowledge vs. Belief

Everybody is entitled to their own opinion, but they're not entitled to their own facts.

–Daniel Patrick Moynihan

Although actual philosophers and epistemologists may quibble with the following definitions, we characterize knowledge and belief as follows. Knowledge is the comprehension or awareness of a verifiable idea, proposition, or concept, and the representation thereof. Belief refers to what one accepts as true, rejects as false, or withholds judgment about its truth-value (probabilistic). Furthermore, belief is a representational mental state that could be part cognitive and part affective. Knowledge typically has no affective aspects.

Sometimes the words 'know' and 'believe' are used interchangeably, but they are actually quite different. Belief typically applies to something that you are either unsure about or for which there is insufficient proof. For instance, one might say, "I believe that dogs make better pets than cats." This belief may (or may not) be true, and may be based on an overgeneralization or otherwise inadequate evidence. Knowledge, however, applies to things that are true (or that at least have a reasonable amount of supporting evidence). Therefore, it may be inappropriate to say, "I know that dogs make better pets than cats" because there is an element of doubt (i.e., disputable evidence) involved with this assertion. Knowledge implies belief.

How does knowledge relate to truth? Consider the following: until 1610, nobody *knew* that Jupiter had moons. Then there was a brief period of time when Galileo was the only person who knew Jupiter had moons. Eventually, larger numbers of people knew that Jupiter had moons. This shows that knowledge can change and be unevenly distributed, although the truth did not change in 1610. So, truth is something to be discovered while knowledge is something to be invented. In fact, much of scientific activity revolves around coming up with models that capture some aspects of the truth with some degree of fidelity. And that is just what we're attempting to accomplish with the ideas in this chapter. Now, going back to the example of Galileo's claim that Jupiter had moons, he had difficulty persuading others of this fact. Many simply did not want to believe that Jupiter has moons, and some people have a powerful ability to be blind to what they don't want to see.

So people hold all sorts of beliefs about the world around them. Some beliefs are more accurate than others—depending on the goodness of the evidence underlying the nodes in the belief network. As educators, we would like to be able to make valid inferences about what a person knows and believes, analyze how well that meshes with the body of knowledge and concepts to be learned, and then try to adjust errant or unfounded beliefs toward more reasonable and well-founded ones.

Having defined relevant terms, we now turn our attention to the current state of research in the area of mental models.

Current Research

Seel's contributions to the mental-model *researchscape* has direct relevance to our work in terms of the assessment of mental models. His approach opens up ways to capture important pieces of evidence relevant to aspects of knowing and learning that we have not done with ECD—namely modeling conceptual (or system) understanding. Heretofore, our assessment expertise and development efforts have focused on modeling declarative knowledge and procedural skills. However Seel et al.'s assessment tasks involve externalizing internal representations of conceptual and functional relatedness. Our tasks have tended to be more specific (defined) from an assessment point of view—capturing clear evidence directly from task performances (or from log files—see Shute, Ventura, Bauer, & Zapata-Rivera, in press). Representative tasks of this type include multiple-choice problems or constructed responses, where the key is a clear, known response. Cognitive models permit the analysis and comparison of responses to keys for diagnostic purposes.

Seel (2003) reported on the results from a long-term analysis of model-based teaching and learning. Among the important findings, the basic research on the development of mental models has shown that the models tend not be fixed structures of the mind, but are constructed by learners on an as-needed basis in response to a specific learning situation and associated cognitive demands. Seel thus concluded that mental models are situation-dependent constructions (or reconstructions) of previously generated models, are essential for problem solving, and may be captured via concept maps. Because concept maps are dynamic, adaptable, and interactive, they are well-suited for this purpose, and may be created and used by single persons or by small groups (Weinberger & Mandl 2003). Furthermore, the idea of using such flexible models to make inferences about what a learner knows and believes, to what degree, and the underlying reasons for these beliefs, comprises a great challenge to people who model how the mind works.

In previous assessment and learning research, the authors of this chapter have focused mostly on topics and tasks that (a) are typically well-defined, (b) have a correct solution (or constrained set of solutions), and (c) are free of controversial issues or indirect evidence. But leveraging Seel's research with mental models (and the progression thereof), provides us with an intriguing way to assess much richer mental representations, and to use that information to inform and update our proficiency models. These more comprehensive proficiency models can include information not only about procedural and declarative proficiencies, but also conceptual understanding and the underlying belief structures.

Assessing Concept Maps

In addition to providing a glimpse at internal mental models, concept maps help in organizing learners' knowledge by integrating information into a progressively more complex conceptual framework. When learners construct concept maps for representing their understanding in a domain, they reconceptualize the content domain by constantly using new propositions to elaborate and refine the concepts that they already know. More importantly, concept maps help in increasing the total quantity of formal content knowledge because they facilitate the skill of searching for patterns and relationships among concepts.

A variety of simple measures have been developed to measure completeness and structural complexity of concept maps. These indicators include the number of nodes, number of links, number of cross links, number of cycles, number of hierarchy structures, and number of examples (Vo, Poole, & Courtney, 2005; Novak & Gowing, 1984). Structural matching indicators, such as the deep structure measure from Seel's research, have also been used to determine how close a concept map is to a reference map (i.e., a concept map crafted by an expert) (Ifenthaler, Pirnay-Dummer, & Seel, 2007). Some of these simple indicators have been shown to be reliable and effective measures of the completeness and structural complexity of concept maps, and have been used to support research in the area (e.g., in relation to learning and intelligence). However, although such reliable and simple indicators play an important role in assessing certain characteristics of a concept model, they do not always provide enough information at the right granularity level to support instructional feedback (i.e., feedback that can be used by students to improve their learning).

Understanding the semantics or meaning of a concept map is a very challenging endeavor. The complexity of this problem can be handled by employing approaches that limit the scope of the concepts and relationships that can be represented and require the user to participate in the process to some extent, such as collaborative diagnosis (e.g., Cimolino, Kay, & Miller, 2004). Some of these approaches include: (a) asking students to select from a list of predefined organizational templates (organizers) representing various reasoning patterns (e.g., Ifenthaler & Seel, 2005; Jonassen, Beissner, & Yacci, 1993; Zapata-Rivera, Greer, & Cooke, 2000); (b) using a logic representation of the concept map, dialogue games, and sentence openers (e.g., Dimitrova, 2003; Jeong & Juong, 2007); and (c) using ontologies and teacher feedback to create a knowledge representation middle layer of the concept map that can be used to provide customized feedback to students (Cimolino, Kay, & Miller, 2004).

Both structural and semantic information can be combined in an evidence-based assessment framework (i.e., ECD). Computer-based learning tools developed on top of this framework can then use the information embedded in student concept maps to adapt their interaction. Monitoring the progress of concept maps over time (Seel, 1999; Ifenthaler & Seel, 2005) is an important goal to be achieved. But while the methods employed by Seel et al. are useful for tracking macro-level (or summative) changes in models over time, we also need a more micro-analysis approach to examine the factors that promote and inhibit specific concept mapping behaviors. We now present an extension of ECD to illustrate our plan for modeling student belief structures and their change over time.

Flexible Belief Networks

The basic idea we want to communicate herein concerns our approach to representing a learner's current set of beliefs about a topic as Bayesian networks (Pearl, 1988) that have been overlaid on top of concept maps. By overlaying a probabilistic network (i.e., a Bayesian network) on top of a concept map structure, we can model and question the degree to which relationships among concepts/nodes hold as well as the strength of the relationships. In addition, prior probabilities can be used to represent preconceived beliefs. A probabilistic network provides us with a richer set of modeling tools that we can use to represent the degree to which people ascribe to a particular belief pattern.

Accomplishing this goal would involve incorporating an assessment layer on top of the concept maps to flesh out the maps more fully. This approach would result in a collection of evidence from students in terms of their evolving mental models as indicated by their relationship to the strength and relevance of associations, directionality of the stated relations, and the specified type or nature of the relationship. The result should be a set of flexible belief networks (or FBNs).

To derive these FBNs, we would need to conduct a domain analysis on the topic in question, and use ECD to (a) model belief structures, and (b) design embedded assessments to gather evidence on learners' concepts, misconceptions, and beliefs. By employing embedded assessments, we will be able to infer a learner's current belief structure (via Bayesian networks) based on performance data (evidence) for a variety of purposes—e.g., to modify thinking, or increase cognitive flexibility and perspective taking. The benefits of such an approach are that it would render tacit (unobservable) knowledge and beliefs visible, and permit, if not actively encourage examination. Models (one's own and alternatives) may be displayed via "lenses" to enhance communication and understanding. Each lens would correspond to a particular belief "pattern" that was representative of, and fairly common in the population. The patterns, as will be discussed later, will be derived from both top-down (e.g., interviews with experts) and bottom-up (e.g., data mining) methods. This approach is expected to enable the modeling of changes in beliefs over time.

Figure 1 illustrates a simplified example of the progression from concepts to concept maps to belief nets when Bayesian networks are overlaid to specify structure, node size, and links (i.e., type, directionality, and strength of association). Evidence is attached to each node-relationship which either supports or counters a given claim. The example used here, for illustrative purposes only, represents some of the concepts and relations among variables related to the war in Iraq.

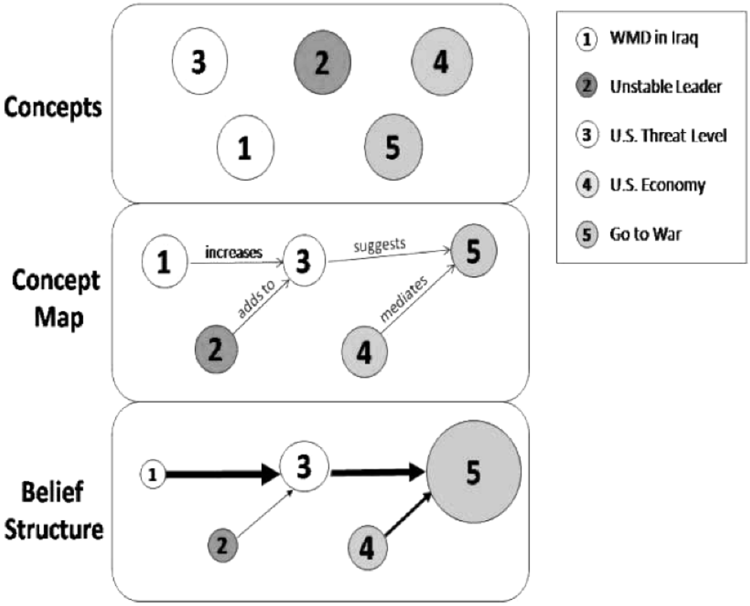

Fig. 1. Progression from concepts to concept map to belief structure.

Note that the *size* of the node in the belief structure indicates a given node's marginal probability (e.g., *p*(node 1 <weapons-of-mass-destruction> = True) = 0.1—a tiny node with a low probability of being true). *Links* illustrate the perceived relationships among the nodes in terms of *type, direction,* and *strength*. *Type* refers to the probabilistic or deterministic representation—defining the nature of the relationship. The s*trength* of the relationship is shown by the thickness of the link, and the *direction* indicates that the relationship has an origin and a destination. The belief structure in Figure 1 models the beliefs of a person (or group of people) that, for example: (a) nodes 1 and 3 exist, (b) the current probabilities of node 1 and node 3 are fairly low (0.1 and 0.3 respectively), and (c) there is a positive and strong relationship between nodes 1 and node 3 (represented by a thick line). So, if the low probability of node 1 (existence of weapons of mass destruction in Iraq) turned out to be true, then the effect on node 3 (U.S. threat level) would be a substantial elevation of the threat level.

A *belief pattern* (BP) is our term for a representative set of nodes and relations. lations. Continuing with the illustrative war in Iraq theme, following are two hypothetical BPs through the eyes of two fictitious persons who differ quite a bit in their respective beliefs about the war (see Figures 2 and 3).

When comparing the two BPs, they contain basically all of the same concepts, but the size of the respective nodes, the directionality of relations, and the strength

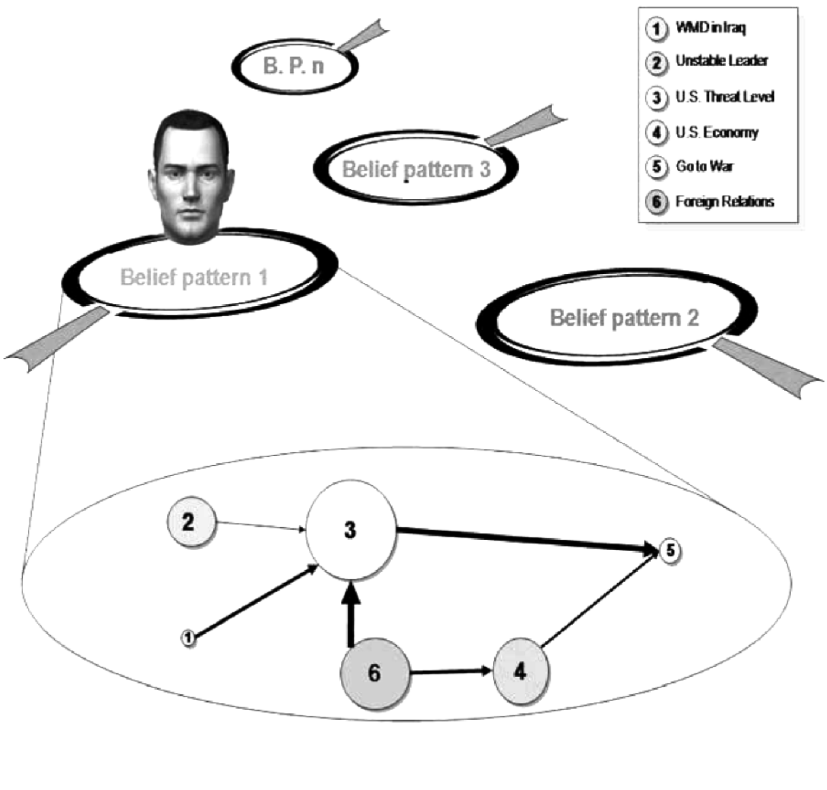

Fig. 2. BP through the lens of Person 1.

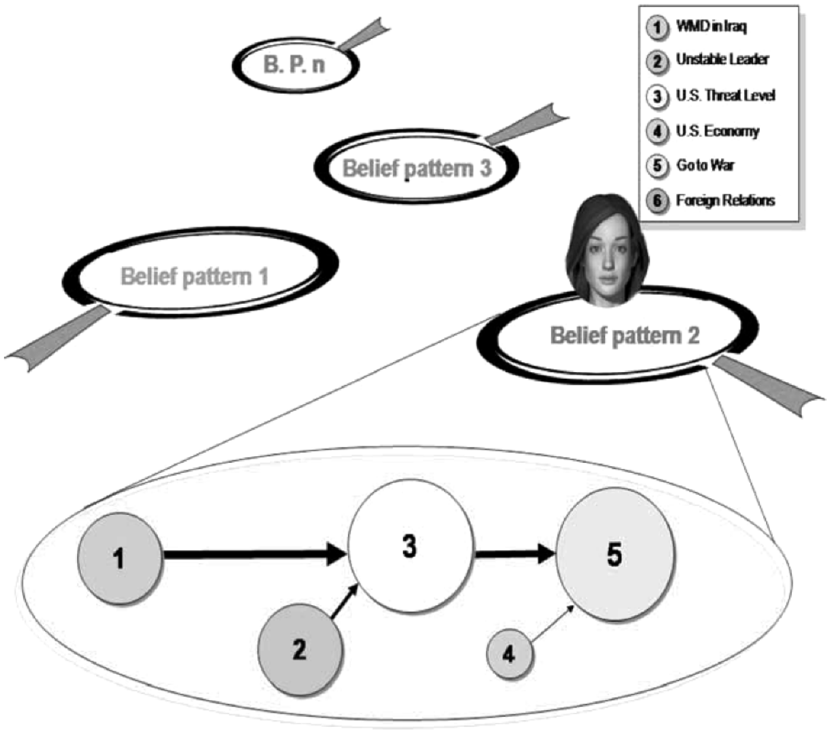

Fig. 3. BP through the lens of Person 2.

of the links are very different. Because we have chosen to use Bayesian networks to represent belief structures, this enables us to examine not only (a) the structure of the map, but also (b) the content (nodes and links), as well as (c) the underlying evidence that exists per structure (and per node). That is, as part of creating a current belief structure, the student arranges concepts and establishes links, and he or she includes specific evidence (sources) per claim (i.e., arguments and documentation in support of, or in opposition to a given claim). The credibility of the evidence, then, should match the strength of the links established in the structure. For instance, if a student made a strong claim about the existence of WMD in Iraq, and cited a dubious source as the only evidence, then that would not count as being credible evidence—and would imply that the student needed some assistance in his critical thinking/analysis skills. In short, we not only want to model the structures, but also the supporting evidence that lives underneath. Figure 4 shows a generic model with its supporting evidence attached.

So how do we accomplish this kind of modeling? There are five main parts to our proposed BP modeling approach:

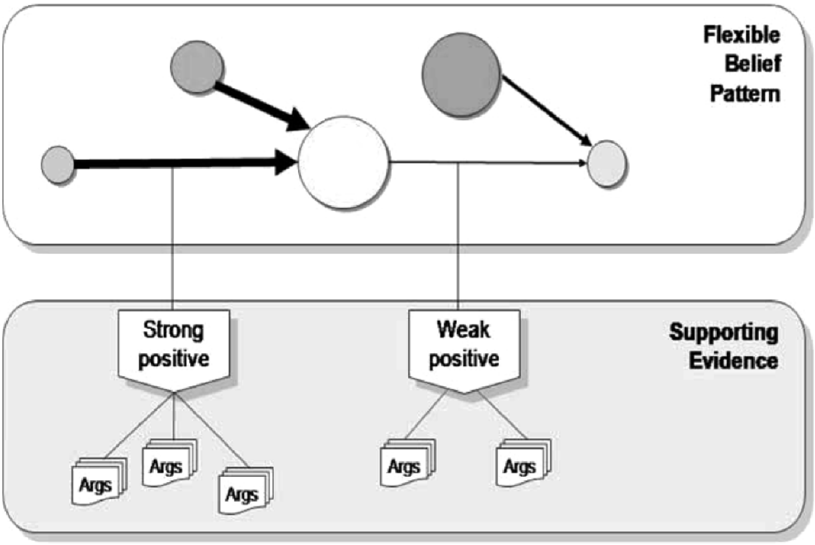

Fig. 4. Supporting evidence underlying an example BP.

1. Analyze the domain and integrate belief nets from human experts and data mining efforts.
2. Create an initial set of belief patterns (BPs).
3. Model beliefs, evidence, and assessments via extended ECD.
4. Infer individual BPs via assessment.
5. Model changing BPs over time via Dynamic Belief Networks (DBN).

The first step is to analyze the domain. This involves defining and structuring information about the topic area. It is instantiated as an FBN. Data sources for domain analysis include: (a) top-down creation of FBNs via ECD (e.g., subject-matter experts, research papers), and (b) bottom-up data mining to yield a large collection of variables relating to the topic, their relations from different perspectives, supporting arguments, claims, and so on. Data to be mined include: journal articles, blogs, listservs, newspapers, public documents, and data from surveys and tasks that students complete to further feed the models. This analysis phase is analogous to conducting a factor analysis on data to discern patterns.

The second step is to generate BPs. This may also be accomplished via top-down and bottom-up processes to effectively merge data from the analysis step – from data mining activities and subject-matter experts. This step informs the creation of the FBPs – both initial and alternative belief patterns.

The third step entails modeling using the proposed extended-ECD approach, and it has two main foci: designing valid assessments and diagnosing knowledge and beliefs. The assessment design process begins with defining three main models: (1) the belief model (BM)—What do you want to say about the person—what does she know and what does she believe is true?, (2) the evidence model (EM)—What observations would provide the best evidence for what you want to say?, and (3) the task model (TM)—What kinds of tasks or scenarios would allow you to make the necessary observations?. ECD thus provides a systematic way of laying out assessments, complete with evidentiary arguments that explicitly link performance data to claims about underlying knowledge and beliefs. Figure 5 shows three ECD models for a particular Belief Pattern (in this case, BP 1). Flowing from left-to-right, we depict the assessment design process from the Belief Model (labeled 'Current BP' in the Figure) to Evidence Model to Task Model. The Current BP model represents the initial *organization of concepts* (including preconceptions and misconceptions), beliefs, and relationships. Tasks will ultimately be designed to impose structure. Next, the Evidence Model specifies the criteria or rubrics needed for *evidence* of the current BP (i.e., specific student performance data, or observables). Finally, the Task Model contains a range of templates and parameters for *task development* to elicit data needed for the evidence model.

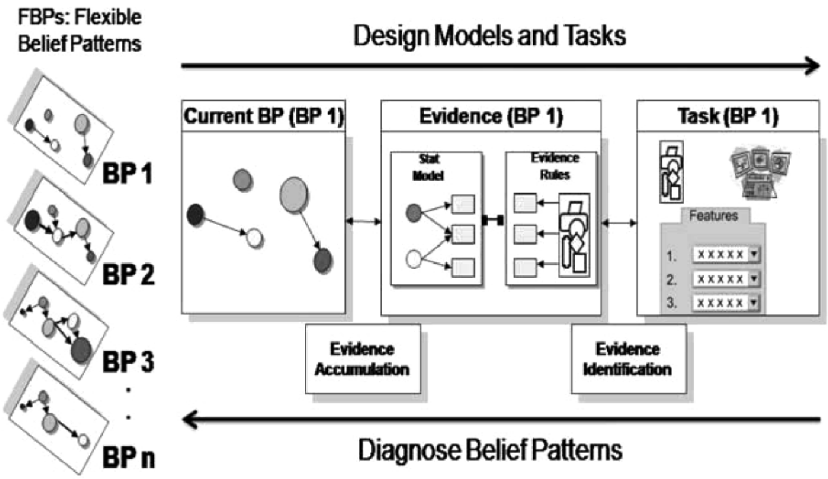

Fig. 5. Designing models and tasks based on extended ECD and diagnosing belief patterns based on students' performance data.

Reversing the flow (from right-to-left) permits diagnosis of what the learner knows/believes, and to what degree as related to each of the BPs. In Figure 5, "evidence identification" refers to the collection and scoring of data to analyze how the student performed while the "evidence accumulation" process refers to the derivation of inferences about what the student knows/believes, and how strongly it is known or believed.

The fourth step concerns the inference of belief patterns. That is, after links are inferred (direction, strength, and type), each student is associated with a particular

BP, at a particular point in time. Here, we propose to use embedded assessments to infer BPs from users' performance data (observables). These BPs may be mapped to initial BPs derived from the Domain Analysis part of the process. Assessment of BPs will include knowledge, concepts (preconceptions, misconceptions), links, argument structures, biases, and so forth. Environments (i.e., embedded tasks and interventions) may include virtual reality, simulations, and tasks like IAT (implicit association tasks; see Greenwald & Banaji, 1995) to reduce "faking" and get at deeply hidden beliefs. We expect that entrenched beliefs will be relatively easy to assess given strong and consistent response patterns. However, research has suggested that entrenched beliefs are harder to modify than existing knowledge (Griffin & Ohlsson, 2001).

The final step involves modeling BPs over time. To accomplish this goal, we plan to (a) use the extended ECD approach to track changes in BPs over time; (b) use ECD models to provide parameters to create different interventions (e.g., VR, simulations, etc.), and (c) assess each user at the beginning and end of a given "session" to see the effects of the intervention(s) on the students' BPs. Figure 6 depicts the modeling over time.

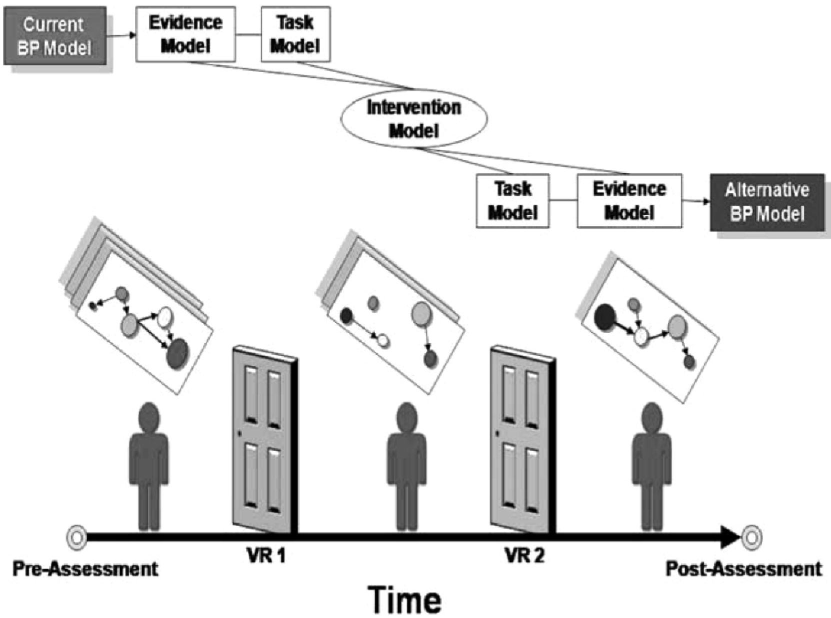

Fig. 6. Modeling BPs over time.

After the modeling part is accomplished, the next challenge will be to design effective and integrated interventions (e.g., making belief nets visible, showing others' nets, highlighting misconceptions, and so on) which must be coordinated because assessments will be embedded directly within the interventions.

Instructional Interventions

Once we know how mental models develop, and we can assess them syntactically and semantically, we will be able to set up interventions to foster their growth and development. That is, determining a learner's current BP is the critical first step in designing and delivering appropriate interventions. These interventions could include exposing the learner to an external representation of the current BP (e.g., an active belief map) and letting the learner react to it. For example, the learner could explain whether the current representation truly reflects what he/she believes, or if it does not, then why not. This can be done by allowing the learner to directly manipulate and annotate the current BP (Zapata-Rivera & Greer, 2004). We can also show the learner someone else's BP and ask her to compare it to her own BP. In fact, our intervention model, which serves to link ECD-based models (see Figure 6) can leverage Seel's model-based learning and instruction framework (Seel, 2003). That is, we can employ Seel's framework to design a variety of interventions that will help the learner analyze and reflect on his/her beliefs by using BP maps that change over time (e.g. Seel's "progression of mental models" concept). In short, we plan to combine Seel's framework with ideas underlying formative assessment as part of the instructional interventions.

Conclusion

Norbert Seel's foundational contributions to the areas of assessment and instructional use of mental models has informed and inspired many of our current ideas. There are still many challenges that lie ahead including: testing our FBN ideas across several "wicked" (i.e., ill-structured) topics, identifying conditions or factors that encourage or inhibit the processes of creating complex links between concepts/arguments, and creating effective interventions that make use of the rich mental model information.

We have described our idea for creating and using evidence-based flexible belief networks and their potential for serving as valid models for instructional intervention as well as communication tools that can be used to enhance learning, argument structures, and cognitive flexibility (e.g., Spiro et al., 1991). Some remaining research questions in this area include the following: If the ultimate goal is to diagnose entrenched BPs in order to help people acquire new knowledge and/or well-founded beliefs, how can we best exploit the information from the various models to create appropriate interventions? Also, how should the belief nets integrate knowledge and possibly affective aspects into the BPs? How broad and/or flexible should these FBNs be in relation to the scope of link types, node types, and so forth to be included in our BPs? Obviously, much more research is needed in this area, but we are very grateful for the firm foundation laid by

Norbert Seel and colleagues with their important research on mental models throughout the years.

Acknowledgments

In addition to Norbert Seel, we would like to acknowledge the following persons for their sage comments on an earlier draft of this paper: Allan Jeong, J. Michael Spector, Dirk Ifenthaler, Eric Hansen, and Jody Underwood.

References

Almond, R. G. (2007). *Cognitive modeling to represent growth (learning) using Markov Decision Processes*. Princeton, NJ: ETS Research Report.

Almond, R. G., & Mislevy, R. J. (1999). Graphical models and computerized adaptive testing, *Applied Psychological Measurement, 23,* 223-238.

Almond, R. G., Mislevy, R. J., Steinberg, L. S., Williamson, D. M., & Yan, D. (in press). *Bayesian networks in educational assessment,* manuscript in preparation to be published by Springer.

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213-238.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-71.

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Cimolino, L., Kay, J., & Miller, A. (2004). Concept mapping for eliciting verified personal ontologies, Special Issue on concepts and ontologies in WBES of the *International Journal of Continuing Engineering Education and Lifelong Learning, 14*(3), 212-228.

Clarke, S. (2001). *Unlocking formative assessment—Practical strategies for enhancing pupils' learning in the primary classroom*. London: Hodder and Stoughton.

deKleer, J. & Brown, J. S. (1981). Mental models of physical mechanisms and their acquisition. In J. R. Anderson (Ed.), *Cognitive Skills and their Acquisition.* Hillsdale, NJ: Erlbaum.

Dimitrova, V. (2003). StyLE-OLM: Interactive open learner modelling, *International Journal of Artificial Intelligence in Education, 13*(1), 35-78.

Forbus, K., Carney, K., Harris, R., & Sherin, B. (2001). *Modeling environment for middle-school students: A progress report*, in Biswas, G. (Ed.): Papers from the 2001 *Qualitative Reasoning Workshop*, The Printing House, Stoughton, WI.

Gentner, D., & Stevens, A. (1983). *Mental models*. Hillsdale, NJ: Erlbaum.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4-27.

Griffin, T., & Ohlsson, S. (2001). Beliefs versus knowledge: A necessary distinction for explaining, predicting, and assessing conceptual change. In J. D. Moore and K. Stenning (Eds.), *proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 364-369). Mahwah, NJ: Lawrence Erlbaum.

Hoska, D. M. (1993). Motivating learners through CBI feedback: Developing a positive learner perspective. In V. Dempsey & G. C. Sales (Eds.), *Interactive instruction and feedback* (pp. 105-132). Englewood Cliffs, N.J.: Educational Technology Publications.

Ifenthaler, D., Pirnay-Dummer, P., & Seel, N. M. (2007). The role of cognitive learning strategies and intellectual abilities in mental model building processes. *Technology, Instruction, Cognition and Learning,* Special Issue on Applying Cognitive Modeling.

Ifenthaler, D., & Seel, N. M. (2005). The measurement of change. Learning-dependent progression of mental models. *Technology, Instruction, Cognition and Learning*, *2*(4), 317-336.

Jeong, A. & Juong, S. (2007). Scaffolding collaborative argumentation in asynchronous discussions with message constraints and message labels. *Computers & Education, 48*, 427-445.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge: Cambridge University Press.

Jonassen, D., Beissner, K., & Yacci, M. (1993). *Structural knowledge techniques for representing, conveying, and acquiring structural knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers Hove & London.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research, 59*(1), 43-64.

Mayer, R. E., Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communication by minimizing cognitive load. *Journal of Educational Psychology, 91*(4), 638-643.

Mislevy, R. J. (1994). Evidence and inference in educational assessment, *Psychometrika, 12,* pp. 341-369.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessment. *Measurement: Interdisciplinary Research and Perspective, 1* (1), 3-62.

Novak, J. D. (1995). Concept mapping: A strategy for organizing knowledge. In S. M. Glynn, & R. Duit (Eds.), *Learning Science in the Schools: Research Reforming Practice* (pp. 229-245), Mahwah, NJ: Erlbaum Associates.

Novak, J. D., & Gowin, D. R. (1984). *Learning how to learn*. New York: Cambridge Press.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference.* San Mateo, CA: Kaufmann.

Seel, N. M. (1999). Educational diagnosis of mental models: Assessment problems and technology-based solutions. *Journal of Structural Learning and Intelligent Systems, 14* (2), 153-185.

Seel, N. M. (2003). Model-centered learning and instruction. *Technology, Instruction, Cognition and Learning, 1*(1), 59-85.

Shute, V. J. (in press). Tensions, trends, tools, and technologies: Time for an educational sea change. To appear in C. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning*. Mahwah, NJ: Erlbaum Associates.

Shute, V. J. (2007). *Focus on formative feedback.* ETS Research Report, RR-07-11 (pp. 1-55), Princeton, NJ.

Shute, V. J., Hansen, E. G., & Almond, R. G. (2007). An assessment for learning system called ACED: *Designing for learning effectiveness and accessibility*, ETS Research Report, RR-07-26 (pp. 1-54), Princeton, NJ.

Shute, V. J., Ventura, M., Bauer, M. I., & Zapata-Rivera, D. (in press). Melding the power of serious games and embedded assessment to monitor and foster learning: Flow and grow. To appear in U. Ritterfeld, M. J. Cody, & P. Vorderer (Eds.), *The social science of serious games: Theories and applications.* Philadelphia, PA: Routledge/LEA.

Spector, J. M., Dennen, V. P., & Koszalka, T. (2006). Causal maps, mental models and assessing acquisition of expertise. *Technology, Instruction, Cognition and Learning, 3*, 167-183.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1991). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. *Educational Technology*, 24-33.

Sterman, J. D. (1994). Learning in and about complex systems. *System Dynamics Review, 10*(2-3), 291-330.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning, *Phi Delta Kappan Professional Journal, 83*(10), pp. 758-765, 2002.

Vo, H. V., Poole, M. S., & Courtney, J. F. (2005). An empirical comparison of causal mapping approaches. In D. J. Armstrong (Ed.), *Causal Mapping for Information Systems and Technology Research: Approaches, Advances and Illustrations* (pp. 142-173), Battle Creek, MI: Idea Group Publishing.

Weinberger, A. & Mandl, H. (2003). Computer-mediated knowledge communication, *Studies in Communication Sciences*, 81-105.

White, B. & Frederiksen, J. (1985). Qualitative models and intelligent learning environments. In R. Lawler & M. Yazdani (Eds.), *Artifical Intelligence and Education*. Norwood, NJ: Ablex.

Zapata-Rivera, J. D., & Greer, J. (2004). Interacting with Bayesian student models. *International Journal of Artificial Intelligence in Education, 14*(2), 127-163.

Zapata-Rivera, J. D., Greer, J., & Cooke, J. (2000). An XML-based tool for building and using concept maps in education and training environments. *Proceedings of the 8th Conference on Computers in Education*, ICCE 2000, 755-762.

3. Practical Solutions for the Diagnosis of Progressing Mental Models

Dirk Ifenthaler

Albert-Ludwigs-University Freiburg, Germany

Abstract: The question of how to diagnose the learning-dependent progression of mental models has been discussed extensively over the last 20 years. However, many questions about the diagnosis of changes in mental model construction still remain unsolved. This chapter begins with a critical analysis of assessment techniques in order to illustrate a variety of methods for assessing changing externalized knowledge structures such as mental models. However, in addition to choosing an adequate assessment technique for the diagnosis of changing models it is also necessary to define a satisfactory statistical procedure for the measurement of change. The chapter thus continues with a historical synopsis of measurement of change which illustrates the general statistical concerns involved in quantitative studies over time. The dilemma of how to measure progressing mental models adds even more complexity to the ambitious project of diagnosing progressing models. Therefore, the chapter closes by presenting eight empirically tested methodological solutions for further investigation of the progression of mental models.

Keywords: Measurement of change; progression of mental models; methodology.

Introduction

Diagnosis is the systematic and theory-based collection and preparation of information with the aim of justifying, controlling and optimizing conclusions and procedures. In the social and behavioral sciences, the diagnosis of change and the detailed investigation of why and how change takes place is of particular interest. Investigating changes of knowledge structures and understanding how to influence them is the key to well-designed and effective learning environments. Seel (1999, p. 180) concludes that “each subject should be measured repetitively over extended periods of time in order to understand the continuous progression of learning and thinking.”

For the last 20 years, the diagnosis of the learning-dependent progression of mental models and how to influence them through instruction has been discussed extensively (see Johnson-Laird, 1989; Seel, 1991; Al-Diban, 2002). However, many questions about the diagnosis of the learning-dependent progression of mental models still remain unsolved (Ifenthaler & Seel, 2005). Are there reliable and valid assessment techniques for capturing changing models? Which statistical procedures meet the requirements for an analysis of longitudinal data? Are tools and computer software available for these statistical procedures? Which experimental designs that use repeated measurements are suitable for a precise investigation of changing models? A researcher will be confronted with these and other problems whenever mental models and their learning-dependent progression are the focus of empirical study. However, before going into the problems involved in diagnosing the progression of mental models, we will briefly describe the frequently used concepts *measurement of change*, *variability*, and *change* from a methodological point of view.

According to Kleiter (1987), the *measurement of change* comprises all data collection and statistical procedures used to test the same phenomenon in multiwave measurements over a defined period of time. Since changes in learning and thinking take place continuously, educational research needs to move beyond the traditional two-wave design in order to capture these changes more precisely (see Willett, 1988). It is therefore necessary to conduct multiwave longitudinal experiments when complex changes over time are of central interest (Seel, 1999).

But when will a researcher report *intraindividual change* over time and when will she or he report *intraindividual variability*? For Nesselroade (1991, p. 94), the answer is straightforward: "*Intraindividual variability* identifies short-term, relatively reversible changes or fluctuations. […] *Intraindividual change* designates long-term changes that usually are relatively not so reversible." According to Eid (2003), *intraindividual changes* can interfere with situation-dependent *variability*. Therefore, the experimental design for the *measurement of change* must exclude such *variability*. These definitions of concepts provide the methodological foundation for the later sections of this chapter.

First, this chapter will provide an overview of different techniques for assessing changing models. The following historical synopsis of measurement of change will focus on doubts, problems and solutions of various statistical approaches during the last century. We will then discuss special dilemmas involved in the measurement of changing models and present methodological solutions from recent empirical research studies. The chapter will conclude with suggestions for future research on the diagnosis of changing models.

Assessment Techniques

Seel (1997) claims that the question of valid and reliable measurement of changing models is one of the central problems of mental model research. As mental models are not directly observable, individuals have to externalize their mental models before researchers can analyze and interpret them. Hence, externalization is defined as a conscious process of communicating internal knowledge structures using adequate sign and symbol systems (see Le Ny, 1993; Ifenthaler, 2006).

However, externalizing mental models requires a dual process of encoding (Wygotski, 1969; Stachowiak, 1973; Seel, 1991; Galbraith, 1999; Stylianou, 2002; Hanke, 2006). The first process of encoding is described as *internal encoding*. Within this process, a mental model is constructed out of one's actual available world knowledge in order to create subjective plausibility. This mental model is represented as an internal knowledge structure using adequate sign and symbol systems. The second process of encoding only occurs if a person needs to communicate her or his mental model. This process not only requires the use of adequate sign and symbol systems but also a format of communication which other people can understand (see Seel, 1991; Ifenthaler, 2006). Accordingly, an externalization of a mental model is a *re-representation* of a person's understanding of the phenomena to be explained (see Pirnay-Dummer, 2006).

Nevertheless, a *re-representation* of a mental model is the only source from which we can learn about how mental models are constructed and how they change over time. Accordingly, diagnoses of progressing mental models – or externalizations of them-have to be assessed with valid and reliable techniques. Ifenthaler (2006) illustrates and critically analyzes a variety of techniques for assessing changing externalized knowledge structures such as mental models. These techniques include *Thinking Aloud Protocols* (see Ericsson & Simon, 1993, 1998; van Someren, Barnard & Sandberg, 1994), *Structure Formation Techniques* (Scheele & Groeben, 1984, 1988), *Concept-Mapping Tools* (Eckert, 1998; Weber & Schumann, 2000; Reiska, 2005), *Causal Diagrams* (Seel, 1999; Al-Diban, 2002), and *DEEP methodology* (Spector, 2004; Spector, Dennen, & Koszalka, 2005). From a methodological point of view, these techniques have their individual strengths and weaknesses.

Thinking Aloud Protocols enable the verbalization of individual cognitive processes. However, the data collected represents only a small amount of the cognitive processes which occur when one solves a complex problem. Accordingly, the quantification of the collected data and the explicit relation to cognitive processes call the validity and reliability of this technique into question (see Nisbet & Wilson, 1977).

Scheele and Groeben (1984, 1988) developed a *Structure Formation Technique* (SLT) to represent a test person's subjective theories using paper and pencil concepts and named relations. The test person's individual paper and pencil networks can be compared with expert networks and solutions. Although the time-consuming consuming technique requires for the test person to have persistent cognitive ability and high motivation, the networks can be analyzed with regard to their content and their structural shape.

The computer-based *Concept-Mapping Tools* enable the assessment of a person's individual structural and semantic knowledge (Mandl & Fischer, 2000). The collected data can be quantified and processed directly for further statistical analysis. Additionally, the use of the computer enables persons to easily reconstruct and reorganize their individual concept map during the experimentation.

Causal Diagrams are considered to be re-representations of the test persons' mental models (see Al-Diban, 2002). Three criteria – "Goodness of Causal Diagrams", "Depth of Connectivity", and "Complexity" – have been developed to quantify changes in a person's re-representation. The *Test for Causal Diagrams* has been found in a series of quasi-experiments to be satisfactory in terms of objectivity, reliability, and validity (see Al-Diban, 2002). However, the data collection is very time consuming.

Spector (2004) developed the *Dynamic Evaluation of Enhanced Problem-solving* (DEEP) methodology as a web-based knowledge mapping technique for problem-based assessment (see Figure 1). The technique identifies gaps between novice and expert decision making in complex domains. Further developments of the *DEEP* methodology are to include indicators for precise quantitative analysis.

The five techniques discussed above provide different indicators for the analysis of changing models. The following four techniques represent recent developments for the assessment of externalized knowledge structures.

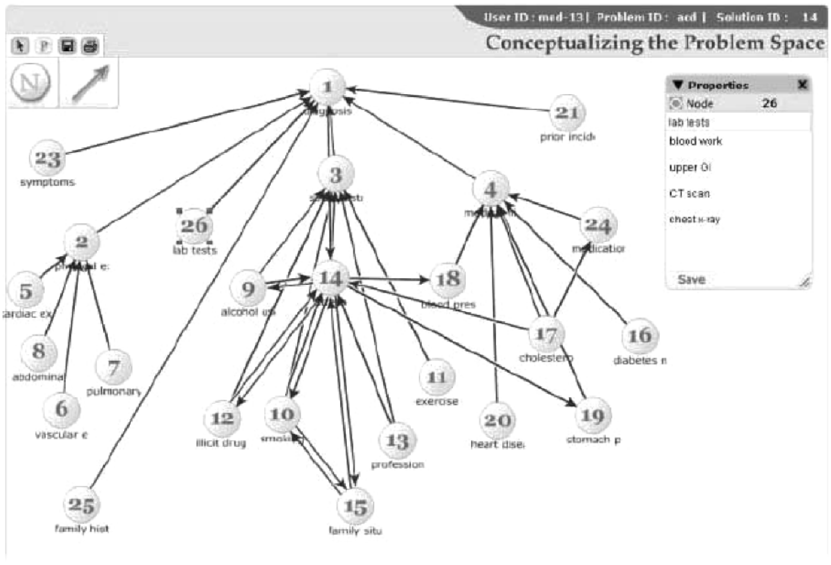

Fig. 1. DEEP methodology (Spector, 2004).

Taricani and Clariana (2006) describe a technique for automatically scoring open-ended concept maps using the *Pathfinder Networks analysis* method (Schvaneveldt, 1990). The indicator focuses on the correctness of propositions, i.e. the correctness of a link between two concepts. The new technique will be further developed as a software tool that can be used to construct and automatically score concept maps.

The *ACSMM* methodology (O'Connor & Johnson, 2004; Johnson, O'Connor, Spector, Ifenthaler, & Pirnay-Dummer, 2006) identifies sharedness among team members. The technique consists of five phases using concept maps as knowledge representation: (1) *Elicitation Design*, (2) *Individually Constructed Mental Model Elicitation (ICMM)*, (3) *ICMM Coding*, (4) *Shared Analysis*, and (5) *ACSMM Construction*. The indicators of the *ACSMM* technique provide information about the sharedness of individual models with a team model and their changes over time.

Pirnay-Dummer (2006) introduced the computer-based *Model Inspection Trace of Concepts and Relations* (MITOCAR), which uses natural language expressions as input data for model re-representations instead of using graphical drawings. The technique is made possible by parsing and corpus linguistics technologies (see Maedche et al., 2002). In two phases of analysis, it generates an automated report including parameters based on graph theory and *SMD Technology* (Ifenthaler, 2006), tests for multidimensional scaling, tests for homogeneity and model complexity, as well as sorted lists of statements. Additionally, the reports give automatically generated assistance in the interpretation of all statistical tests, which makes the test suitable for non-experts on research methods, e.g. instructional designers, teachers, etc.

The *SMD Technology* (Ifenthaler, 2006) uses graphical representations or concept maps to assess individual processes in persons solving complex problems at multiple intervals over time. The computer-based and automated *SMD Technology* is composed of three levels – Surface, Matching, and Deep-Structure. The first level of the *SMD Technology* is the Surface Structure, which enables a rapid and economical assessment of the relational structure of graphical representations and concept maps. The assessment of the structural properties of the externalized cognitive systems is realized on the Matching Structure level. This second level indicates the complexity of the graphical representations or concept maps. The change in complexity of the individuals' representations is another key indicator for the assessment of learning and problem solving processes (see Seel, 1991). The third level of the *SMD Technology* is defined as the Deep Structure. This is the level on which the semantic structure of the graphical representations or concept maps is assessed. The Deep Structure is calculated with the help of the similarity measure (Tversky, 1977) as the semantic similarity between an individual representation and a reference representation (e.g. expert solution to a problem).

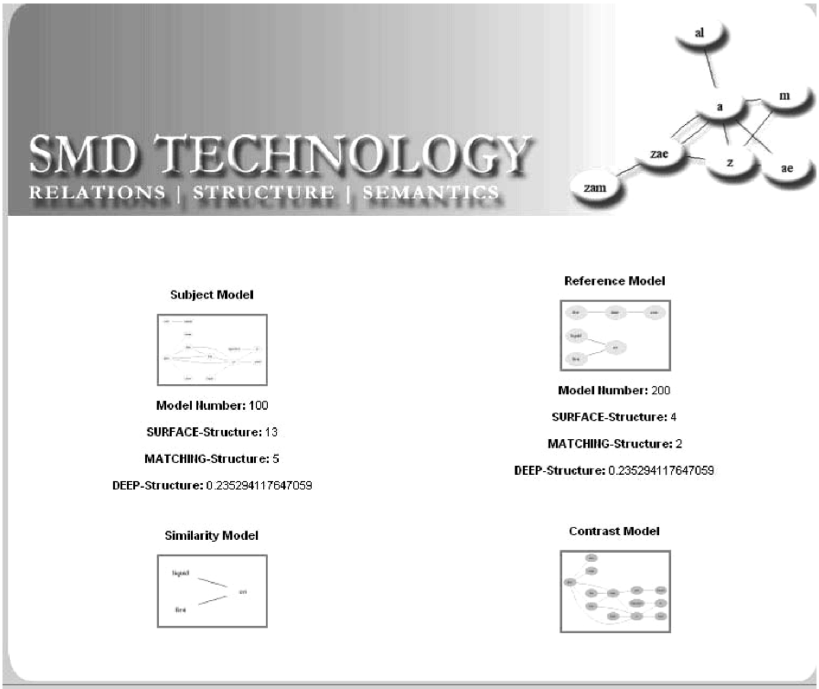

Fig. 2. Automated output of the SMD Technology.

The assessment of progressing individual graphical representations or concept maps is made possible by the computer-based and automated *SMD Technology*, which generates three indicators: (1) sum of propositions, (2) complexity of an individual model, and (3) the semantic similarity between a domain-specific expert representation and an individually constructed model. Additionally, a standardized graphical representation of each individual model is generated automatically (see Figure 2). The application of the instrument is considered to be fast, economical, and domain independent.

A further development of these assessment techniques will include automated data analysis and elaborated feedback on individual model representations. As most of these assessment techniques are applicable for the assessment of changing models, it is essential to prove their validity and reliability before using them (Seel, 1999).

However, an adequate assessment technique for the diagnosis of changing models does not necessarily include a satisfactory statistical procedure for the measurement of change. Therefore, the following synopsis describes how statistical approaches for the measurement of change have developed over the past decades.

Historical Synopsis of Measurement of Change

Fahrenberg (1968) states of the origins of psychological process analysis that Oehrn (1889) and Seashore and Kent (1905) were the first to give an account of a quantitatively observable intraindividual change. However, he also points out that this early work was limited due to "the low amount and biased selection of research variables" as well as "the mathematical-statistical difficulties" (Fahrenberg, 1968, p. 44). Process analysis was continued by the publications of personality researchers such as Heiß and Cattel.[1]

The methodological discussion on *measurement of change* is attributed great significance in a collection of essays edited by Harris (1963) entitled "Problems in measuring change."[2] In his article on the fundamentals of the topic, Bereiter (1963) tackles questions concerning the "over-correction-under-correction dilemma," the "unreliability-invalidity dilemma," and the "physicalism-subjectivism dilemma." Harris (1963) includes discussions of univariate and multivariate analysis models (see Gaito & Wiley, 1963; Bock, 1963; Horst, 1963) as well as statistical models for single-case analyses (see Holtzmann, 1963).

Cronbach & Furby (1970, p. 68) take up the discussion on *measurement of change* and label it "a persistent puzzle in psychometrics." Several years later, "some probabilistic models for measuring change" are introduced by Fisher (1976).

A further concern of the publications is the development of methods for detecting change. Fahrenberg, Kuhn, Kulick, and Myartek (1977) consider a wealth of vastly different methods and also shed light on the methodological problems presented by repeated observations in tests. The relevance of measurement of change for classical test theory is discussed by Petermann, Hehl, and Schneider (1977). The comprehensive spectrum of methods offered by Petermann (1978) includes exemplary introductions to *probabilistic test models*, *Makroff models*, *computer simulations*, and *single-case observations*. Tack (1980) endeavors to formalize measurement of change in order "particularly to shed light on the problems arising from the situation-bound nature of measurements within the context of measurement of change" (Tack, 1980, p. 105). Fischer & Formann (1982) discuss a linear-logical model of measurement of change. While the LLRA[3] does not require for observations to be quantified, it does provide quantitative effect parameters. Spada (1983, p. 83) is skeptical as to whether the detection of changes is "a sad chapter in the history of psychological process analysis."

---

[1] Fahrenberg (1968, p. 44) recounts that Heiß and Cattel published articles on personality psychology dealing with process research simultaneously in 1947.

[2] The publication is the product of a three-day conference held in Madison, Wisconsin in 1962. The goal of the conference was to discuss and reflect on current problems in measurement of change (see Harris, 1963, p. vii).

[3] LLRA stands for "linear logistic model with relaxed assumptions" (see Fischer & Formann, 1982).

In Petermann (1986), an analysis of the literature provides the point of departure for an overview of problems and recent developments in measurement of change. The study focuses on the conceptual difficulties of measuring change, test theory, single-case diagnosis and single-case statistics, and formal models such as LISREL and computer simulations. Kleiter (1987) discusses the state of and prospects for research on measurement of change from a formal-representational perspective. Willett (1988, 1989) laments that most publications on measurements of change have contributed only little to the solution of empirical problems of learning.

Collins & Horn (1991) present the "best methods for the analysis of change" and report on current developments, unanswered questions, and the future course of measurement of change. Renkl & Gruber (1995) use examples from the psychology of memory to demonstrate the opportunities for an adequate statistical modeling of changes by means of hierarchical linear models. Collins & Sayer (2001) report that enormous progress has been made in the domain of measurement of change since the appearance of the publication by Collins & Horn (1991). The collection of essays edited by Moskowitz & Hershberger (2002) places emphasis on the current discussion on the methods and applications of measurement of change.

The 2005 spring conference of the AEPF[4] focused on measurement of change and longitudinal studies. Held at the Free University of Berlin, Germany, the conference included 23 symposia, 83 paper presentations, and 25 poster presentations and met with an "unusually positive response" (AEPF, 2005, p. 7). In addition to current research studies, the conference focused on several special problems of measurement of change which will be dealt with in more detail in the following.

The latest publication of Ittel & Merkens (2006) emphasizes the need for more research designs with repeated measurements in the educational sciences. However, a stronger focus on the methodology of measurement of change requires revised approaches and a sharpened awareness of the problems of longitudinal research designs.

Dilemmas in the Diagnosis of Progressing Mental Models

The historical synopsis of measurement of change provided above illustrates the general concerns about quantitative studies over time. This section, on the other hand, will focus on special problems in the diagnosis of progressing mental models. On the basis of classical test theory it is assumed that a cognitive construct (such as a mental model) can be described by a measured quantity x_i, which is comprised of the real quantity X_i and the corresponding measurement error e_{xi}. Hence, the measured quantity is determined by adding the real value and the cor-

[4] AEPF is short for "Arbeitsgruppe empirisch-pädagogische Forschung" (Empirical-Pedagogical Research Work Group).

responding error of measurement (Petermann, 1978, p. 26). From a mathematical perspective it seems appropriate to represent changes in a cognitive construct between two points of measurement by way of a differential measurement (see Stelzl, 2005). In addition to simple differential measurements, a distinction is made between correlation, regression, and residual measurements (Petermann, 1978, p. 33). If, however, one follows the argumentation of Bereiter (1963, p. 3), three "persisting dilemmas in the measurement of change" should not be left out of account. These dilemmas are the *over-correction-under-correction dilemma*, the *unreliability-invalidity dilemma*, and the *physicalism-subjectivism dilemma*. Some problems, for example the raw-score differences, are eliminated by using (a) modern approaches such as structural equation modeling for measuring true score changes or (b) item response models (e.g., LLMC) for measuring changes in ability underlying subjects' performance on a test. According to Nesselroade (1991) and Raudenbush & Bryk (2002), many of the dilemmas with the measurement of change have been dispelled and new conceptions and approaches provide a wide range for the diagnosis of psychological phenomena over time (see Willet, 1997).

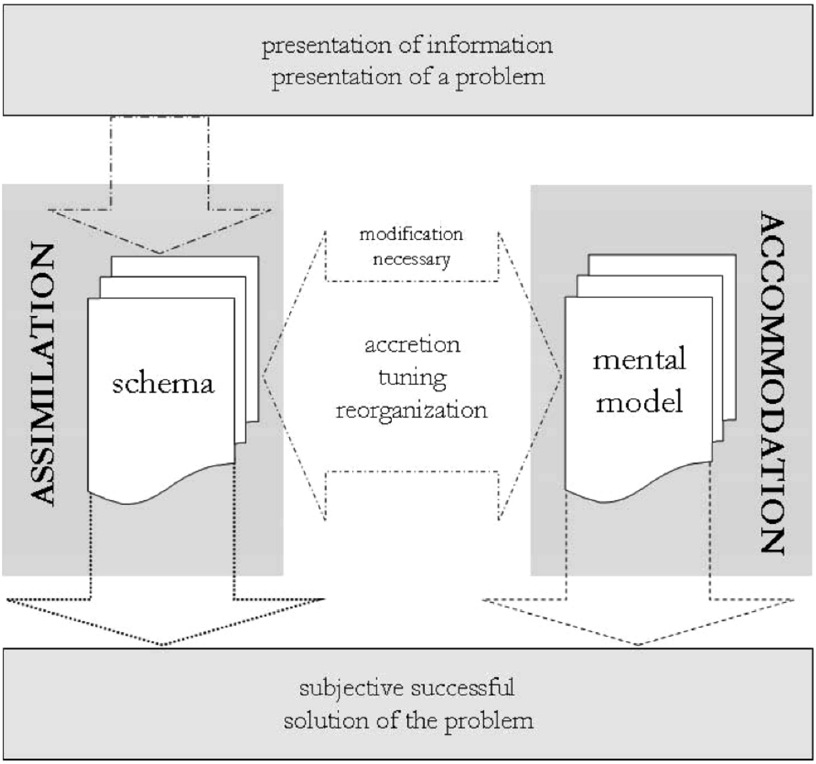

Fig. 3. Activation of schemata and mental models.

However, the *dilemma of a measurement of changing mental models* adds even more complexity to the ambitious project of diagnosing progressing models (see Al-Diban, 2002; Ifenthaler, 2006). Besides the problems discussed above concerning the externalization of mental models, valid and reliable assessment techniques, and adequate statistical procedures, a longitudinal diagnosis of changing models also requires specific situations which activate the construction of mental models.

According to Piaget's epistemology and the basic functions of assimilation and accommodation (Piaget, 1976; Seel, 1991), there is no need for a mental model as long as a person can assimilate new information and activate adequate schemata. If the activated schema does not fit the requirements of the situation exactly it can be adjusted by means of *accretion*, *tuning*, or *reorganization*. However, if accretion or tuning is not successful or if no schema is available, accommodation must take place in order to reorganize and structure a person's knowledge concerning the construction of a mental model (see Figure 3).

Hence, the dilemma of an educational researcher interested in the diagnosis of progressing mental models is how to implement and provide adequate situations and instructions to foster the construction of mental models. In exceptional experimental settings, researchers have implemented a variety of situations and instructions for the investigation of mental models. By providing subjects adequate *conceptual models* before, within, and after a lesson, Mayer (1989) showed that it is more effective to provide the *conceptual model* before or within a lesson than after. Seel & Dinter (1995) also reported a significant effect by providing *conceptual models* before a lesson. In various experiments on logical reasoning, Johnson-Laird (1983, 1989, 2006) showed that certain conclusions yield systematic fallacies that can be explained only by the mental model theory. Additional exceptional experimental situations and instructions are explored in studies by Anzai & Yokoyama (1984), Gentner & Stevens (1983), Al-Diban (2002) and Seel, Darabi, and Nelson (2006).

Although these experimental settings provided interesting results for explaining the complex procedures of progressing mental models, situations emphasizing problem solving tasks grant more detailed insight into the transition of mental models during learning. Such situations can be realized as learning environments with problem solving tasks (Ifenthaler, 2006). However, problems vary in terms of how structured they are, which should be taken into account when implementing such learning environments. According to Jonassen (1997), problems can be classified on a continuum from *well-structured* to *ill-structured*. *Well-structured problems* tend to be fairly stable and have a well-defined initial state, a known goal state or solution, and a limited number of known procedures for solving a class of problems. Therefore, a learning environment with *well-structured problem* solving tasks is not appropriate for the diagnosis of changing models.

In contrast, *ill-structured problems* have no predictable solutions, often possess aspects which are unknown, allow multiple solutions, have a high degree of connectivity among variables, and tend to be highly dynamic as variables change over time while solving the problem. Accordingly, a research design for the diagnosis of changing mental models must contain some sort of *ill-structured problems* in order to activate the *construction*, *modification*, and *reorganization* of mental models.

Another *dilemma of a measurement of changing mental models* has to be considered. As mental models are relatively complex ad hoc constructions, the assessment technique has to be highly dynamic in order to investigate the individual changes precisely. As already discussed above, not every assessment technique can meet these requirements. A researcher would not be able to assess mental models – or their externalizations – with a set of questionnaires (see Al-Diban, 2002). Accordingly, the number of adequate assessment techniques is very limited.

Timing is a further *dilemma of a measurement of changing mental models*. The observation of the learner's externalized mental models "only" before and after instruction does not take important changes during the learning process into account (Ifenthaler & Seel, 2005). Only by conducting an assessment over numerous points in time can the researcher reveal important changes in mental models and implement effective learning environments. Additionally, the gap between numerous measurement points is a very important criterion for designing an experiment for the assessment of changing models. The measurement points should neither be too close to each other, nor should the gaps between them be too large to reveal important changes during the learning process.

A final *dilemma of a measurement of changing mental models* is the learning-dependent progression of mental models from novice to expert models (Johnson-Laird, 1989). As mental models are subjectively plausible ad hoc constructions (Seel, 1991), there is not only one single expert mental model for a specific subject domain. Accordingly, simply comparing the mental models of novices with those of experts will not provide enough information about how a novice has to change to become an expert. Furthermore, not every novice needs to become an expert in a specific subject domain. Therefore, Ifenthaler (2006) and Pirnay-Dummer (2006) introduced *explanation* and *utility models* to describe the learning process between novices and experts more precisely. *Explanation* or *utility models* enable fast, efficient, and successful conclusions. The complexity and range of such models is dependent on (a) the epistemic state of the model builder and (b) the minimum amount of knowledge necessary to successfully solve a problem (see Ifenthaler, 2006; Pirnay-Dummer, 2006). Depending on the situation and subject domain, an *explanation* or *utility model* might be very similar to a novice or to an expert model, or it might mediate between the two. Accordingly, the focus on a learning-dependent progression of mental models from novices to experts neglects a very important condition of mental models – the *explanation* or *utility model*.

Methodological Solutions

It has been made clear that mental models are temporary internal structures created by a person confronted with new information or new situations that cannot be mastered with available conceptual structures or schemata. From a methodological point of view, different dilemmas concerning the diagnosis of changing models must be taken into account. A systematic diagnosis of the progression of mental models focusing on learning and instruction therefore requires at least the following eight principles.

1. A diagnosis of changing models is *embedded in a complex ill-structured problem situation*. Such a situation requires for the problem to be (a) unfamiliar to the person, (b) complex, i.e. contain many different variables, (c) cross-linked, i.e. variables are connected and influence each other, (d) dynamically changing over time, (e) intransparent, i.e. required information is not fully available, as well as to have, (f) staggered effects, i.e. new effects caused by the dynamic development during the problem solving process, and (g) polytely, i.e. different criteria and goals must be taken into account to solve the problem (see Funke, 1991, 1999).

With regard to the requirement of complex ill-structured problem situations, learning environments have been implemented for the diagnosis of changing models (Al-Diban, 2002; Dummer & Ifenthaler, 2005; Ifenthaler, 2006). Empirical findings show that these learning environments must contain authentic situations and that information should be presented in different formats and in a non-linear way. Additionally, the learning environments should provide ways to make a person's structural and semantic knowledge explicit.

2. A diagnosis of changing models is *applied in different subject domains*. The construction of mental models is dependent not only on available heuristics or procedural knowledge, but also on declarative knowledge (Seel, 1991). Consequently, different subject domains require different mental models. Thus, in order to influence the learning-dependent progression of mental models through instruction one must also take the phenomenon being explained into account, and this should not be neglected in experimental research on changing models.

3. The assessment technique for a diagnosis of changing models *allows the construction, modification, and reorganization of mental models*. As discussed above, assessment techniques can hinder people from externalizing their mental models. Additionally, the requirements of the specific research question, the underlying experimental design, and the statistical procedure applied need to fit the characteristics of the chosen assessment technique.

The assessment technique should facilitate the iterative processes of constructing, modifying, and reorganizing mental models. Accordingly, the instrument should assess the subject's changes dynamically at defined points in time during experimentation. A best practice example is the DEEP methodology (Spector, 2004), which enables subjects to change their externalized models dynamically while solving a specific problem.

Al-Diban (2002) and Ifenthaler (2006) used different assessment techniques in their experimental research on changing models. However, in both research designs the subjects were given training on how to externalize mental models using the assessment technique before the actual data collection began. This training resulted in reduced difficulties and uncertainties with regard to the assessment technique during the data collection phase.

4. A diagnosis of changing models *collects mental model data in a longitudinal design*. Only through systematic observation over numerous points in time can important changes in the individually constructed mental models be revealed. The number of measurement points depends on the research question, the subject domain, and the assessment technique. Accordingly, only a strong focus on the theory of mental models enables a sufficient implementation of the research design and the necessary measurement points. In a quasi-experiment conducted by Ifenthaler & Seel (2005), the short gaps between individual measurement points caused frustration in various test subjects. The motivation to continue working in the multimedia learning environment was rather low because the subjects had to externalize their mental models too often from scratch. In order to avoid such problems during experimentation, pre-tests of longitudinal designs for the diagnosis of changing models could provide useful information for optimizing the research design and numbers of measurement points.

Accordingly, a systematic design for the diagnosis of changing models includes (a) more than two measurement points, (b) adequate gaps between measurement points, (c) experimental variation in order to control direct and carryover effects, and (d) if necessary a stability test which provides useful information about long-term effects (see Seel et al., 2006).

5. A diagnosis of changing models *indicates the successive model construction and completion from different perspectives*. Depending on the assessment technique applied, the collected data may give very specific insight into the progression of changing models. In order to investigate the complex processes of model construction and completion over time, multiple perspectives of quantitative and qualitative data should be collected. Pirnay-Dummer (2006) uses two measures derived from graph theory to describe changes in externalized models. With the help of the computer-based and automated SMD Technology (Ifenthaler, 2006), externalized models can be analyzed on three different levels – the *Surface*, *Matching*, and *Deep Structure*. On the *Surface Structure* level, the sum of all propositions (node – link – node) in an individual model is calculated. This level of analysis gives a fast and economical insight into quantitative changes in constructed or dismissed propositions at multiple intervals over time. The complexity of an externalized model is measured at the level of the *Matching Structure* as the

quantity of links of the shortest path between the most distant nodes. In accordance with the theory of mental models, a change in the complexity of a mental model is an important indicator for the process of reorganizing a model. On the *Deep Structure* level, the semantic structure of the externalized models is assessed. The semantic similarity between a subject's externalized model and a domain-specific reference model indicates the changes in semantic quality and the growth of knowledge over various points in time. Multi-perspective approaches such as the SMD Technology enable precise diagnosis of mental model progression and help to describe the complex processes of mental model construction and completion.

6. A diagnosis of changing models *considers characteristics of novice, explanation*, *utility*, *and expert models*. As not every training aims to create *expert models*, a comparison between a subject's mental model and that of an expert is not always appropriate. Many experimental designs should rather focus on individual changes in *novice models* or compare individual models with adequate *explanation* or *utility models* in order to reveal more about effective instructional treatments (see Pirnay-Dummer, 2006; Ifenthaler, 2006).

7. The assessment technique for a diagnosis of changing models *provides valid and reliable quantitative or qualitative data*. Not every assessment technique's output guarantees valid and reliable data for further analysis (Seel, 1999). Therefore, one should conduct extensive tests of validity and reliability before using an assessment technique for the diagnosis of changing models. Johnson et al. (2006) conducted experiments for cross validating different quantitative and qualitative assessment methods. The results indicate promising valid and reliable assessment techniques which are applicable for various research designs. Additionally, Ifenthaler (2006) showed methods for measuring the reliability and validity of an assessment technique with various outside criteria.

8. A diagnosis of changing models *enables a methodologically straightforward analysis and interpretation of the data collected*. Raudenbush & Bryk (2002) introduced *hierarchical linear models* for the research and analysis of individual change data. A two-level hierarchical model is used to represent each test subject's progression over time (level 1) and the variation across subjects, measured either as a characteristic of the person's background or of an experimental variation (level 2). The computer software is easy to handle and supports various types of input data. LISREL (Jöreskog & Sörbom, 1989) is another powerful method for analyzing longitudinal data since it can handle many of the problems of measurement of change discussed above (Seel, 1999). Nevertheless, the best statistical procedures are ineffective unless the research design and the assessment techniques guarantee high validity and reliability.

Conclusions

The implementation of longitudinal research designs for the analysis of changing models involves a series of methodological dilemmas for every educational researcher. Seel's research on mental models addresses many of these dilemmas and has contributed various promising approaches for the solution of the problems discussed above (Seel, 1991, 1997, 1999, 2003). This chapter has explored a wide range of methodological issues associated with the diagnosis of progressing mental models – theoretical foundations, assessment techniques, statistical procedures, and experimental designs.

The theory of mental models (Seel, 1991) enables educational researchers to explain the processes of complex problem solving within a given subject domain. A central goal of educational research on changing mental models is to establish more effective learning environments and instructional techniques in order to improve individual task expertise. Therefore, the experimental research design for changing models should be closely related to the theoretical foundation of mental models.

With the help of computer based and automated assessment techniques, even larger samples of re-presentations of mental models can be assessed in a rapid and economic way. Additionally, these techniques allow different formats of re-representations, such as concept maps (SMD Technology, Ifenthaler, 2006) or natural language expressions (MITOAR, Pirnay-Dummer, 2006). Since the quality of research on changing mental models is highly dependent on the assessment technique one chooses to apply, it is essential that the instruments be valid and reliable.

However, the assessment techniques developed to date have their limitations in that they do not explain internal mental processes in sufficient detail. The diagnosis of the external representation of mental models embodies only a small part of the complex internal processes which occur while problems are being solved (see Al-Diban, 2002). Therefore, well devised experimental designs are necessary if we are to acquire a deeper understanding of the progression of mental models and design more effective learning environments that support the construction of mental models.

The historical synopsis of measurement of change in this chapter provided insight into a long debate about statistical problems with longitudinal data. Nevertheless, more and more researchers have been highlighting the need for further experimental research designs with repeated measures (Ittel & Merkens, 2006; von Eye, 1990a, 1990b). New statistical procedures and computer software (e.g. HLM, Raudenbush & Bryk, 2002) offer a wide spectrum of data analysis for longitudinal data. Accordingly, most of the concerns over problems of measurement of change can be limited to a manageable size (see Willett, 1997).

To conclude, the eight methodological solutions provide a general guideline for planning and conducting future research on the progression of changing mental models. Accordingly, the more detailed our measurement of the information contained in changing mental models is over extended periods of time, the better we can understand the complex processes of mental model construction, which will in turn enable us to implement more effective learning environments and instructional techniques.

References

Al-Diban, S. (2002). *Diagnose mentaler Modelle*. Hamburg: Kovac.

Anzai, Y., & Yokoyama, T. (1984). Internal models in physics problem solving. *Cognition and Instruction, 1*, 397-450.

Arbeitsgruppe empirisch-pädagogische Forschung (2005). *Tagung der Sektion Empirische Bildungsforschung. Veränderungsmessung und Längsschnittstudien. Neue Datengrundlagen für Bildungsforschung und Bildungsmonitoring.* Berlin: Tagungsband.

Bereiter, C. (1963). Some persisting dilemmas in the measurement of change. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 3-20). Madison, WI: The University of Wisconsin Press.

Bock, R.D. (1963). Multivariate analysis of variance in repeated measurements. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 85-103). Madison, WI: The University of Wisconsin Press.

Collins, L., & Horn, J. (1991). Best methods for analysis of change: Recent advances, unanswered questions, future directions. Washington, DC: American Psychological Association.

Collins, L.M., & Sayer, A.G. (2001). *New methods for the analysis of change.* Washington, DC: American Psychological Association.

Cronbach, L.J., & Furby, L. (1970). How should we measure "change" – or should we? *Psychological Bulletin, 74*(1), 68-80.

Dummer, P., & Ifenthaler, D. (2005). Planning and assessing navigation in model-centered learning environments. Why learners often do not follow the path laid out for them. In G. Chiazzese, M. Allegra, A. Chifari, & S. Ottaviano (Eds.), *Methods and technologies for learning* (pp. 327-334). Southampton: WIT Press.

Eckert, A. (1998). *Kognition und Wissensdiagnose. Die Entwicklung und empirische Überprüfung des computerunterstützten wissensdiagnostischen Instrumentariums Netzwerk-Elaborierungs-Technik (NET).* Lengerich: Pabst.

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

Ericsson, K.A., & Simon, H.A. (1998). How to study thinking in everyday life. *Mind, Cluture, and Activity, 5*(3), 178-186.

Fahrenberg, J. (1968). Aufgaben und Methoden der psychologischen Verlaufsanalyse (Zeitreihenanalyse). In K.J. Groffmann & K.H. Wewetzer (Eds.), *Person als Prozeß* (pp. 41-82). Stuttgart: Huber.

Fahrenberg, J., Kuhn, M., Kulick, B., & Myartek, M. (1977). Methodenentwicklung für psychologische Zeitreihenstudien. *Diagnostica, 23*, 15-36.

Fischer, G.H. (1976). Some probabilistic models for measuring change. In D.N.M. De Gruijter, & L.J.T. Van der Kemp (Eds.), Advances in psychological and educational measurement (pp. 97-110). New York: Wiley.

Fischer, G.H., & Formann, A.K. (1982). Veränderungsmessung mittels linear-logistischer Modelle. *Zeitschrift für Differenzielle und Diagnostische Psychologie, 3*(2), 75-99.

Funke, J. (1991). Solving complex problems: Exploration and control of complex problems. In R.J. Sternberg, & P.A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185-222). Hillsdale, NJ: Lawrence Erlbaum.

Funke, J. (1999). Komplexes Problemlösen: Ein Blick zurück und nach vorne. *Psychologische Rundschau, 50*, 194-197.

Gaito, J., & Wiley, D.E. (1963). Univariate analysis of variance procedures in the measurement of change. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 60-84). Madison, WI: The University of Wisconsin Press.

Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance, & D. Galbraith (Eds.), *Knowing what to write. Conceptual processes in text production* (pp. 139-160). Amsterdam: University Press.

Gentner, D., & Stevens, A.L. (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum.

Hanke, U. (2006). *Externale Modellbildung als Hilfe bei der Informationsverarbeitung und beim Lernen.* Freiburg: Universitäts-Dissertation.

Harris, C.W. (1963). *Problems in measuring change.* Madison, WI: The University of Wisconsin Press.

Holtzmann, W.H. (1963). Statistical models for the study of change in the single case. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 199-211). Madison, WI: The University of Wisconsin Press.

Horst, P. (1963). Multivariate models for evaluating change. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 104-121). Madison, WI: The University of Wisconsin Press.

Ifenthaler, D. (2006). *Diagnose lernabhängiger Veränderung mentaler Modelle. Entwicklung der SMD-Technologie als methodologisches Verfahren zur relationalen, strukturellen und semantischen Analyse individueller Modellkonstruktionen. [Diagnosis of the learning-dependent progression of mental models. Development of the SMD-Technology as a methodology for assessing individual models on relational, structural and semantic levels]*. Freiburg: Universitäts-Dissertation.

Ifenthaler, D., & Seel, N.M. (2005). The measurement of change: Learning-dependent progression of mental models. *Technology, Instruction, Cognition and Learning, 2*(4), 317-336.

Ittel, A., & Merkens, H. (2006). *Veränderungsmessung und Längsschnittstudien in der empirischen Erziehungswissenschaft*. Wiesbaden: VS Verlag.

Johnson, T.E., O'Connor, D.L. Spector, J.M., Ifenthaler, D., & Pirnay-Dummer, P. (2006). *Comparative study of mental model research methods: Relationships among ACSMM, SMD, MITOCAR & DEEP methodologies*. Paper presented at the CMC 2006 – Second International Conference on Concept Mapping, September 5-8, San José, Costa Rica.

Johnson-Laird, P.N. (1983). *Mental models. Towards a cognitive science of language, inference, and consciousness.* Cambridge, UK: Cambridge University Press.

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

Johnson-Laird, P.N. (2006). Mental models, sentential reasoning, and illusory inferences. In C. Held, M. Knauff, & G. Vosgerau (Eds.). *Mental models and the mind. Current developments in cognitive psychology, neuroscience, and philosophy of mind* (pp. 27-51). Amsterdam: Elsevier.

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

Jöreskog, K.G., & Sörbom, D. (1989). *LISREL 7. A guide to the program and applications*. Chicago, IL: SPSS Inc.

Kleiter, E.F. (1987). Stand und Perspektive der Veränderungsmessung aus formal-gegenständlicher Sicht. *Zeitschrift für Pädagogische Psychologie, 1*(4), 235-258.

Le Ny, J.-F. (1993). Wie kann man mentale Repräsentationen repräsentieren? In: J. Engelkamp, & T. Pechmann (Eds.), *Mentale Repräsentation* (pp. 31-39). Bern: Huber.

Maedche, A., Pekar, V. & Staab, S. (2002). Ontology learning part one – on discovering taxonomic relations from the web. In: *Proceedings of the Web Intelligence Conference, Berlin* (pp. 301-322). Springer: Berlin.

Mandl, H., & Fischer, F. (2000). Mapping-Techniken und Begriffsnetze in Lern- und Kooperationsprozessen. In H. Mandl, & F. Fischer (Eds.), *Wissen sichtbar machen – Wissensmanagement mit Mapping-Techniken* (pp. 3-12). Göttingen: Hogrefe.

Mayer, R.E. (1989). Models for understanding. *Review of Educational Research, 59*(1), 43-64.

Moskowitz, D.S., & Hershberger, S.L. (2002). *Modeling intraindividual variability with repeated measures data. Methods and applications*. Mahwah, NJ: Lawrence Erlbaum.

Nesselroade, J.R. (1991). Interindividual differences in intraindividual change. In D.M. Collins, & J.C. Horn (Eds.), *Best methods for the analysis of change. Recent advances, unanswered questions, future directions* (pp. 92-105). Washington, DC: American Psychological Association.

Nisbett, R.E., & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.

O'Connor, D.L., & Johnson, T.E. (2004). *Measuring team cognition: Concept mapping elicitation as a means of constructing team shared mental models in an applied setting.* First International Conference on Concept Mapping, September 14-17, 2004, Pamplona, Spain.

Oehrn, A. (1889). Experimentelle Studien zur Individualpsychologie. In E. Kraepelin (Ed.), *Psychologische Arbeiten, 1* (pp. 92-151).

Petermann, F. (1978). *Veränderungsmessung*. Stuttgart: Kohlhammer.

Petermann, F. (1986). Probleme und neuere Entwicklungen der Veränderungsmessung – ein Überblick. *Diagnostica,* 32(1), 4-16.

Petermann, F., Hehl, F.-J., & Schneider, W. (1977). Veränderungsmessung im Rahmen der klassischen Testtheorie. In F. Petermann (Ed.), *Methodische Grundlagen klinischer Psychologie* (pp. 129-154). Weinheim: Beltz.

Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen*. Stuttgart: Klett.

Pirnay-Dummer, P. (2006). *Expertise und Modellbildung: MITOCAR*. Freiburg: Universitäts-Dissertation.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models. Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.

Reiska, P. (2005). *Experimente und Computersimulationen im naturwissenschaftlichen Unterricht: Ein empirischer Vergleich*. Frankfurt am Main: Lang.

Renkl, A., & Gruber, H. (1995). Erfassung von Veränderung: Wie und wieso? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 27*(2), 173-190.

Scheele, B., & Groeben, N. (1984). *Die Heidelberger Struktur-Lege-Technik (SLT). Eine Dialog-Konsens-Methode zur Erhebung Subjektiver Theorien mittlerer Reichweite*. Weinheim: Beltz.

Scheele, B., & Groeben, N. (1988). *Dialog-Konsens-Methoden zur Rekonstruktion Subjektiver Theorien: Die Heidelberger Struktur-Lege-Technik (SLT), konsensuale Ziel-Mittel-Argumentation und kommunikative Flussdiagramm-Beschreibung von Handlungen*. Tübingen: Francke Verlag.

Schvaneveldt, R.W. (1990). *Pathfinder associative networks: Studies in knowledge organization*. Norwood, NJ: Ablex Publishing Corporation.

Seashore, C.E., & Kent, G.H. (1905). Periodicity and progressive change in continous mental work. *Psychological Review Mongr. Suppl. 6*, 47-101.

Seel, N.M. (1991). *Weltwissen und mentale Modelle*. Göttingen: Hogrefe.

Seel, N.M. (1997). Pädagogische Diagnose mentaler Modelle. In H. Gruber, & A. Renkl (Eds.), *Wege zum Können: Determinanten des Kompetenzerwerbs* (pp. 116-137). Bern: Huber.

Seel, N.M. (1999). Educational diagnosis of mental models: Assessment problems and technology-based solutions. *Journal of Structural Learning and Intelligent Systems, 14*(2), 153-185.

Seel, N.M. (2003). Model-centered learning and instruction. *Technology, Instruction, Cognition and Learning, 1*(1), 59-85.

Seel, N.M., & Dinter, F.R. (1995). Instruction and mental model progression: Learner-dependent effects of teaching strategies on knowledge acquisition and analogical transfer. *Educational Research and Evaluation, 1*(1), 4-35.

Seel, N.M., Al-Diban, S., & Blumschein, P. (2000). Mental models and instructional planning. In J.M. Spector, & T.M. Anderson (Eds.), *Integrated and holistic perspectives on learning, instruction and technology: Understanding complexity* (pp. 129-158). Dordrecht: Kluwer Academic Publishers.

Seel, N.M., Darabi, A.A., & Nelson, D.W. (2006). A dynamic mental model approach to examine schema development in performing a complex troubleshooting task: Retention of mental models. *Technology, Instruction, Cognition and Learning, 4*(3-4), 303-330.

Spada, H. (1983). Die Analyse von Veränderungen im Rahmen unterschiedlicher testtheoretischer Modelle. In W.-R. Minsel, & R. Scheller (Eds.), *Brennpunkte der Klinischen Psychologie* (pp. 83-105). München: Kösel-Verlag.

Spector, J.M. (2004). Current issues in new learning. In K. Morgan, C.A. Brebbia, J. Sanchez, & A. Voiskounsky (Eds.) *Human perspectives in the internet society: Culture, psychology and gender* (pp. 429-439). Southampton, Boston: WIT Press.

Spector, J.M., Dennen, V.P., & Koszalka, T.A. (2005). Individual and collaborative construction of causal concept maps: An online technique for learning and assessment. In G. Chiazzese, M. Allegra, A. Chifari, & S. Ottaviano (Eds.), *Methods and technologies for learning* (pp. 223-227). Southampton, Boston: WIT Press.

Stachowiak, H. (1973). *Allgemeine Modelltheorie*. Berlin: Springer.

Stylianou, D. (2002). On the interaction of visualization and analysis: the negotiation of a visual representation in expert problem solving. *Journal of Mathematical Behavior 21*, 303-317.

Stelzl, I. (2005). *Fehler und Fallen der Statistik. Für Psychologen, Pädagogen und Sozialwissenschaftler*. Münster: Waxmann.

Tack, W.H. (1980). Zur Theorie psychometrischer Verfahren. Formalisierung der Erfassung von Situationsabhängigkeit und Veränderung. *Zeitschrift für Differentielle und Diagnostische Psychologie, 1*(2), 87-106.

Taricani, E.M., & Clariana, R.B. (2006). A technique for automatically scoring open-ended concept maps. *Educational Technology, Research & Development, 54*(1), 65-82.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327-352.

van Someren, M.W., Barnard, Y.F., & Sandberg, J.A. (1994). *The think aloud method: a practical guide to modeling cognitive processes*. London: Academic Press.

von Eye, A. (1990a). *Statistical methods in longitudinal research. Volume I. Principles and structuring change*. San Diego, CA: Academic Press.

von Eye, A. (1990b). *Statistical methods in longitudinal research. Volume II. Time series and categorical longitudinal data*. San Diego, CA: Academic Press.

Weber, S., & Schumann, M. (2000). Das Concept Mapping Software Tool (COMASOTO) zur Diagnose strukturellen Wissens. In H. Mandl, & F. Fischer (Eds.), *Wissen sichtbar machen – Wissensmanagement mit Mapping-Techniken* (pp. 158-198). Göttingen: Hogrefe.

Willett, J.B. (1988). Questions and answers in the measurement of change. *Review of Research in Education, 15*, 345-422.

Willett, J.B. (1989). Some results on reliability fort he longitudinal measurement of change: Implications fort he design of studies of individual growth, *Educational and Psychological Measurement, 49*, 587-602.

Willett, J.B. (1997). Measuring change: What individual growth modelling buys you. In E. Amsel, & K.A. Renninger (Eds.), *Change and development: Issues of theory, method, and application* (pp. 213-243). Mahwah, NJ: Lawrence Erlbaum Associates.

Wygotski, L.S. (1969). *Denken und Sprechen. Mit einer Einleitung von Thomas Luckmann. Übersetzt von Gerhard Sewekow*. Stuttgart: Fischer Verlag.

4. Constructing Mental Models during Learning from Science Text

Eye Tracking Methodology Meets Conceptual Change

M. Mikkilä-Erdmann, M. Penttinen, E. Anto and E. Olkinuora

University of Turku, Finland

Abstract: The purpose of this chapter is to examine the process of mental model construction while reading science text. This is done by bringing together three research traditions: conceptual change, text comprehension, and eye-tracking research. The results of our experimental pilot study show the usefulness of the eye-tracking methodology in conceptual change research. This line of study can be beneficial for developing science texts and other learning materials.

Keywords: Conceptual change; eye-tracking; text comprehension.

Introduction

The aim of this chapter is twofold. First, we introduce our theoretical framework in which we integrate two theoretical research traditions: conceptual change (Limón & Mason, 2002) and text comprehension (Kintsch, 1988). Second, we introduce our pilot study that used eye-tracking to record and model the conceptual change process during learning from a text concerning photosynthesis. Eye-tracking has been widely used in cognitive psychology (Rayner, 1998), but is still new to conceptual change research.

The learning and understanding of scientific models requires reorganization of existing domain-specific knowledge structures, such as mental models. This reforming of knowledge structures is called "conceptual change" (Mason, 2001, p. 259). It seems that radical conceptual change is almost impossible to achieve without systematic instruction (Hatano & Inagaki, 1997). Thus, the importance of school teaching when inducing conceptual change in different areas is unquestionable. Despite the long research tradition (see, diSessa, 2006), initiating and supporting conceptual change in the classroom is still very challenging for learners, teachers and the designers of learning materials.

Textbooks dominate science instruction (Hynd, 2001). In everyday school practice, a large part of teaching and learning is based on different kinds of texts aiming at transmitting models about scientific phenomena, of which photosynthesis is a good example. The problem is, however, that the scientific models are often presented in textbooks as if the learners have no prior knowledge or only relevant prior knowledge about the topic to be learned; the learners' possible misconceptions are left ignored. Books seem to offer a ready-made conceptual model, and learners are assumed to understand and make effective use of the presented conceptual model although the model presented in a textbook may not fit well with existing mental models that a learner is able to activate.

This practical problem of everyday school teaching has theoretical background in, for example, Vygotsky's ideas. Vygotsky (1962) identified a gap between naïve and scientific concepts. Because naïve concepts are based on everyday experiences and scientific concepts belong to a qualitatively different conceptual system, combining of the two demands some rearrangement of the learners' existing knowledge structures. The possibility of incompatible informal mental models and new scientific conceptual models may cause difficulties for some learners. One result may be that some new scientific knowledge can remain inert and not applicable outside school in the learner's everyday life. Thus, when studying with the help of textbook texts, it is understandable that the experience of conceptual change can be extremely difficult for learners. No reorganization of knowledge structures occurs, when everyday experiences and prior knowledge do not meet with the scientific definitions of the same phenomenon.

The general theme of this edited volume is the understanding of mental models. Our theoretical target is to understand and model how young learners construct and revise mental models during science text comprehension. Furthermore, our long-term pragmatic aim is to design experimental interventions and science texts that will help young learners construct meaningful and relevant mental models appropriate for learning science. Achieving these goals will promote higher quality learning in school settings. In this chapter, we examine the conceptual change in the context of student processing information contained in science texts. We bring together research on mental models and text comprehension through the notion of conceptual change making use of an eye-tracking methodology.

Conceptual Change

Let us take photosynthesis – one of the most important concepts in the biology curriculum – as an example of conceptual change. Most children seem to have a mental model consisting of the misconception that a plant has multiple external sources of food (e.g., Hatano & Inagaki, 1997; Mason, 1994; Roth, 1990). After instruction, children are often able to show that they have learned some characteristics of the scientific model of photosynthesis on a superficial level, and they might, for example, pass a typical school test. But when children are asked to explain photosynthesis, they reveal in their explanations that they assimilate parts of the new scientific model into their naïve model (see, Vosniadou, 1994a). Children might, for instance, confuse new concepts with their old ideas; thus, they might think that plants take in sun, rain, light, soil or minerals, which then gives them energy to live. Herein lies the problem of science teaching; how to help the learners to see the differences and similarities in their own thinking and the presented scientific model, and how to make the scientific model more accessible, when the naïve model seems to have sufficient explanatory power in everyday experiences.

During recent decades, conceptual change has been studied from at least two different perspectives. Cognitive psychology has studied conceptual change as a developmental process and produced rich descriptions of the naïve theories of young learners concerning different science phenomena (Carey, 1985; Vosniadou, 1994a; Vosniadou & Brewer, 1992). Science instruction, on the other hand, has focused on the practical applications of conceptual change and developed instructional models intended to facilitate conceptual change (Posner, Strike, Hewson & Gertzog, 1982).

In conceptual change research there have been two opposing views on the nature of the learners' naïve ideas (diSessa, 2006). On the one hand, it has been argued that even very young children have domain-specific naïve, theory-like mental models about scientific phenomena. These models are constructed in everyday interactions and used when everyday events need to be understood and predicted (see, e.g., Vosniadou, 1994a). Vosniadou (1994a) argues that conceptual structures can be altered either by adding more information to them (so-called *enrichment*) or by replacing old conceptions with new ones (so-called *revision*). On the other hand, prior knowledge may consist of fragmented elements. These elements can be used in forming coherent, scientific models (diSessa, 2006). According to this view, as Smith, diSessa and Roschelle (1993) and Siegler (1996) argue conceptual change is not a sudden but more of a gradual process; it is *evolutionary* rather than *revolutionary* (Sinatra, 2002, p. 190). Still, all theories of conceptual change agree that conceptual change can happen on more than one level (Duit & Treagust, 2003; White, 1994).

Piaget's (1950, 1985; see, Schnotz & Preuss, 1997; Vosniadou, 1994b) theory of intellectual development includes the concept of *cognitive equilibrium*. The cognitive system tries to avoid conflict between newly presented information and already existing knowledge. Cognitive equilibrium has later on been used to explain routes to conceptual change. Hence, when a conflict occurs, the attempt to restore cognitive equilibrium may lead to conceptual change. Hewson (1981) and Posner and colleagues (1982) consider this so-called *cognitive conflict* to be a necessity in experiencing conceptual change. According to them, one has to be discontented with one's old conceptions before they can be altered. In the light of new experiences, old conceptions are no longer sufficient; thus, there is a conflict between existing knowledge and new information. Accepting new information requires that it is considered intelligible, plausible and/or fruitful (Posner et al., 1982).

Both cognitive psychologists and science instructors consider the inducing of cognitive conflict to be an important element of conceptual change (Guzzetti, Snyder, Glass & Gamas, 1993; Kang, Scharmann & Noh, 2004). Nonetheless, producing cognitive disequilibrium does not always result in desired conceptual change; what will surely result is an attempt to achieve cognitive equilibrium, but this may involve an undesired conceptual change. One can also try to achieve cognitive equilibrium by forming new erroneous theories or by creating synthetic models where there are elements from both naïve and scientific models. (Vosniadou, 1994b). Thus, merely producing cognitive conflict is not sufficient when aiming for a particular conceptual change. It has been suggested that motivational factors may also account for lack of success in resolving cognitive conflict to promote learning (Olkinuora & Salonen, 1992). According to studies conducted by Kang and colleagues (2004), cognitive conflict can be one of the important factors of conceptual change, but it is not a necessary precondition.

The theory by Posner and colleagues (1982) requires that one is conscious of the nature of ones thinking; only after that can different conceptions be compared to each other and there are grounds for experiencing cognitive conflict and conceptual change. This awareness of the theoretical nature of ones thinking is called *metaconceptual awareness* (Vosniadou, 1994a). According to conceptual change research this seems to be a necessary, though not sufficient, precondition for conceptual change. In order to change something, one has to both realize the need for it and be willing to do so (Limón, 2001, p. 359; White & Gunstone, 1989).

Mental Models and Text Comprehension

A common notion about text comprehension is that a reader constructs mental representations during text processing. For instance, Kintsch and van Dijk (1978) assumed in their earliest model that the meaning of a text could be parsed into semantic units, so-called *propositions*, which are interconnected according to coherence relationships. Comprehension was perceived adding together semantic units. Inferences were seen to have the function of bridging coherence gaps within the text. With this model, it is possible to model the process of text comprehension adequately to some extent, but only as long as no misunderstanding occurs and the reader is not forced to reinterpret what has read before (Mandl & Schnotz, 1987).

In contrast to this additive-elementaristic view, the later models of text comprehension attempted to describe text comprehension with the help of holistic structures (Johnson-Laird, 1980, 1983; van Dijk & Kintsch, 1983). For instance, van Dijk and Kintsch (1983) modified and extended their earlier model into the so-called *Construction-Integration (CI)* model (Kintsch, 1988; see also Kintsch, 1986, van Dijk & Kintsch, 1983). According to the CI model, the process of text comprehension proceeds in a two-phased cycle. In the *construction* phase, propositions activate in the reader's mind a network of related concepts. This network consists of both relevant and irrelevant concepts. In the next phase, *integration*, this network is established; new propositions that are consistent with the reader's prior knowledge are left active and inapplicabilities are discarded. (Kintsch, 1988.)

In the CI model, there is a distinction between two types of representations built while reading the text (Kintsch, 1988). The *text base* is constructed from propositions and expresses the semantic content of the text at both local and global level. The *situation model*, on the other hand, is the mental representation of the situation described by the text (Kintsch 1986). The situation model may be, for example, the readers' mental map of a town described in the text, an arithmetic structure derived from the text, or, as in our studies, a scientific model of photosynthesis constructed from the content given in the text. These two representations, the text base and the situation model, are not independent of each other, but each has its own characteristics and each supports certain types of learning. (Kintsch, 1986.)

Although these theoretical concepts are all hypothetical, there is empirical evidence for the dichotomy of the text base and the situation model. Kintsch (1986) draws the pedagogical conclusion that if the learner has constructed a text base during reading, the learner can remember the text itself; however, understanding the content of the text demands the construction of a correct situation model. Thus, there is a difference between learning *the* text and learning *from the* text. Also Johnson-Laird (1980, 1983) assumes that in reading the text, besides the propositional representation of the text a mental model is constructed. The mental model is a holistic structure representing the content in a directly analogous manner instead of an indirect digital manner. In text comprehension, mental models are constructed on the basis of propositional representation and of prior knowledge. Thus, comprehension of a text may go beyond the immediate content of the text, and so the text itself loses its individuality, and its information content is integrated into some larger structure as Kintsch (1988) has proposed.

Despite the fundamental dichotomy of the two mental representations in the presented theories, later on it was argued that the reader forms one mental model instead of two separate ones while reading, and this mental model can be observed from two dimensions; the text base and the situation model as presented above (McNamara, Kintsch, Butler Songer & Kintsch, 1996). Thus, these approaches by van Dijk and Kintsch (1983) and Johnson-Laird (1980, 1983) may be called holistic since they assume that from the very beginning of the comprehension process a holistic mental structure or mental model is constructed, evaluated and eventually revised on the basis of the text and prior knowledge. Hence, mental inferences are perceived as less text- dependent in that they also not only serve to fill the coherence gaps in the text but are used to enrich and elaborate the mental model of the learner. Inferences function at the conceptual level rather than the linguistic level (Kintsch, 1986; Mandl & Schnotz, 1987).

One of the most important factors in text processing is the reader's working memory (Kintsch & van Dijk, 1978). The concept of *cognitive overload* is closely related to working memory (Armand, 2001; Sweller, 1994). Cognitive overload occurs when the reader's working memory capacity is not sufficient for the comprehension task, and thus the processing of text becomes much of a strain. When the reader has low prior knowledge, a coherent text might lessen the cognitive load and the effective processing of the contents could succeed.

Text processing is also influenced by text coherence. A coherent text facilitates the remembering of details and the forming of a knowledge structure. (McNamara et al., 1996). When, for example, the reader has low prior knowledge of the presented subject, the text should be coherent enough so that the reader's working memory is not overloaded. On the other hand, high prior knowledge readers have been found to benefit from a text with low coherence, because this forces them to process more actively during reading (Schnotz, 1993). Thus, the results are contradictory. In general, it can be stated that a text that increases mental activity during reading, will benefit all readers irrespective of the level of prior knowledge (Gilabert, Martinez & Vidal-Abarca, 2005).

Thus, it seems that both the reader with his/her personal characteristics and the text influence the process of text comprehension. The reader's prior knowledge of the text's subject has an impact on the understanding of what is being read, and thus on the experience of conceptual change. Apart from this, text comprehension can also be influenced by text coherence.

Eye-tracking Research

Previous studies have demonstrated the connection between the reader's focus of attention and eye movements while reading (Hyönä, 1998). The reading process can be examined extremely accurately with eye-tracking methodology. Hence, the reader's cognitive processes during reading can be inferred on the basis of the reader's eye movements. Previously, the studies on eye tracking have focused mainly on the recognition of words or syllables, but according to Hyönä, Lorch and Rinck (2003), this methodology can also be beneficial when text processing is studied on a macro level.

The reading process is not continuous but consists of short stops, *fixations*, and the transitions from one fixation to another, so-called *saccades*. An adult reader's length of fixations during reading varies from 200 to 300 milliseconds, whereas the saccades are remarkably shorter (Rayner, 1998). It is commonly admitted that the reader acquires information only during fixations (e.g., Kaakinen, Bertram & Hyönä, 2004).

Problems during reading can be caused either by difficulties in word or letter recognition, or in understanding the contents of the text. *Regression* is often considered to be a symptom of problems in the reading process. Hyönä, Kaakinen and Lorch (2002) argue that fixations moving forward and regressions inside a sentence refer more to mechanical reading skills. According to them, eye movements in a wider range, in and out of sentences can indicate problems in understanding the content. (Hyönä et al., 2002). Nevertheless, if the reader experiences difficulties in text comprehension, returning to the previous text is not the only solution to the problem. The reader might continue his/her reading and expect the following text to clarify the part s/he did not understand. Naturally, the reader can also rely on his/her working memory and solve problems of understanding without showing it in the reading process at all. Thus, situations where problems of understanding the contents of the text occur are not always visible. (Hyönä et al., 2003).

Regression in text processing can be observed from two different perspectives; from the target of re-reading or from its starting point. When the reader goes back to a certain part of text and re-reads it that text unit collects what are called *look backs*. Look backs might occur when the reader needs to return to a critical text unit, a unit that caused cognitive difficulties and had content that needed to be clarified. On the other hand, there is always a starting point – e.g., a word or a sentence – for the returnings to previous text. When the reader leaves a text unit to read previous text again, this text unit collects so-called *look froms*. In this way, the critical text unit can be seen as the one that is evoking text processing immediately when cognitive difficulties with text comprehension occur. When look backs and look froms are studied, the observer actually tracks down fixations that have landed on a certain area. The role of a fixation is defined on the basis of the directions of the surrounding saccades (see, Hyönä et al., 2002).

Look back time describes the amount of time spent on re-reading a critical text unit. Look back time is the sum of all fixations that land on a part of text after its first reading (Hyönä et al., 2003). *Look from time* consists of the fixations that land on previous text in the middle of reading the critical text unit (Anto & Penttinen 2006; Hyönä et al., 2002). Look from time starts when the reader leaves an unfinished sentence and ends when the reader returns to the sentence and reads on. A longer look from time can be seen as longer look back times in the sentences to which the reader returns. The concepts describe problem solving in reading process from two perspectives; look froms and look from time tell us where the reader started to process the text, while look backs and look back time describe where the reader returned. Nonetheless, the concepts look from time and look back time are not exactly mirror images, since the look from time of a sentence can be divided into look back times for several sentences.

The concepts used in eye-tracking research still lack unity, though there have been attempts to simplify the terminology (Hyönä et al., 2003). Here we present some variables that can be examined in the reading process.[5] *Total fixation time* describes the time used in reading the text. The reading process can be examined more closely with such variables as *regression path duration* and *selective regression path duration*. The regression path duration of a sentence includes all fixations during the reading of a particular sentence. Thus, it also includes those fixations that land on previous text, i.e., look backs, besides the sentence in question. The selective regression path duration of a sentence only includes those fixations that land on the sentence itself. Thus, selective regression path duration can be

[5] The variables and their definitions are based on the software (DataViewer) used in the experimental study that will be presented later in this chapter.

counted by deducting from regression path duration the fixations that land on previous text, that is, look from time.

Ways of solving cognitive difficulties during reading: (Hyönä et al. 2003)	**Indicators used in eye-tracking research:**
Slow reading	Fixation time
Instant re-reading of the critical part of text	Selective regression path duration
Reading on	Linear reading
Reading on and *turning back to the critical part* of text	Look back time
Returning to previous text *from the critical part* of text	Look from time

Experiencing Cognitive Conflict while Reading Science Text

We believe that combining conceptual change research with theories on text comprehension can bring us nearer to the process of conceptual change. When defining conceptual change in text comprehension process through the CI model, conceptual change can be seen as rearranging and supplementing the knowledge network. One of the possible paths to conceptual change, cognitive conflict (i.e., the possible mismatch between prior knowledge and the scientific models to be learned from the text), can most likely be experienced at different levels of comprehension. On a propositional level, cognitive conflict can be placed between a newly faced, naively interpreted proposition and the more or less scientific-like text base (Mikkilä-Erdmann, Anto, & Penttinen, 2006).

If naïve misconceptions dominate the text comprehension process, parsing a new proposition might activate units in the reader's knowledge network that are false when compared with the scientific information presented in the text. Cognitive conflict or conceptual change cannot occur, since the text comprehension process is misguided at an early phase. Misconceptions can in some cases lead the process of understanding in that the scientific-like conceptions are left ignored, or the reader adapts some scientific elements to his/her naïve theory and creates a so-called synthetic model (see, Mikkilä-Erdmann, 2002; Vosniadou, 1994a). In order to make the experiencing of cognitive conflict in the text comprehension process possible, the reader must be able to give up interpreting propositions using only naïve misconceptions and build the situation model through the information offered by the text instead. This requires a conscious guiding of the text comprehension process, and also favorable motivational-emotional circumstances.

Prior knowledge is an important factor both in text comprehension (Armand, 2001) and conceptual change (Mayer, 2002). Previous studies have proved that while studying from a text, activating the reader's prior knowledge and making the reader aware of the differences between naïve conceptions and the scientific model – that is, inducing cognitive conflict and awakening metaconceptual awareness – will benefit comprehension (Alvermann & Hague, 1989; Hynd, 2001; Hynd & Alvermann, 1986). According to a commentary of conceptual change theories by Mayer (2002), prior knowledge can both hinder conceptual change and create the conditions for it by working as building blocks when constructing new models. Also Alvermann and Hague (1989) argue that merely activating prior knowledge is not enough, since it can hinder the acceptance of the scientific model.

The effects of these so-called *refutational texts* on learning outcomes have been studied during the last couple of decades. In their studies, Alvermann and Hague (1989) used an introductory text passage where readers were warned about the possible inconsistencies between naïve and scientific models. This was done to awake the reader's awareness of his/her own thoughts on the subject. Mikkilä-Erdmann (2002), on the other hand, created a text with *metaconceptual text units* that were embedded within the text. These units were designed to support the reader's metaconceptual awareness by challenging his/her possible misconceptions and again inducing cognitive conflict by pointing out the differences with scientific models. In eye-tracking research the strong effect of *reading perspective* on the reading process has been confirmed (Kaakinen, 2004). It is suggested here that activated metaconceptual awareness could work the same way than reading perspective. Metaconceptual awareness could guide the reader to form the text base by constructing propositions according to the text and not only according to the readers' misconceptions. (Mikkilä-Erdmann et al., 2006.) Hence, the result would be a coherent, meaningful and complemented scientific mental model, and the reader would have solved the problem of cognitive equilibrium so that it leads to conceptual change.

The problem with earlier studies on inducing conceptual change through text has been that the focus has been mainly on the learning outcomes, whereas the process itself has been left unexplored. This has been mostly due to the limitations of the used methodologies. In this chapter, the problems of constructing mental models in text comprehension process will be examined with eye-tracking research methodology. This methodology offers the possibility to observe the reading process more closely.

When tracing cognitive conflict in text comprehension process, the critical moment is when a reader has problems constructing either the text base or the situation model. Problems on the text base level can occur when, for example, naively interpreted conceptions do not fit in the more or less scientific-like text base. When forming the situation model, the reader might have difficulties if new information contradicts with the reader's prior knowledge. According to van Dijk and Kintsch (1983), while constructing the text base, the reader has to compare a new proposition to the already formed text base. This can be done by relying on working memory or, if this is not enough, by re-reading previous text for support. The term presented in this chapter, look from time, describes this returning to previous text precisely when the problem occurs. Solving difficulties experienced when forming a situation model, on the other hand, could demand different strategies, such as returning to the critical text unit after reading on. Thus, it is argued here that look from time would be beneficial when observing the text processing more on a text base level, and look back time might be better in observing the cognitive problems occurring during the formation of the situation model.

Experimental Study

The target of our research project is to model cognitive processes such as experiencing cognitive conflict during science text comprehension. We are conducting experiments with the aim of making the process of conceptual change visible. Furthermore, we are investigating instructional tools, such as texts, that could facilitate conceptual change in classroom situations. We present the pilot study of this research stream next.

Research Objectives and Method

The target of the experimental study presented here was to find out if cognitive conflict experienced during text processing can be traced using eye movement research methodology (Anto & Penttinen, 2006; Anto, Penttinen, & Mikkilä-Erdmann, 2007). Based on earlier findings, the hypothesis was that the experience of cognitive conflict in the reading process might cause more text processing. In this study, such variables as *total fixation time, selective regression path duration* and *look from time* were examined in each participant's reading process. The other goal of this study was to examine the effect of text type on inducing cognitive conflict. The eye-movement methodology was used to examine the reading processes of two different text types.

Thirty sixth-graders participated in this study. The participants were randomly divided into two treatment groups. In a laboratory setting, the participants read a science text concerning photosynthesis and did some written exercises. Written pre- and delayed posttests were used to identify changes in the participants' conceptions on photosynthesis. The pre- and delayed posttests were identical and were carried out in the participants' own classrooms.

Two different treatment texts were used. The *refutational text* was based on Mikkilä-Erdmann's (2002) studies about children's naïve conceptions of photosynthesis. The text included so-called metaconceptual text units, which were planned to both activate the readers' metaconceptual awareness and point out the possible differences between the readers' misconceptions and the scientific model (e.g., "It's important to understand that a plant does not take ready-made food from the soil and therefore it does not eat."). Metaconceptual text units were assumed to cause more text processing, and help when experiencing cognitive conflict. The *explanatory text* was otherwise identical to the refutational text but the metaconceptual text units were replaced with other sentences. The refutational text was slightly longer than the explanatory text (2540 vs. 2369 letters and punctuations marks), due to the necessary added words in the metaconceptual text units.

In a laboratory setting every participant read two texts on a computer screen. Reading processes were recorded using EyeLink II (SR Research Ltd., Toronto, Ontario, Canada). First, all participants read a control text on a related subject. The purpose was to make the participants familiar with the head-mounted eye-tracking camera. After that, every participant read the treatment text. The length of time allowed for reading the texts was not limited.

The reading processes were examined both on the word and sentence level through multiple variables. Based on the theoretical background and the preliminary analysis, some variables were chosen for a more accurate analysis. *Total fixation time* and *selective regression path duration* were converted into milliseconds per sign. Letters and punctuation marks were counted as signs. *Look from time* was converted into milliseconds per sentence, since the whole sentence was considered to enable the possible processing, while the length of the sentence was of no importance. The eye-movement data were analysed by DataViewer.

The written pre- and delayed posttests consisted of text-based *fact finding* (e.g., "What are stomata?") (see, van Dijk & Kintsch, 1983) and so-called *generative questions* (e.g., "Do a carrot and a rabbit get energy the same way?") (see, Vosniadou, 1994a). Correct answers to the latter demanded a scientific mental model on photosynthesis. The scales for rating the written answers were formed on the basis of the participants' answers. Thus, the questions were rated either from 0 to 3 or from 0 to 4. The maximum score always demanded a scientific answer. The interrater reliability was 87.6%. Due to the two different scales, the ratings were changed into percentages, the maximum score (3 or 4) being 100%.

Results

The two treatment groups (n=15/group) were found comparable when the participants' scores in the pretest, the grade in science, the GPA of theoretical subjects, and the total fixation time in reading the control text were examined using the independent samples T-test. Because of unsuccessful eye-tracking data five participants (2 refutational text and 3 explanatory text readers) had to be excluded.

The effect of text type on the change of scores in generative questions and the reading process was tested with the repeated-measures ANOVA. No significant interaction effect was found. The two treatment groups were then combined for the following analysis.

The participants' level of conceptual change was defined on the basis of the performance in generative questions from pretest to delayed posttest. The participants were divided into three groups. Group 1 showed either no change in their scores or even occasional regression, and thus its members did not experience a change in their conceptions about photosynthesis. Group 2 slightly improved their scores from pretest to delayed posttest, the maximum improvement being 13.33%. Thus, for the members of group 2, some elements of photosynthesis were more scientific-like after this intervention, but no conceptual change as such occurred. Group 3 improved their scores by at least 20%, and were seen to have experienced conceptual change to some extent.

The reading process of the treatment text was compared in relation to the participants' level of conceptual change. An independent samples T-test was conducted to compare the reading processes of groups 1 and 3, the no conceptual change and conceptual change groups. Group 2 did not fulfill the null hypothesis of the normal distribution and was thus left out of the following analysis.

Table 1. Means and standard deviation of groups 1 and 3 for total fixation time, selective regression path duration and look from time

Groups	N	Total fixation time *(ms/sign)* M/SD	Selective regression path duration *(ms/sign)* M/SD	Look from time *(ms/sentence)* M/SD
1 no conceptual change	12	63,8/15,8	61,9/14,8	142,6/102,6
3 conceptual change	4	75,5/17,9	68,7/15,0	432,3/223,4

Total fixation time (ms/sign) was slightly longer for those who experienced conceptual change to some extent, although the difference is not significant, t(14)= –1,248; p> 0,05 (see, table 1). Selective regression path duration (ms/sign) is slightly longer for the conceptual change group, but the difference is not significant, t(14)= –0,794; p> 0,05. The most important finding was that look from time (ms/sentence) was significantly longer for the conceptual change than the no conceptual change group, t(14)= –3,644; p< 0,01. The results suggest that those readers who experienced conceptual change spent more time in re-reading previous text than the readers who did not experience conceptual change. Thus, it seems that look from time may work as an indicator of cognitive conflict in science text processing.

General Discussion

The theoretical aim of this chapter is to combine two research traditions, conceptual change and text comprehension. This is done in order to investigate how learners experience conceptual change during learning from science text. Both research traditions deal with cognitive processes, which are not visible or easily observable, and are mostly examined on the basis of performance measurement from pretest to posttest. Thus, the results of different interventions are well documented, but the process of conceptual change (or not succeeding in it) has been left unexplored.

On the basis of our explorative study, we consider eye-tracking an appropriate method for examining conceptual change processes during text comprehension due to its long research tradition in examining cognitive processes, and its extremely accurate online results. Although eye tracking gives us new possibilities for investigating cognitive processes online, this is not enough when examining the construction of mental models. A multi-method approach seems necessary; for example, the think aloud method, stimulated recall interviews etc. could complete and validate the research design. Nevertheless, on the basis of our explorative study we found possible indicators of cognitive conflict in the reading process, and progressed a little in the analysis of the conceptual change process during text comprehension. There are still many challenges – both theoretical and methodological – that have to be dealt with in future studies.

The importance of well-written texts as one component of school teaching is unquestionable. Thus, studies on both the texts and the ways they are studied are extremely relevant, when trying to help learners to achieve a true understanding of difficult concepts. In their studies, Hyönä et al. (2002) categorized four reading strategies[6], which differ in, for example, the time used for reading and the way the text (headlines, beginnings of paragraphs etc.) is processed. The ability to write summaries after reading a text varies between different strategies; those who processed the text systematically produced the best summaries. (Hyönä et al., 2002). These personal ways of processing a text bring a new perspective to our studies on conceptual change in text comprehension. The role and stability of reading strategies have yet to be examined more closely. They also need to be compared to the possibilities to affect the reading process by building the text in a way that would promote beneficial reading strategies, enabling a better understanding of the subject, and thus create a good basis for experiencing conceptual change. In addition to the reading strategies, the reading perspective already discussed in this chapter

[6] The strategies are named as fast linear readers, nonselective reviewers, slow linear readers and topic structure processors (Hyönä et al., 2002).

(e.g., Kaakinen, 2002) offers interesting possibilities when trying to induce conceptual change in the text comprehension process.

As mentioned above, textbook texts still need to be investigated and developed further. This is also one of the most important applications of our line of study. The development of the refutational text design must be continued and tested with larger samples. For a learner, the awareness of a cognitive conflict is an important factor in promoting conceptual change. Our future task is to reflect on the question whether it is possible to provide the readers with more effective metaconceptual monitoring through text design. Hence, we have to develop and test a text that would lead the learner from enrichment to revision of mental models by supporting the readers' metaconceptual awareness. Later on, relevant pictures and other learning materials (e.g., multimedia learning environments) will be added to the created text design and their interaction will be studied.

However, in order to promote radical conceptual change at school, a textbook text and integrated pictures are not enough. The whole instructional process should be planned so that it supports conceptual change (Vosniadou & Ioannides, 1998). The main idea would be to design the whole instructional environment so that it would be possible to make the misconceptions visible to the learners and teachers. This would suggest a knowledge-building community in which learners would have the possibility to become aware of their prior knowledge in discussions and collaborative learning settings (Chann, Burtis & Bereiter, 1997). The role of text in a knowledge-building community or in problem-based learning has been left unexplored. Texts are already used as tools, though much more information on the text comprehension process is needed.

As Seel (2004) sums up, in a learning situation, besides the learner's preconceptions, also his/her motivation towards the task has to be taken into consideration. There has been criticism of conceptual change research being too "cold" and not taking learners' motivational-emotional factors into account (Sinatra & Mason, 2007; Sinatra & Pintrich, 2003). It has been suggested that the students' motivation, the way they approach a learning task, also plays an important role in experiencing conceptual change (Hynd, Holschuh & Sherrie, 2000). There is a research tradition on *motivational orientations*, which has shown that, at school, the learners perform in very different ways depending on their orientations (Olkinuora & Salonen, 1992). Motivational orientations seem to be related to the responsiveness or resistance to conceptual change; the orientations of the learners will either lend support or function as a filter in the conceptual change process (Mikkilä-Erdmann & Lepola, 2004). Hence, our global target would be to design instructional interventions which not only try to promote conceptual change but also foster task orientation. The motivational factors of the learning situation also have to be taken into consideration, when planning the way to present the learning task to the learner. The complex process of learning has to be studied as a whole, since no text or any other single factor of school teaching alone can produce the best outcomes.

References

Anto, E., & Penttinen, M. (2006). *Kognitiivinen konflikti ja silmänliikkeet oppikirjatekstin lukemisprosessissa*. [Cognitive conflict and eye movements in reading textbook text.] Master's Thesis, University of Turku, Department of Teacher Education.

Anto, E., Penttinen, M., & Mikkilä-Erdmann, M. (in press). Käsitteellinen muutos ja silmänliikkeet tekstiymmärtämisprosessissa. [Conceptual change and eye movements in text comprehension process.] *Psykologia*.

Alvermann, D.E., & Hague, S.A. (1989). Comprehension of counterintuitive science text: Effects of prior knowledge and text structure. *Journal of Educational Research, 82*, 197-202.

Armand, F. (2001). Learning from expository texts: effects of the interaction of prior knowledge and text structure on responses to different question types. *European Journal of Psychology of Education, 16*, 67-86.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT press.

Chann, C., Burtis, J., & Bereiter, C. (1997). Knowledge building as a mediator of conflict in conceptual change. *Cognition and Instruction, 15*, 1-40.

diSessa, A.A. (2006). A history of conceptual change research. Threads and fault lines. In R.K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 265-281). Cambridge: Cambridge University Press.

Duit, R., & Treagust, D.F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education, 25*, 671-688.

Gilabert, R., Martinez, G., & Vidal-Abarca, E. (2005). Some good texts are always better: Text revision to foster inferences of readers with high and low prior background knowledge. *Learning and Instruction, 15*, 45-68.

Guzzetti, B.J., Snyder, T.S., Glass, G.V., & Gamas, W.S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly, 28*, 117-159.

Hatano, G., & Inagaki, K. (1997). Qualitative changes in intuitive biology. *European Journal of Psychology of Education, 12*, 111-130.

Hewson, P.W. (1981). A conceptual change approach to learning science. *European Journal of Science Education, 3*, 383-396.

Hynd, C.R. (2001). Refutational texts and the change process. *International Journal of Educational Research, 35*, 699-714.

Hynd, C., & Alvermann, D.E. (1986). The role of refutation text in overcoming difficulty with science concepts. *Journal of Reading, 29*, 440-446.

Hynd, C., Holschuh, J., & Sherrie, N. (2000). Learning complex scientific information: motivation theory and its relation to student perceptions. *Reading and Writing Quarterly, 16 (1)*, 23-57.

Hyönä, J. (1998). Lukeminen havaintoprosessina. [Reading as a perceptual process.] *Psykologia, 4-5*, 276-282.

Hyönä, J., Kaakinen, J.K., & Lorch, R.F. Jr. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Education 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.

Johnson-Laird, P.N. (1980). Mental models in cognitive science. *Cognitive Science, 4*, 71-115.

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

Kaakinen, J. (2004). *Perspective effects on text comprehension: Evidence from recall, eyetracking and think-alouds*. Annales Universitatis Turkuensis, ser. B, tom. 275.

Kaakinen, J., Bertram, R., & Hyönä, J. (2004). Lukijoiden silmänliiketutkimus. Sanojen tunnistamisesta tekstinymmärtämisstrategioihin. [Eye tracking research of readers. From word recognition to text comprehension strategies.] *Psykologia, 2,* 92-103.

Kang, S., Scharmann, L.C., & Noh, T. (2004). Reexamining the role of cognitive conflict in science concept learning. *Research in Science Education, 34*, 71-96.

Kintsch, W. (1986). Learning from text. *Cognition and Instruction, 3 (2)*, 87-108.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163-182.

Kintsch, W., & van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363-394.

Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction, 11*, 357-380.

Limón, M., & Mason, L. (2002). Introduction. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change. Issues in theory and practice* (pp. 15-20). Dordrecht: Kluwer Academic Publishers.

Mandl, H., & Schnotz, W. (1987). New directions in text processing. In E. De Corte, H.G.L.C. Lodewijks, R. Parmentier & P. Span (Eds.), *Learning and instruction. European research in an international context: Vol. 1* (pp. 321-338). Leuven/Oxford: Leuven University Press/Pergamon Press.

Mason, L. (1994). Analogy, metaconceptual awareness and conceptual change. A classroom study. *Educational Studies, 20 (2)*, 267-291.

Mason, L. (2001). Introduction. In L. Mason (Ed.), Instructional practices for conceptual change in science domains. *Learning and Instruction, 11 (4-5)*, 259-262.

Mayer, R.E. (2002). Understanding conceptual change: A commentary. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change. Issues in theory and practice* (pp. 101-111). Dordrecht: Kluwer Academic Publishers.

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

Mikkilä-Erdmann, M. (2002). *Textbook text as a tool for promoting conceptual change in science*. Annales Universitatis Turkuensis, ser. B, tom. 249.

Mikkilä-Erdmann, M., Anto, E., & Penttinen M. (2006). Learning from science text – Eye movement measurement as a method for investigating science text comprehension. Poster presentation. In *Proceedings of EARLI SIG 2 2006*, August-September 2006.

Mikkilä-Erdmann, M., & Lepola, J. (2004). *Responsiveness vs. resistance to conceptual change concerning photosynthesis*. Paper presented at EARLI Conceptual Change SIG meeting. Delphi, Greece, May 2004.

Olkinuora, E., & Salonen, P. (1992). Adaptation, motivational orientation, and cognition in a subnormally performing child: a systemic perspective for training. In B.Y.L. Wong (Ed.), *Contemporary intervention research in learning disabilities. An international perspective* (pp. 190-213). New York: Springer-Verlag.

Posner, G.J., Strike, K.A., Hewson, P.W., & Gertzog, W.A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*, 211-227.

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

Roth, K. (1990). Developing meaningful conceptual understanding in science. In B. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 139-175). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schnotz, W. (1993). Brief research report. Adaptive construction of mental representations in understanding expository texts. *Contemporary Educational Psychology, 18*, 114-120.

Schnotz, W., & Preuss, A. (1997). Task-dependent construction of mental models as a basis for conceptual change. *European Journal of Psychology of Education, 12*, 185-211.

Seel, N.M. (2004). Model-centered learning environments: Theory, instructional design, and effects. In N.M. Seel & S. Dijkstra (Eds.), *Curriculum, plans, and processes in instructional design. International perspectives* (pp. 49-74). Mahwah, NJ: Lawrence Erlbaum Associates.

Siegler, R.S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Sinatra, G.M. (2002). Motivational, social, and contextual aspects of conceptual change: a commentary. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change. Issues in theory and practice* (pp. 187-197). Dordrecht: Kluwer Academic Publishers.

Sinatra, G.M., & Mason, L. (in press). Beyond knowledge: learner characteristics influencing conceptual change. In S. Vosniadou (Ed.), *Handbook on conceptual change*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sinatra, G.M., & Pintrich, P.R. (2003). The role of intentions in conceptual change. In G.M. Sinatra, & P.R. Pintrich (Eds.), *Intentional conceptual change* (pp. 1-18). Mahwah, NY: Lawrence Erlbaum Associates.

Smith, J.P., diSessa, A.A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences, 3(2)*, 115-163.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295-312.

van Dijk, T., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Orlando, FL: Academic Press.

Vosniadou, S. (1994a). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*, 45-69.

Vosniadou, S. (1994b). Introduction. *Learning and Instruction, 4*, 3-6.

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., & Ioannides, C. (1998). From conceptual development to science education: A psychological point of view. *International Journal of Science Education, 20 (10)*, 1213-1230.

Vygotsky, L.S. (1962). *Thought and language*. Cambridge, MA: MIT Press.

White, R.T., & Gunstone, R.F. (1989). Metalearning and conceptual change. *International Journal of Science Education, 11*, special issue, 577-586.

White, R.T. (1994). Conceptual and conceptional change. *Learning and Instruction, 4*, 117-121.

5. Progress in the Diagnostics of Mental Models

Sabine Al-Diban

Technical University Dresden, Germany

Abstract: Mental models are designed in individually meaningful problem situations, and are specific phenomena that enable individuals to develop new knowledge. With regard to external phenomena, mental models show internally a subjective plausibility and externally an explanation value for the reality. Knowledge diagnostics of mental models, that is standardized and simultaneously, allow content related statements, only recently beginning to be done. The newly developed and standardized Test for Causal Models (TCM) as a combination of structure formation techniques and causal diagrams appears to be a workable structure discovering method in the context of an extensive empirical test of validity. Generally with regard to content the complexity measure had a higher validity than the formal one. Nearly diminution-free qualitative and quantitative data analyses and evaluations become feasible in connection with content structuring and lattice analysis. So the problem of the structure isomorphism arising in structure formation techniques is overcome, and comparisons as well as data based typifying of the studied dependent changes, become possible for intra- and inter-individual comparisons. The subjective plausibility content is examined in detail. The meaning of measuring – validity – dilemma with regard to an individually centred data elevation and with low information losses connected with data inquiry is discussed. A perspective for further research is given.

Keywords: Mental models; validity; test of causal models; structure isomorphism.

"The descriptive phase of the investigation on naive conceptions has now been concluded, while the micro-genetic analysis of processes in response to instructional interventions is currently booming and highly promising" Caravita (2001, p. 421).

Introduction

Learning processes are connected with individual changes of the structure and content of knowledge representations. One kind of complex knowledge representations is mental models. "They were used to explain or simulate specific phenomenon of objects and events. [...] There are different concepts of mental models, but their same starting point is, that they will be constructed on the basic of recallable knowledge." (Seel, 2000, p. 265) Mental models represent and organize the subject's knowledge in such a way that even very complex phenomena of the reality become plausible.

From an instructional point of view is it very important to know what characterizes the preconceptions and what the resulting post conceptions of mental models are after a learning intervention. Mental models are a central theoretical construct of the situated cognition (Reynolds, Sinatra & Jetton, 1996) and a moderate constructivist approach. In the last one, the role of a teacher is to be a "coach" and the learning environment should be designed as an opportunity to initiate active and self regulated learning processes (Collins, Brown & Newman, 1989). An advised method of the latter approach for an effective change of mental models is to initiate 'conceptual conflicts' (Nussbaum & Novick, 1982). Comprehension and reasoning in learning and real-life situations necessarily involve the use of mental models (Greeno, 1989). That's one main reason why is it so important to diagnose the mental models of the students. More theoretically sounded, an important precondition to know more about the human knowledge acquisition process is a valid diagnosis of the mental model representations.

The study which is reported in this paper was designed to investigate processes of change of mental model representations in such a moderate constructivist context. The learning environment was strictly designed after the principles of the cognitive apprenticeship approach (Collins, Brown & Newman, 1989). The instructional design is reported in detail in Seel, Al-Diban, and Blumschein (2000).

The study was engaged with problem solving in the high school subject civics as one representative example of use. When we know more about the processes of change of mental model representations during learning processes, it becomes possible to increase the efficiency of instructional designs in general and their individual specificity. For domain specific and complex learning tasks it is necessary to know not only quantitative criteria but also qualitative criteria about the contents and the quality of the mental model representations too. This would allow for the differentiating between low and high change resistant attributes of preconceptions. A better diagnosis of mental models is highly relevant to facilitate complex learning processes. The content related diagnosis of mental model representations should be developed based on the scientific quality criteria objectivity, reliability and validity. This is the first step to the long term aim of applying such diagnosis instruments in various knowledge domains and fields of practice.

Theoretical Foundations

Never can the diagnostic of mental models be better than its theoretical foundation and the operationalization of the construct, which is to be measured.

While the current research is focused mainly on micro-genetic analyses of learning dependant processes of mental models (Caravita, 2001), most of the research lacks a consequent theoretical foundation. As a consequence my empirical work was concentrated on one theoretical concept (Seel, 1991; 2004). This made it possible to arrange a systematic empirical testing strategy concerning the validity of the measuring instruments.

The following theoretical foundations of mental models (Seel, 1991) were explicitly included and realized in the empirical testing strategy (see Figure 1).

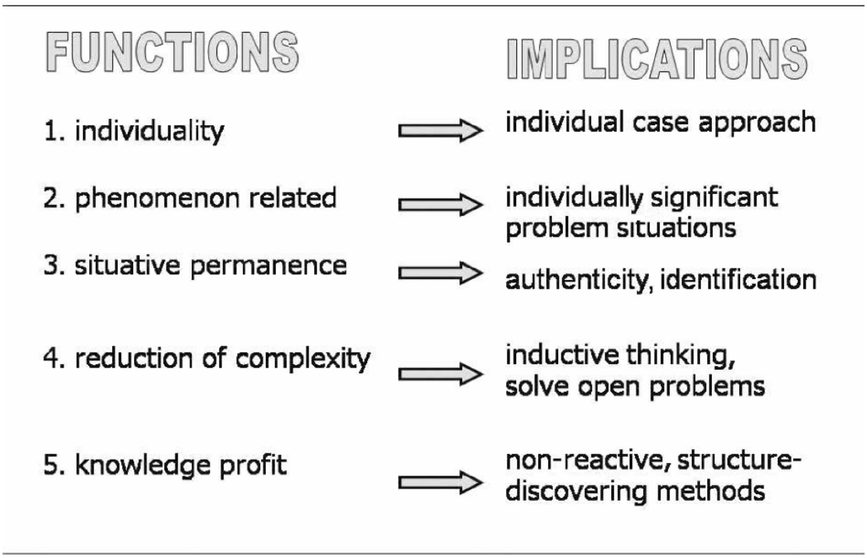

Fig. 1. Implications for the diagnostic of mental models.

1. Complex knowledge representations like mental models can be characterized as highly individual, because they are dependent on the prior knowledge, idiosyncratic experiences, interactions and the every day language of a human being. The most important implication for the diagnosis is to use an individual case approach and qualitative investigation methods to actually represent the wide range of possible knowledge contents.
2. Mental models are area specific and phenomenon related. That's why it is necessary to arrange an individual meaningful problem situation. This situation should include the requirements to explain and/or predict a phenomenon or problem. If there is no area specific or phenomenon related task used, then no mental model representations, merely descriptions or factual knowledge, were examined.
3. Mental models do not exist permanently; they are merely considered to be situational permanent cognitive and functional constructs. This implies that mental models have to be observed in individually significant and meaningful problem

situations, with reference to the participants of a study. The researcher should design tasks and problems with intensive authenticity or possibilities of identification. The reported study realized a high curricular validity through participating high school students and a curriculum tasks. In addition, there was an up-to-date and authentic topic included – the implementation of the new currency EURO in Europe.

4. People construct mental models for such requirements generally only when they help to reduce complexity. That is why it is important to check the complexity as objective as possible with task analysis (Jonassen, Hannum & Tessmer, 1989). It is advisable to take care of different important and controllable influences. For instance, task analysis revisers have to be true experts in this subject; their horizon of understanding and actual ability to solve the referring problem should be observed. Revealing and recommendable is a view of further dimensions in the task analysis like 'dynamic' and 'required knowledge' (Hacker, Sachse & Schroda, 1998). Another hint for the processing of a diagnosis is the avoidance of problem specifications and direct supports. Tasks should be created which require inductive thinking and accommodations to solve open problems.
5. A further main function of mental models is that they serve the knowledge profit. For this reason the diagnosis methods should enable the participants to acquire new knowledge, make new combinations of available knowledge, think inductively, realize accommodations and find the best solutions for open and/or new problems (Seel, 2004). It's possible no reactive and structure-discovering methods are necessary to afford the conditions for an adequate representation of these functions.

When recapitulating all these functions it can be said, that mental models are a cognitive construct to 'run in the mind's eye'. This enables the potential to find new solutions, analogies, and generalisations with the aim to explain or prognosticate the reality or predict future events.

This work was focused on the assessment of the contents of learning dependent changes of mental model representations (Seel, 1991, 2001, 2004). Subjective causal models can be understood as a subset of mental models concerning causal explanations of causes and consequences of the phenomenon of inflation on the macro-economic system. It will be assumed, that temporal sequences of causes- and consequences (or if- then- relations) represent the subjective causal thinking of the students. This is called the 'dynamic hypothesis' (Seel, 1994, p. 2). A large psychological comprehension of causality also includes action orientated (Gasking, 1981) and temporally sequenced relations (van der Meer & Schmid, 1992).

In addition, the process of situational model construction takes place in an interaction with conceptual models from experts and their established scientific theories (see Figure 2).

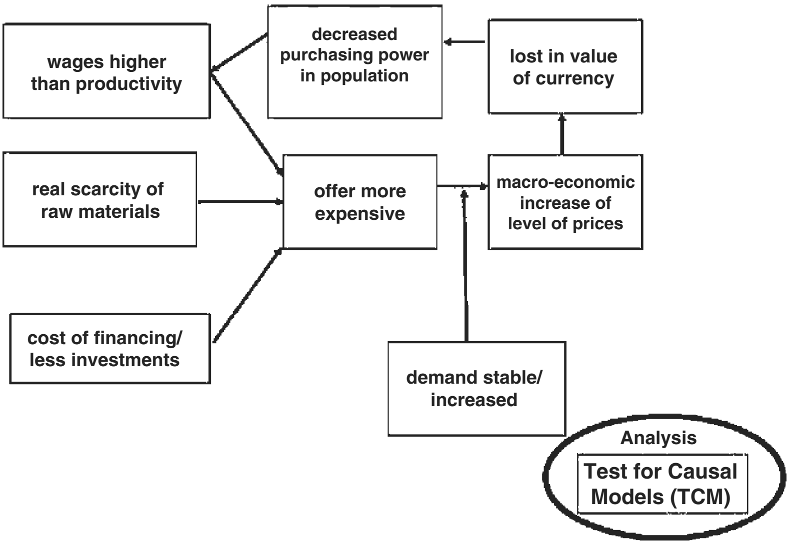

Fig. 2. Part of a conceptual expert model: offer caused genesis of inflation.

On one hand, mental models possess plausibility which is reality related. This is called their 'explanatory value'. This study indicates the degree of accordance with the 'dynamic hypothesis' (if- then- relations) of a conceptual expert model. In Figure 2 it is shown that one part of the conceptual expert model, concerning the type of offer- caused- genesis of inflation with eight if- then- relations (Al-Diban, 2002). The entire expert model of all types of inflation possessed 47 if- then- relations. The expert model of inflation was developed in collaboration with a doctorate economist, who has substantial experience in applied research.

On the other hand, mental models contain subjective plausibility measured with indicators like consistency, creativity, impression and coherence in relation to every individual case. According to Seel (1991) the subjective plausibility is seen as another central feature of mental models. The plausibility is understood as accurateness and coherence in regard to the entirety of the domain specific 'world knowledge' of one person.

Summarized, mental models simultaneously possess internally subjective plausibility and externally a more or less high explanation value for the reality.

The studies of (Thagard, 1992) also differentiate between a subjective and a so called 'explanatory coherence'. The explanatory coherence of mental models from experts like Lavoisier, Wegener, Keller or Newton, in comparison with their contemporary contra ends, brought the following results:

> "Propositional systems are primarily structured via relations of explanatory coherence. New theoretical hypotheses generally arise by abduction. The transition to new conceptual

and propositional systems occurs because of the greater explanatory coherence of the new propositions that use the new concepts" (Thagard, 1992, p. 9).

It is necessary to emphasize that the domain of validity of Thagard's studies exclusively were conceptual models from scientists and experts. For educationalists and instructors the much more interesting questions should be:

- How do the contents of mental models by students change, especially if they dispose of low or inadequate preconceptions only?
- What scientific quality criteria like objectivity, reliability, and validity can be measured and checked by intra- and inter-individual comparisons?

The exact object of investigation theoretically can be described as the decoded representation fragments of subjective causal models of conscious and recallable dynamic hypotheses to explain and/or prognosticate the phenomenon of inflation. The dependant variables for the content of these subjective causal model representations were evaluated as the summation of the self-formulated terms and subjective causal relations for each individual mental model representation in comparison with conceptual expert models.

In the next chapter I want to introduce the standardized and content related diagnostic strategy.

Diagnostic Strategy

The knowledge reconstruction was arranged by an analysis with the newly developed Test for Causal Models (Al-Diban, 2002). This test is inspired by the idea that structure discovering thinking is unlike structure testing thinking. In other words, there are no direct or indirect target settings neither concerning the relevant problem variables nor concerning the connection between these variables. This standardized test consists of a combination of Structure Lay Down Techniques (Scheele & Groeben, 1992) and Causal Diagrams (Funke, 1990). The newly developed test is based on the theoretical functions, which are relevant to the mental model construction process (Figure 2). The test is adapted to the concrete learning intervention.

This newly developed Test of Causal Models (TCM) uses an open structure lay down method and allowed for a correspondence of findings in comparison with an interview. This interview, at the beginning, was used as a reactivity poor knowledge acquisition method and aimed to reconstruct the subjective problem space and the general prior knowledge of the participants. The interview results served as a comparison between the subjective problem space and the subjective causal preconceptions. Further more the interview group was used for the checking of the differential validity to a group without interview.

With the Test of Causal Models inter- and intra-individual comparable indicators of the formal and content based complexity can be analysed. The following

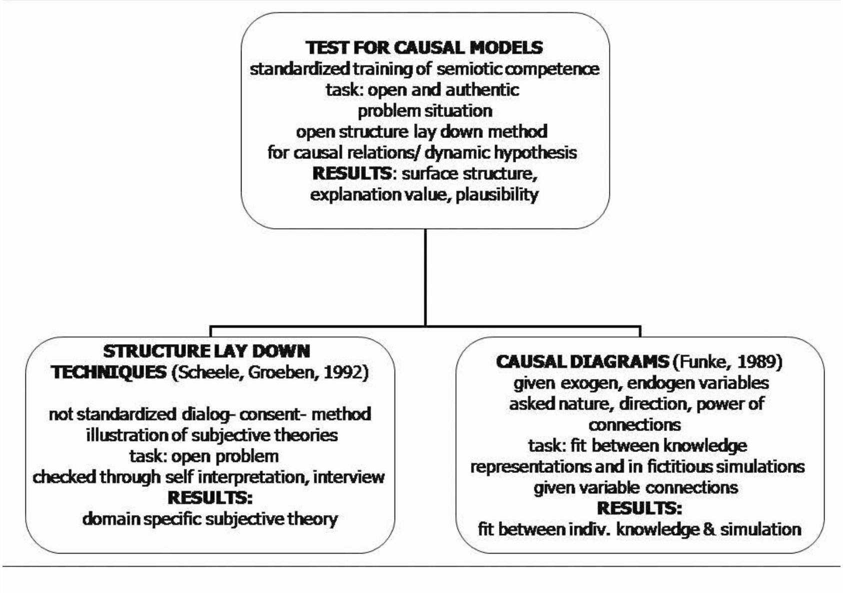

Fig. 3. Elements of the new developed Test for Causal Models.

picture (see Figure 3) shows the elements of the Test of Causal Models. There are to find more details in Al-Diban (2002).

Open knowledge acquisition methods already existing, like Structure Lay Down Techniques (Scheele & Groeben, 1992), have no theoretical test criteria. Participants with low intelligence, education, or linguistic abilities can not be tested because they would need good language skills. Facing the interviews, the content analysis is very labour-intensive. Likewise standardized questionnaires are not qualified for a diagnosis of mental models. Most of them use closed questions only. Such items can not represent the entirety of inductive thinking and accommodations to solve phenomenon related, open problems. They are not exhaustive. At the best case questionnaires represent a central section of the entire mental model. With questionnaires the individuality of mental models can't be reproduced.

The main disadvantage of structure testing methods like Causal Diagrams, (Funke, 1990) is the fact that those methods only establish a correspondence between a reconstructed and a given answer. In Causal Diagrams there is proof of the fit between knowledge representations and given variables, and the connections among them. In fact the subjects actually have different subjective problem spaces and preconceptions. So they use inductive thinking to solve the problem with their available knowledge. This is not asked at all in this method. As a logical consequence, there is no empirical relative for inductive – only adaptive thinking to a given target structure – in this method. An example is, when all variables, which are not included in the target structure, won't be asked or measured.

This study was an integrated part of a bigger research project[7] and is based on a small, not randomised sample of 26 German 11th class grammar school students. A complete identical multimedia teaching program about cycle systems in a macro-economy was developed and used as a standardized intervention. The instructional design used was the principles of the Cognitive Apprenticeship (Collins et al., 1989). There was a considerably high internal and curricular validity through curricular relevant topic in civics – implementation of the new currency EURO in Europe. In this research design the results can only be transferred very limitedly. They exclusively deal with internal derivation based, subjective causal mental model representations of inflations.

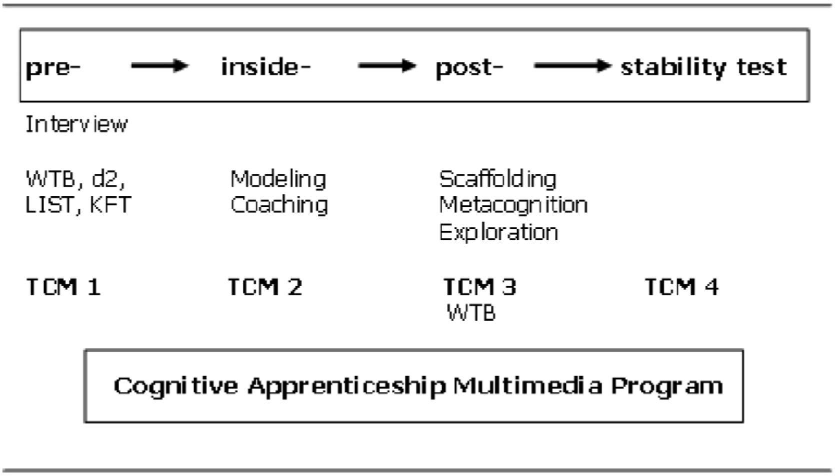

Fig. 4. Learning process accompanying study, N = 26 students, 11th class.

The empirical testing strategy of validity (see Figure 4) included the following points: retest reliability between dependent and independent measures of TCM 2 and TCM 3 under the assumption of outwearing mental models during the learning intervention, differential validity between participants with and without 'cognitive conflicts' activated by an interview and conformity validity between all instruments, which collect data of the causal complexity, especially the explanatory value – TCM 1, 2, 3, 4, the scaffolding problem answers, and a dimension of verbal analogy thinking in the Test of Cognitive Abilities (KFT). Last but not least, the prognostic validity was measured as transfer between causal mental models in the pre-test (TCM 1) and the solutions of very complex problems in exploration,

[7] This project was very successfully guided by N.M. Seel and sponsored by the German Research Society from 1997-2001 at Technische Universität in Dresden and Albert-Ludwigs-Universität in Freiburg, Germany.

the last part of the Cognitive Apprenticeship and the causal mental models in stability test (TCM 4). For this purpose cross-lagged-correlations (Lazarsfeld & Barton, 1955) and regression analyses were applied.

Findings – Single Approach

One remarkable progress in the diagnostic of mental models is that nearly diminution-free qualitative and quantitative data analysis and evaluations became feasible. Therefore two data analyses steps were realized. The first step concerned a qualitative content structuring analysis (Mayring, 1997) of all terms after the semantics. Here an overview is given on the results from more than one participant, about commonly used terms to construct mental models of inflations.

Table 1. Overview of in the Test of Causal Models commonly used terms

pretest	inside test	post test	stability test
6 wages	12 German Central Bank	12 German Central Bank	3 demand
4 prices	11 inflation	13 inflation	3 production
3 purchasing power	10 firms	7 state	
3 customer	6 housholds	6 households	
3 scarcity of raw materials	6 state	6 firms	
3 demand	6 products	6 trade: import	
	5 money supply	6 trade: export	
	4 wages	6 wages	
	3 prices	4 production	
	3 offer	4 purchasing power	
	3 market	4 commercial bank	
	3 production	3 prices	
	3 purchasing power	3 demand	
		3 market	
		3 abroad	
		3 environment protection	

new term contents in inside and posttest only

Obviously you can only find shared, new term contents in inside and postest (see Table 1). The term "demand" seems to be very stable, but terms like "wages, prices, purchasing power, customer, scarcity of row materials" are not remembered in the Test of Causal Models four months later in the stability test. Clearly you find less shared terms in stability test than in the pre-test.

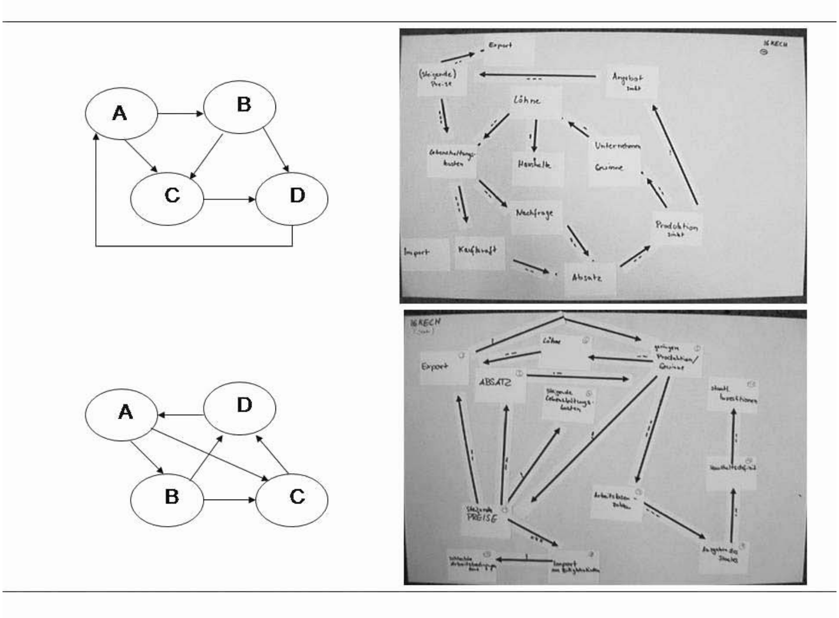

Fig. 5. Structure isomorphism – same or different structures? (The example on the right side is the same participant in pre- and stability test, row data).

The second step concerned the data and expert model based application of formal concept analysis (Ganter & Wille, 1996). The main problem of standardizing the knowledge data collection methods, which use structure formation techniques, is structure isomorphism (Nägler & Stopp, 1991). For example the same four elements can be connected in 24 arrays (see Figure 5). A true content based comparison is nearly impossible. The already reported two analysis steps – qualitative content analysis and formal concept analysis – overcame this problem. The main progress consists of the fact that content based comparisons, as well as data based typifying of the student dependent changes, become possible.

In which way does the formal concept analysis help? Mental models can systematically structured after objects and the entirety of all true attributes based on the mathematical lattice theory (Birkhoff, 1973). The central assumption of formal concept analysis is that a systematic structuring after all attributes helps to survey complex qualitative data. The analysis principle is data evolvement in contrast to data aggregation. How is a concept lattice graphic drawn? A formal context (G, M, I) consists of two sets G and M and of a binary relation $I \subseteq GxM$. The elements of G are called the objects, those of M are the attributes of (*G, M, I*). If $g \in G$ and $m \in M$ are in relation *I*, we write $(g,m) \in I$ or *g I M* and read this as *"the object g has the attribute M"*. For any formal context (*G, M, I*) the set of extents is a closure system on *G* and the set of intents is a closure system on *M* (Ganter, Wille, 1996). The Formal Concept Analysis was used here for a systematic data evolvement of the contents in the Test of Causal Diagrams. The result is a systematic structuring of objects (in this example of use the participants) after all true attributes (here: entirety of contents firstly on the level of terms, secondly on the level of causal relations in the test causal models). All resulting graphics are defined as "term content" – you can reach all attributes through upward lines and as "term volume or amount" – you can reach all objects through downward lines related to the point where the single object is located in the graphic. In other words – an attribute on the top is applicable to all objects; an attribute on the bottom is respectively applicable to no object participant.

Figure 6 shows a formal concept analysis (Ganter & Wille, 1996) with in preand stability test called terms from the same participant like in Figure 5. The attributes were mental models representations reconstructed in the TCM, here analyzed on the level of terms. This analysis is based on a cross-classified data table for seven participants of one subgroup as objects and all their true attributes in the background. In the foreground one participant is pointed out.

The intra-individual comparison shows a high stability as well in the number of used terms (12) as in a constant core of 6 terms (gain, production, earnings, prices,

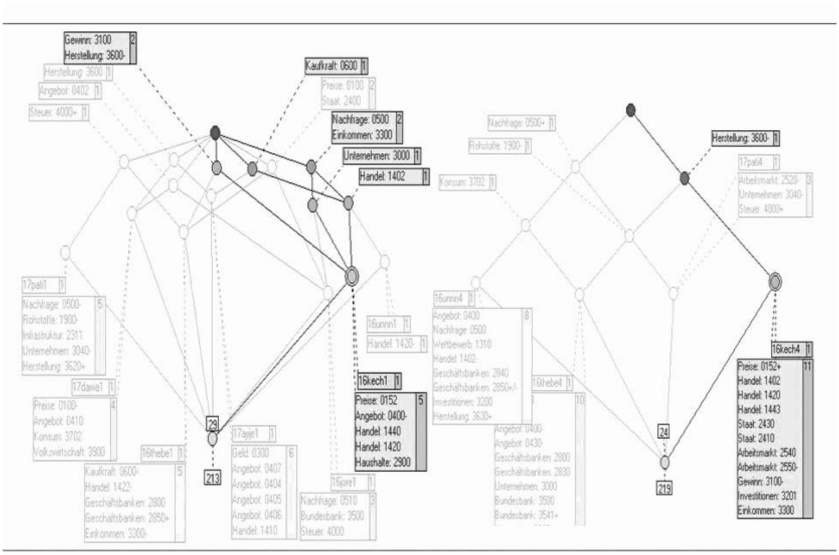

Fig. 6. Formal concept analysis – individual case approach on the level of terms by the same participant as Figure 5: 16kech (pre- and stability test).

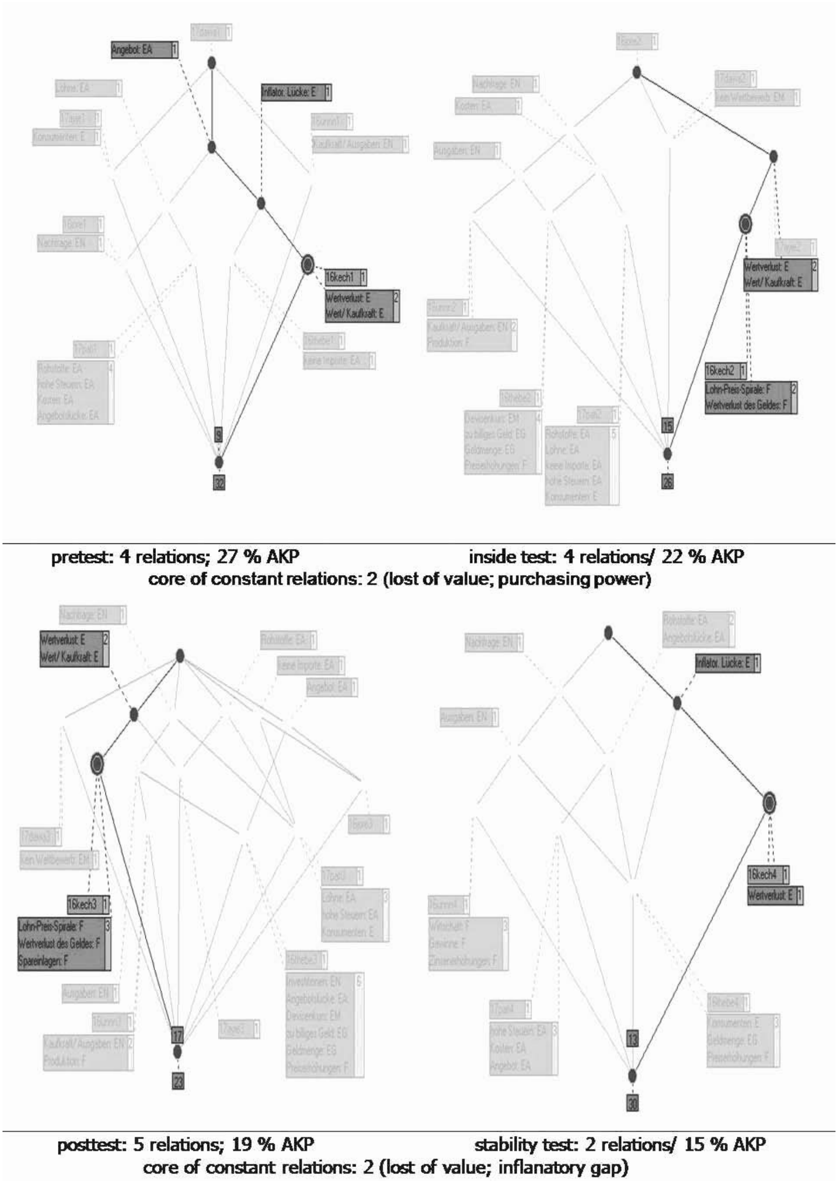

Fig. 7. Individual case approach – Level of causal relations by 16kech, AKP: proportion of conceptual expert model based explanatory value.

export trade, import trade) by this student. The preconception refers strongly to firms and is classified as a genesis- no consequence-model of inflation, which is shown in Figure 6. The explanation value (AKP) has a low level, a constant core of 2 relations and decreases continuously (Figure 6). Inter-individual perspective: The originality of used terms in comparison with a group of 7 students is low, especially in the preconception (Figures 2–5), but the originality of causal relations (Figure 7) is higher. Altogether the results of participant 16kech were classified as an external add-on changing type of the mental model for the phenomenon inflation. This student constructed a pure genesis model of inflation with 75% originality of terms and 86% proposition of explanatory value (AKP) in the pre-test. The array of raw data is equivalent to a star. The described formal criteria do not change over the 4 times, but the content related criteria originality and explanatory value increased continuously (Figure 8). There is a correct rethinking to find in inside test from a "lacking demand" to a "strong demand" as a reason for the genesis of inflations. Including the starting conditions of the learning process the participant 17pati (see Figure 9) was classified as a data based integrative reconstruction type in contrast to the external add-on type in the example 16kech. When recapitulating all single approaches the following permanent striking observations were obvious. The formal and the content related criteria develop quite independent from each other. In most cases an increase of formal (account of terms, of relations, depth of connectivity) and a decrease of content related criteria (explanatory value) and vice versa could be observed simultaneously. The developing types of the observed mental models are widely independent from the cognitive starting points (area specific previous knowledge, concentration). There are some hints that the teach back interview in the beginning benefits a favourable change of mental models. In summary, six external add-ons, six integrative reconstructions, two absolute constants, some unclassified change types were found.

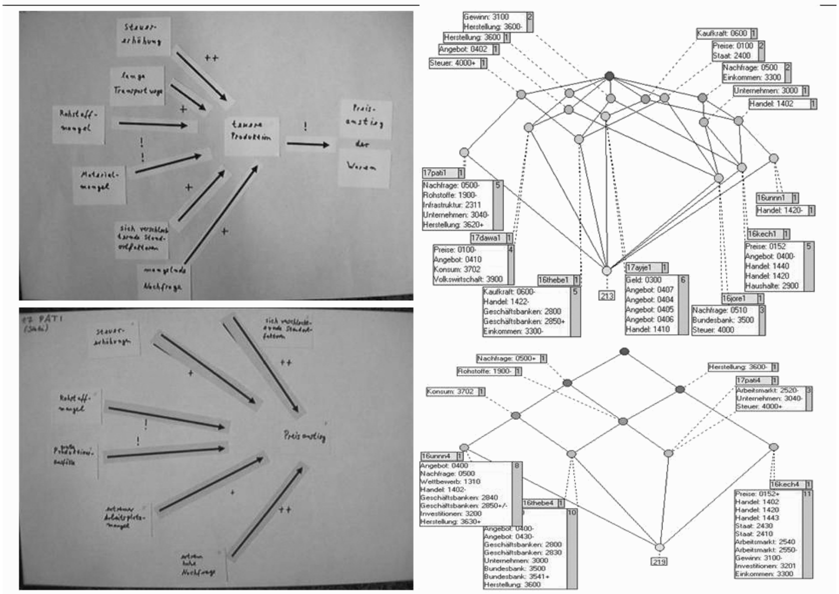

Fig. 8. Individual case approach 17pati – Level of row data and of terms (pre-, stability test).

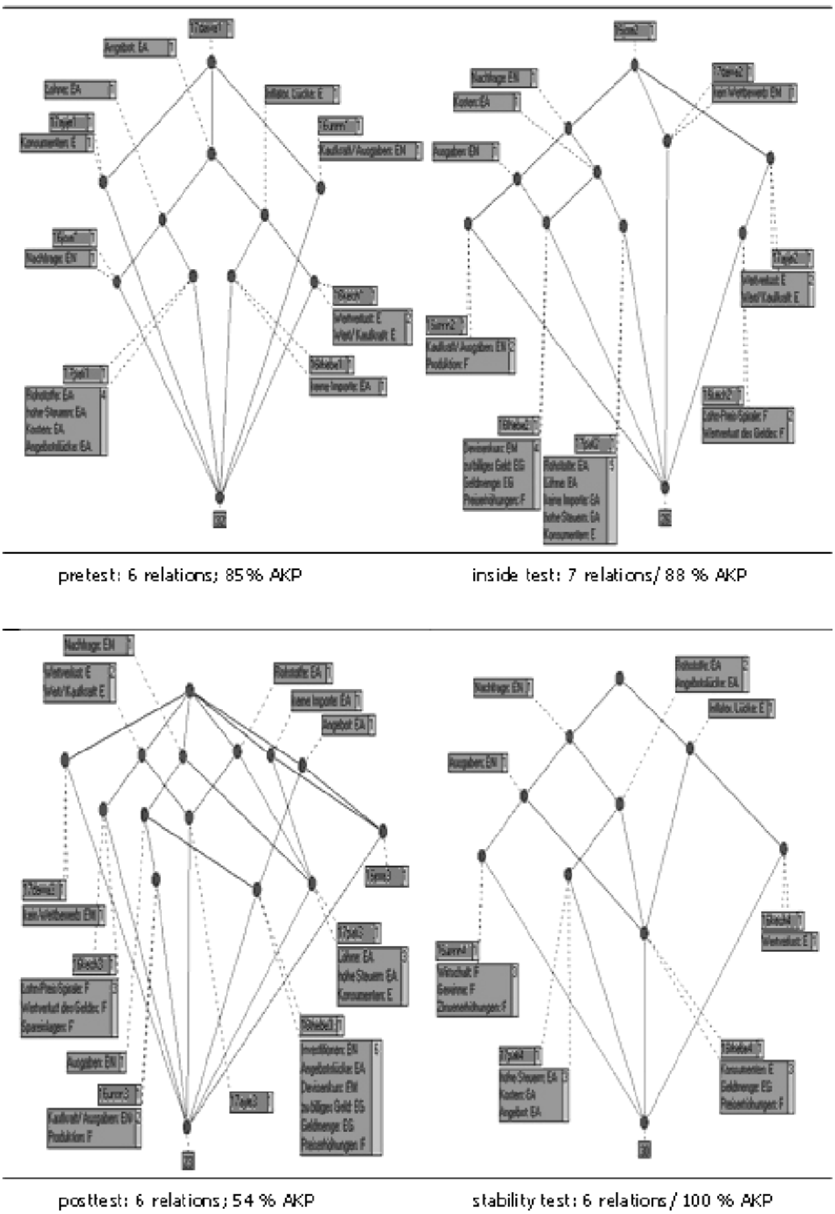

Fig. 9. Individual case approach 17pati – Level of row data and of terms (pre-, stability test).

Findings – Plausibility

The subjective plausibility content was examined in a sense of correspondence with the conceptual expert model (Figures 2–5) as proportion of conceptual expert model based explanatory value (AKP) or absolute explanatory value on the one hand and in a sense of subjective plausibility, consistency, coherence (Thagard, 1989) on the other hand.

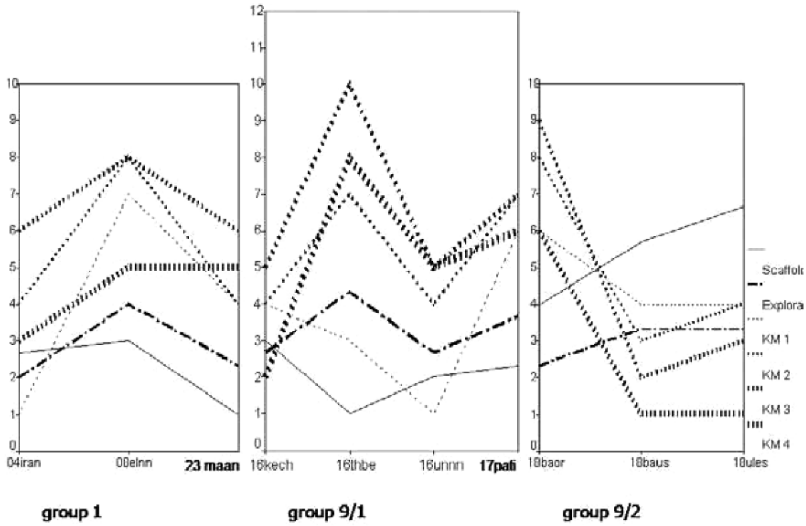

Fig. 10. Change of plausibility: Explanatory value absolute: Scaffolding, Exploration, Causal Models 4 times (KM), 10 single approaches.

Almost all single approaches show correspondent changing tendencies in the absolutely explanatory value of the test of causal models, the scaffolding and exploration problems. That means if an increasing explanatory value is found from scaffolding to exploration increasing explanatory values in the test of causal models will be recorded from pre- to post-test too. Also corresponding decreasing tendencies are found. Despite of the identical learning intervention there are 4 participants with a higher, 2 a constant and 4 with lower explanatory value in the stability then in pre-test (see Figure 10).

For the further research it is necessary to develop more exact and selective operationalizations of the different aspects of plausibility. This is shown in Tables 2 and 3, which illustrate these aspects on the basis of the solutions in Scaffolding and Exploration.

Table 2. Correlation Analysis inside Scaffolding

	Expla-natory value	subjective content plausibility		subjective formal plausibility		decl. know-ledge
		consistency	creativity	impression	coherence	
Explanatory value	1	0.47*	0.50*	0.64*	0.53*	0.14
consistency		1	0.37	0.80*	0.54*	0.58*
creativity			1	0.42	0.44*	0.05
impression				1	0.76*	0.72*
coherence					1	0.59*
decl. knowledge						1

* $p < 0.01$

Table 3. Correlation Analysis inside Exploration

	Expla-natory value	subjective content plausibility		subjective formal plausibility		decl. know-ledge
		consistency	creativity	impression	coherence	
Explanatory value	1	0.31	0.33	0.32	0.45	0.21
consistency		1	1.29	0.38	0.39	0.26
creativity			1	0.19	0.09	0.27
impression				1	0.75*	0.66*
coherence					1	0.52*
decl. knowledge						1

* $p < 0.01$

In Scaffolding the highest correlations between consistency as well as coherence and impression of the answers were found. Continuing high and significant correlations exist between creativity, impression, coherence, and consistency and the explanatory value as well as impression, coherence, consistency, and the declarative knowledge. Noticeable are generally lower correlations inside Exploration. The highest one is between coherence and impression. Alone between impression and declarative knowledge exists further significant correlation. That means the means the subjective aspects of plausibility are connected with the explanatory value in scaffolding but not in exploration. Impression and coherence are connected with declarative knowledge in both problem solutions, here could exist an operationalization problem.

Findings – Validation Strategy

At first the retest reliability between dependent and independent measures of TCM 2 and TCM 3 under the assumption of intervention outwearing mental models during learning was tested. All content based criteria – absolute and proposition of explanatory value – showed comparable high and significant retest correlations (r_{max} =.925; r_{min} =.789) but all formal criteria – number of terms, arrows, depth of connectivity – were on a low insigni-ficant level (r_{max} =.691; r_{min} =.391). This is a hint that an accurate measuring of mental models is possible by content based criteria like explanatory value.

Second it was checked the differential validity between participants with and without 'cognitive conflicts' activated by an interview. Only some students took part there. So here are reported only hints that the explanatory value developed in comparison of pre-/post- and pre-/stability test in correspondence with the assumptions for a more intensive and long term learning in the group with the teach back interview at the beginning.

Third the conformity validity was tested between all instruments, which collect data of the causal complexity, especially the explanatory value – TCM 1, 2, 3, 4, the scaffolding and exploration problem answers and a dimension of verbal analogy thinking in the Test of Cognitive Abilities (KFT). It was already reported (Tables 2 and 3), that inside Scaffolding and Exploration were found some problems with operationalizations of the different aspects of subjective plausibility. Furthermore in correspondence with the hypothesis closed correlations between verbal analogy thinking in the Test of Cognitive Abilities (KFT) and the coherence (r =.555**), the creativity (r =.397*) and the explanatory value (r=.330) of the scaffolding answers.

By the exploration answers significant correlations (r =.498**, r =.496**) are found with the pro-portion of explanatory value in the Test of Causal Models in

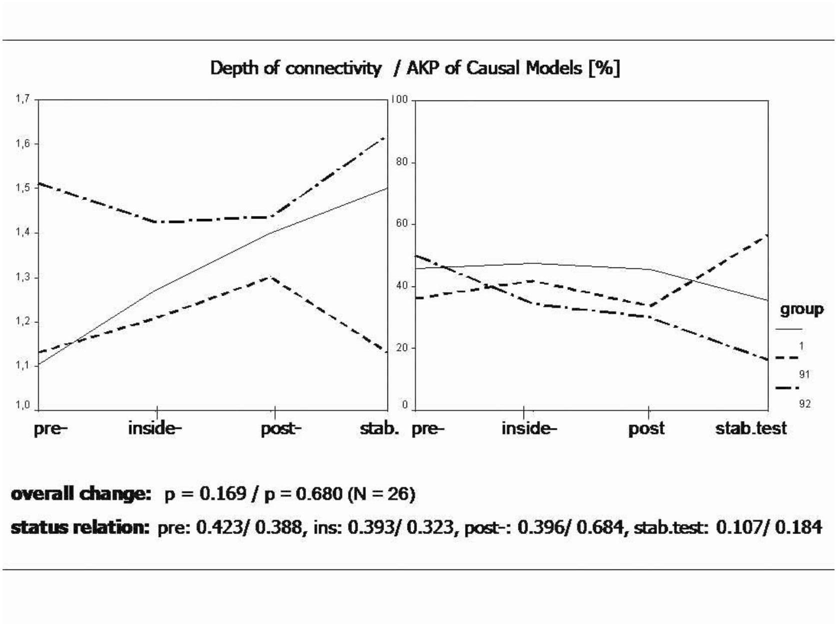

Fig. 11. Depth of Connectivity/AKP of Causal Models per group in 4 times.

inside- and post test, (stability test: r =.485) and the KFT verbal analogy thinking (r =.426*).

Fourth, the prognostic validity was measured as transfer between causal mental models in the pre-test (TCM 1) and the solutions of very complex problems in exploration, the last part of the Cognitive Apprenticeship as well as the causal mental models in stability test (TCM 4). The cross-lagged-correlations showed the following results. Between AKP of the mental model representations in the pre-test and the depth of connectivity in the post test were found negative correlations r = -.443* (pre-/stability test: r = -.281). Otherwise strong positive correlations exist between the AKP of the mental model representations in the pre-test and the post test (r =.606**) and the stability test (r = .477). The correlation values are stronger than between the absolute explanatory value to the four times. That means over the accompanying learning process there are time delayed proportional connections between the explanatory value and reciprocal connections between explanatory value and depth of connectivity of mental model representations (see Figure 11).

To prognosticate the transfer to explanatory value of the exploration problem the strongest correlations were found for the AKP of the mental models in inside test (r =.498**), then AKP in post test (r =.489*) and AKP in stability test (r = .485). So this vertical transfer problem can called a prognostic outside criterion.

The numerous stepwise multiple regression analyses which were calculated under the assumption of internal construct validation: explanatory value of causal models in pre-, inside-, post- and stability test, KFT, depth of connectivity, and the different aspects subjective plausibility. Under the assumption of elimination, domain specific knowledge (WBT), a questionnaire of learning strategies (LIST), mistakes in coaching and declarative knowledge in scaffolding, were included. The dependent variable was the explanatory value in scaffolding and in exploration.

While the internal construct validation for scaffolding failed, the explanatory value of exploration reached good results. The explanatory value of causal models in inside and post-test conformed to the hypotheses which were included in the model. Together with the coherence 24% explained variance of the exploration answers were found. A step forward into a content based empirical validation of mental models was done.

The conclusions are: mental models show internally a subjective plausibility and externally a more or less high explanation value for the reality. In some cases they are related with each other. Generally complexity measures concerning the contents – like explanatory value or AKP – have an acceptable validity; formal ones have a low or no validity. The empirical construct validation showed contended results as well for retest reliability as for differential, conformity and prognostic validity. The results for prognostic validity with regression analyses showed 24% explained variance for vertical transfer problem solutions in exploration. The explaining variables were causal model representations in inside and post test, combined with the coherence of exploration answers. Declarative knowledge was excluded. Change resistance is higher on the level of explanations than on the level of terms. Typical change types over the 4 times were six "integrative reconstructions", six "external add on", two "constant mental models", the others could not be classified. These change types were not linked with the cognitive starting conditions, but successful supported by a teach-back interview in the beginning. The combination of structure content analysis and formal concept analysis made possible content based inter- and intra- individual comparisons of mental model representations. The problem of structure isomorphism was overcome. In summary only this small study is able to give factual based orientations for further research questions.

Perspective

Replications of this study have already taken place in actual studies with bigger samples, model-based interventions and including causal model representations of different phenomena too (Seel, Ifenthaler & Pirnay-Dummer, i.p.) and are also reported in this book.

Especially for a data and content based diagnosis of mental models the measuring- validity- dilemma is highly relevant. In consideration of an acceptable validity exclusively of content based criteria of mental models it should be asked after a compromise for the empirical research. The compromise should respect that data aggregation reduces the amount of data but includes a loss of information (Figure 12). On the other hand should be respected that the individual approach (Figure 1) has priority. Data on the level of groups are not adequate to assess mental models. In this contribution it is suggested to make a compromise while categorical data are used. Categorical data allow approximate content based information and a lot of statistical analyses[8] as well as the application of test theoretical models. The criticism that most of empirical studies concern one sided organized only "...initial learning in well structured domains without fathoming […] complex changes of knowledge structures in later development phases (Stark, 2003, p. 138) is emphasized. In the majority of studies there is an absence of a standardized change measurement of mental model representations. Furthermore, the here reported study shows the high impact of content based criteria for a valid measuring of mental model representations. This basic prerequisite is essential for an application-oriented conceptual change research with the highly significant aim to support a goal-oriented teaching.

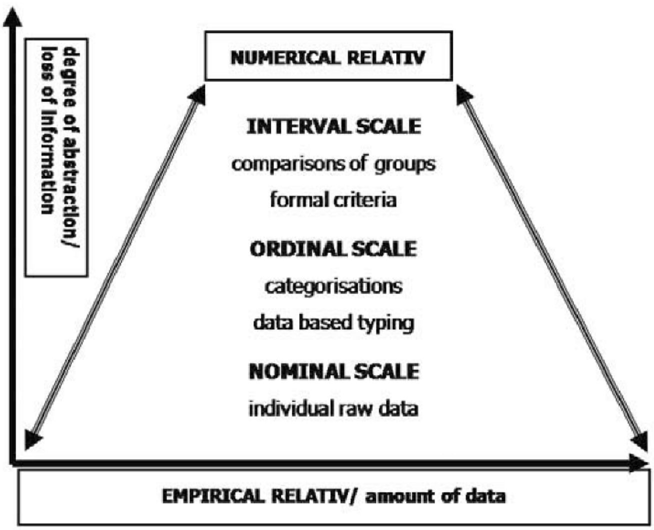

Fig. 12. Measuring- Validity- Dilemma by diagnosis of mental models.

[8] f.i. Rasch- Models of persons, times and situations (Rost, 1996)

References

Al-Diban, S. (2002). *Diagnose mentaler Modelle*. Hamburg: Kovac.

Birkhoff, G. (1973). *Lattice theory*. 3th ed. American Mathematic Society Coll. Publ. 25: Providence, R.I.

Caravita, S. (2001) A re-framed conceptual change theory? *Learning and Instruction 11*, 421-429.

Collins, A., Brown, J.S., & Newman, S.E. (1989). Cognitive Apprenticeship. Teaching the crafts of reading, writing and mathematics. In L.B. Resnick (Ed.), *Knowing, learning and instruction* (pp. 453-494). Hillsdale, NJ: Erlbaum.

Funke, J. (1990). Systemmerkmale als Determinanten des Umgangs mit dynamischen Systemen. *Sprache & Kognition, 9*(3), 143-153.

Funke, J. (1992). *Wissen über dynamische Systeme: Erwerb, Repräsentation und Anwendung.* Heidelberg: Springer.

Ganter, B., & Wille, R. (1996). *Formale Begriffsanalyse. Mathematische Grundlagen*. Berlin: Springer.

Greeno, J.G. (1989). Situations, mental models and generative knowledge. In D. Klahr, & K. Kotovsky (Eds.) *Complex information processing* (pp. 285-318). Hillsdale, NJ.: Lawrence Erlbaum.

Hacker, W., Sachse, P., & Schroda, F. (1998). Design thinking-possible ways to successful solutions in product development. In H. Birkhofer, P. Badke-Schaub & E. Frankenberger (Eds.) *Designer – the key to successful Product Development*. London: Springer.

Lazarsfeld, P.F., & Barton, A.H. (1955). Das Verhältnis von theoretischer und empirischer Analyse im Rahmen qualitativer Sozialforschung. In Ch. Hopf, & E. Weingarten (Eds.), *Qualitative Sozialforschung*. Stuttgart: Klett.

Mayring, P. (1997). *Qualitative Inhaltsanalyse-Grundlagen und Techniken* (Vol. 6). Weinheim. Beltz: Deutscher Studien Verlag.

Nägler, G., & Stopp, F. (1996). *Graphen und ihre Anwendungen. Mathematik für Ingenieure und Naturwissenschaftler*. Stuttgart: B.G. Teubner.

Nussbaum, J., & Novick, S. (1982). Alternative frameworks, conceptual conflict and accommodation: Toward a principles teaching strategy. *Instructional Science, 11*, 183-200.

Reynolds, R.E., Sinatra, G.M., & Jetton, T.L. (1996). Views of knowledge acquisition and representation: A continuum from experience centered to mind centered. *Educational Psychologist, 31*(2), 93-104.

Rost, J. (1996). *Lehrbuch Testtheorie Testkonstruktion*. Bern: Huber.

Scheele, B., & Groeben, N. (1988). *Dialog-Konsens-Methoden zur Rekonstruktion subjektiver Theorien. Die Heidelberger Struktur-Lege-Technik (SLT), konsensuale Ziel-Mittel-Argumentation und kommunikative Flussdiagramm-Beschreibung von Handlungen*. Tübingen: Francke Verlag.

Seel, N. M. (1991). *Weltwissen und mentale Modelle*. Göttingen: Hogrefe Verlag.

Seel, N. M. (2003). Model-centered learning and instruction. *Technology, Instruction, Cognition and Learning, 1*(1), 59-85.

Seel, N.M., Al-Diban, S., & Blumschein, P. (2000). Mental models and instructional planning. In J.M. Spector, & T.M. Anderson (Eds.), *Integrated and holistic perspectives on learning, instruction and technology: Understanding complexity* (pp. 129-158). Dordrecht: Kluwer Academic Publishers.

Seel, N.M., Ifenthaler, D., & Pirnay-Dummer, P. (2007, i.p.). Mental models and problem solving: Technological solutions for measurement and assessment of the development of expertise. In J. Strobel, D. Jonassen, & W. Hung (Eds.) *Learning by modelling systems*. Mahwah, NJ: Lawrence Erlbaum.

Stark, R. (2003). Conceptual Change - kognitiv oder situiert? *Zeitschrift für Pädagogische Psychologie, 17*(2), 133 - 144.

Thagard, P. (1989). Explanatory coherence. *Behavioural and Brain Sciences, 12*, 435-502.

Model-Based Theories of Instruction

Dirk Ifenthaler and Pablo Pirnay-Dummer

University of Freiburg, Germany

Baking the Bread of Learning and Instruction

Even for the simplest tasks of instruction, interpretation and transfer from fundamental research to theories of instruction is not very often a straightforward procedure. Common misunderstandings are not only the result of differences in terminology, but also depend on how precisely findings are transferred to the field. In the preface we showed the transfers between the different fields. At this point, we are about to approach the first transition. The following part will cover the first transition and enter the field of model-based theories of instruction. It contains contributions from Pablo Pirnay-Dummer, David Jonassen & Young Hoan Cho, Andrew S. Gibbons, and Ulrike Hanke.

Pirnay-Dummer starts with a critical review of theory based experimental research and different alternatives for interpreting results with a special focus on common ways of interpreting effects, e.g. within studies. The idea of Design Experiments is then discussed on this basis and Extended Design Experiments are introduced. Methodological strengths and weaknesses are discussed against the background of important paradigms from philosophy of science, leading to constraints for research and to a new perspective on selected kinds of hypotheses. Having sorted out what can and what can not be concluded from specific kinds of results, the author creates a new scenario for learning and instruction which uses effect interpretation metaphors which are only slightly different from the (implicitly) common interpretation techniques. However, the differences of transfer for learning and instruction are not as small as researchers usually assume.

The chapter by Jonassen and Cho focuses on mental model externalization. After a review of assessment methodology, the authors provide an understanding of the complexity and multi-dimensionality within mental models which has to be taken into account in model assessment oriented toward the complex cognitive tasks of learning. They distinguish between structural, procedural, and episodic knowledge and show how the differences affect the assessment of each type of knowledge.

The impact of models about dynamic systems are the key focus of Gibbons' chapter. With an emphasis on simulation, the author introduces different layers of design which aim directly at model-centered design (content layer, strategy layer, control layer, message layer, representation layer, media-logic layer, data management layer). Hence, design constraints are derived from the theoretical foundation for all design layers separately, allowing designers to follow a matrix for every given set of constraints to make the right decisions for the specific design task.

Hanke introduces a theory-based design sequence designed to allow specific model building capabilities to be used for the acquisition of schemata and thus support stable learning of a given content or skill. A model of model-based instruction is introduced to support layers of subprocesses of learning, which leads to a sequence of instruction. The author then discusses the realization of this sequence in the practical field, which provides the bridge to the technological theories of instructional design.

6. Rendezvous with a Quantum of Learning

Effect Metaphors, Extended Design Experiments and Omnivariate Learning Instances

Pablo Pirnay-Dummer

University of Freiburg, Germany

Abstract: This chapter is about central ideas about how research on the bridge between cognitive psychology and theories of learning and instruction can be conducted in general and what methodological and logical traps may come with this special endeavor. These conclusions can be made on the basis of decades of consistent and complementary research on mental model theory and on model based learning and instruction. The chapter begins with a presentation and discussion of design experiments and extended design experiments from the tradition of experimental research and their relations to practical feasibility and to different traditions and paradigms from the philosophy of science. Then, common and new metaphors for the interpretation of effects for the empirical levels of the methodological assumptions are introduced and discussed against the backdrop of applied theory of learning and instruction.

Keywords: Learning and instruction; design experiments; mental model theory; learning instances; effect metaphors; research methodology; philosophy of science.

Introduction

The theory of mental models (Johnson-Laird, 1983; Seel, 1991) is the basis for a generalized theory of model centered learning and instruction (cf. Seel, 2003). Findings from mental model theory can help with this deductive process. However, the assumptions for MCL&I are (and have to be) investigated separately with specified research questions concerning the embedded processes of learning and instruction. Both theories are linked and have been investigated for a long time. They are still among the few theories available which contain strong, complex and well-investigated internal (model coherent) assumptions which bring about consistent hypotheses leading to data which do not falsify the assumptions. In this chapter I would like to share some central ideas about how research on the bridge between cognitive psychology and theories of learning and instruction can be conducted in general and what methodological and logical traps may come with this special endeavor. The ideas which I introduce are based on the problems and solutions from our research on mental models and model centered learning and instruction as well as on the methodologically sound traditions within these fields. The solutions presented in this chapter are thus completely applicable to the theory of mental models and to MCL&I. I have kept the statements and conclusions as general as possible in order to ease the transfer to other domains and theories, which will in turn also have referring fields (e.g., in cognitive science) and theories of learning and instruction.

Basic Definitions

In this paragraph I will give some basic and very short definitions of terms as I use them in this text. It is too incomplete to serve as a general glossary. For a more detailed overview see Kuipers (2007).

- *Axiom*: A self-evident *assumption* which can not be proven or demonstrated but which is considered to be true for the time being. Axioms are always needed for theory building. One of the goals of good theory building is to specify the axioms which are needed, to keep their number low and their content as basic as possible.

- *Theorem*: Anything which necessarily follows from a set of axioms and/or other theorems when strict logic (deduction) is applied. Simple theorems can be derived from a small set of axioms. Complex theorems may consist of numerous combinations of axioms and theorems. Theorems have to be formally provable.

- *Assumptio*n: A consideration about something that *may be* true on the basis of what we already know from existing axioms and theorems (induction). An assumption is like a theorem and has to be either basic, simple and necessary to be considered an *axiom* or to be tested in the real world by *hypotheses*.

- *Theory*: A self-contained set of *axioms*, *theorems*, and strictly bound *assumptions* all belonging to a defined subject domain. The goal of a theory is to systematically and unambiguously explain a simple or complex, abstract or concrete phenomenon or sets of such phenomena in the world.

- *Theoretical Construct*: A theoretical construct is generated when a theory combines sets of phenomena and subsumes them to describe them as if they were a single object. Famous theoretical constructs can be found in physics, e.g., the black hole, the atmosphere (of a planet), and light. If theoretical constructs describe phenomena which can not be observed or measured directly (at least at the moment), then they often use analogies or metaphors to consistently describe causes of a certain behavior which is observable, e.g., "memory", "storage and retrieval", "mental model", "anchor" (as in anchored instruction), or "bottle neck" (as in some theories on attention). Most theoretical constructs describing internal (cognitive) states are metaphors.

- *Hypothesis*: A prediction which either a theorem or an assumption necessarily induces. A hypothesis is something which can be empirically tested (in the real world) and which can either support evidence or be proven wrong.

- *Design*: A bound subset of assumptions from a theory which serve a specific goal. In industrial design goals may serve the purpose of improving, e.g., cost, functionality, aesthetics, and originality. In instructional design these are in most cases learning goals. Goals are derived from an intention (e.g., the intention of teaching or learning a certain skill).

- *Application*: The concrete result when the design is implemented in a single case. The application is the actual product which works, e.g., in a classroom and/or runs on a computer. Specifically well-working applications (determined by evaluation) can be considered as "best practice examples" to illustrate the benefits of a *technology*.

- *Technology*: A generalized set of common applications (e.g., learning systems which are based on mental model theory) or a guideline for the development of *applications*. Both follow a specific *design*. Technology in this sense can include but is not limited to the use of computers and other machines.

All Good Design is Based on Theory

Although they can appear randomly or naturally, learning environments usually are designed on the basis of an intention to teach or to instruct. Intentions are useful when they are implemented in the real world. We call this transfer *instructional design* (including everything that is needed to be done to call the transfer a success). The translation from an intention to the act of implementation can be seen as a reasoning process. Success has to be carefully monitored by summative and formative evaluation and the appropriate methodology. A good learning environment enables a learner to reach goals, to solve problems, to gain knowledge and experience, etc. An environment is designed to mediate between the learner and a given goal or set of goals. In the design phase the success of a learning environment is mostly hypothetical. Since the application does not yet exist, its real outcome is always vague, no matter how convinced we are about our design. Former experiences do not always apply; at least we can not be sure to what end they will hold true for each new instance (of experience), especially when they are related on the basis of intuitive experiences rather than common traits. Evaluation can help one to improve over each cycle of redesign so that the applications get better every time. The hypothetical nature of the design phase explains why good instructional design is always based on theory: Every design carries multiple risks of failing partially or completely. Every theoretical assumption can be made into a design constraint through deduction and will thus improve the chances for success, assuming that we apply the right theories. In other words: We can be surer about the success if we use assumptions which have already been carefully investigated over the years. The translation from theory to design constraints is also a process of reasoning. Real learning applications and technologies can then be based on theories of learning and instruction, e.g., on *model centered learning and instruction*. Good theories can be found wherever these three steps are available:

1. *Versatile*: Transfers (transitions) between intention and implementation are open for a large variety of applications.

2. *Strict*: Transfers between theories of learning and instruction and design are (logically and methodologically) as strict as possible.

3. *Strong*: Assumptions have been investigated and tested carefully in subsequent experimental research.

It makes sense to have a closer look at what this means for the *domain of cognition and learning*, for the *field of instructional design* and for the *cooperation* of both.

Aiming at the Domain of Cognition and Learning

Researchers in the domain of *cognition and learning* usually start their interest at step three. Of course a theory has to be operationalized, self-contained and not replaceable by a longer existing theory which describes the same or by a new theory which describes the same effects in a better way (true only in rare cases). But additionally a good theory is always specifiable (without necessarily being reducible) and likewise generalizeable. This includes the notion that a theory does not always already have to be specified or generalized. When developing theories in the domain of cognition and learning, both directions are of great importance.

Specifiability is needed to pass the constraints from the underlying cognitive mechanisms through to the level of applications. This includes theories about applications and the technologies themselves. If we know, for instance, that a four-year-old child is unable to process abstract reasoning it would not be wise in any case to try to teach it theoretical ethics. This is an obvious example for illustration purposes (and may also turn out to be wrong), but effects like this can easily constrain all kinds of applications in a similar but not so obvious way. It is the task of a good theory on cognition and learning to pass constraints like this actively and within the theoretical framework through to instructional designers who rely on the theory: If things do not work cognitively they can not be implied for the design. Generalizability, on the other hand, is needed even more directly – and it differs from something what we would discuss as *applicability* or *feasibility*. It means that the assumptions of the theory can be generalized in terms of their effects on and within the actual technology and that its hypotheses can lead to assumptions about applications *directly*. This is where step two is of utmost importance: in preventing the use of the theory merely as a loose source of association and justification of professional action.

In order to work properly within the span of basic research and applications, theories (as a whole) and their assumptions as well as all their refinements will have to be:

1. Specifiable (but not necessarily already specified)
2. Generizable (but not necessarily already generalized)
3. Operationalizable (*and* necessarily operationalized).

Following the Field of Instructional Design

Instructional designers will start looking at theories which can serve as the basis for not only single applications but a whole class of applications – thus starting at step one. Unfortunately, theories which are not generalizable and therefore are used for "wild" association and justification have – exactly because of their vagueness – a tendency to create the illusion of fulfilling step one perfectly. But theories with a strong empirical foundation, strong assumptions and clear findings (step three) are more likely to support good technologies, although the transfer (needed in step one) is usually more demanding. To identify such theories, a good understanding of the methodologies which are used to test the theories is crucial for the whole design process as well as for the implementation: Theories coming from the domain of cognition and learning do not always fully cover step two because in progressing with the work on the theory, no researcher can possibly think of all future applications. Fulfilling step two is an ongoing process and maybe the most crucial one when linking theory to practice. Since we do not have the luxury of working in interdisciplinary teams on every design task, we need for the theories to be “in shape” for the reasoning needed for the transfer. And “in shape” means that they have been tested and reported in a way which makes strict conclusions for the transfer easier or even possible. Identifying single or simply combined effects may suffice for many general questions about cognition and learning. However, from a design perspective and having the whole application in mind, a bunch of separate effects rarely gives consistent guidance for the assembly of an application. Likewise, they can not be combined like a set of independently working effects. This may raise a demand for new methodological approaches to test theories. Some of the existing methodological ideas for simplifying the transfer process will be introduced and developed later on (e.g., the design experiment and the extended design experiment).

For the sake of completeness (although it may already be obvious): There is no other way around. It is by far not advisable to construct an application and search for a spontaneously fitting theoretical foundation afterwards to justify the design. This holds true *even if* an application randomly turns out to have good evaluations. What is at stake, especially in everyday decisions about design, is not just making something work – it is about *improving* what already works somehow.

Classical Experiments and the Common Effect Metaphor

All good theories are based on methodologically sound research. For me as an empiricist there is no specific paradigm for instruments for empirical research. What is to be measured determines how it has to be measured. What is to be observed determines how it is observed. I would not measure heart rate with a ruler nor would I measure flight distances with a micrometer screw just because I like micrometer screws or because I’m used to them. The methodology is dictated by the research question. It’s not the other way around.

Measurement and Instruments

Measurement assigns numbers to properties in a meaningful way, and it does nothing else. Neither can it describe research objects directly (only properties), nor can a procedure which circumvents the systematic observation of properties be called “measurement”. The instrument which allows a measurement follows the research question. If a research question requires an instrument, researchers will have to learn how to use it. If there is not yet an instrument to measure the required properties then this instrument has to be developed before further research is carried out. There are properties of research objects which do not yet have instruments. Speculation about the properties or about the objects may fill some (but not all) gaps between research and practice. However, pure speculation will never lead to persistent and coherent theories – as long as they are theories about this something we like to call the “real world”. For some theoretical constructs there is more than one instrument for measuring the right things, and the discussion about the right assessment, diagnosis, and measurement is an ongoing process which will never come to an end as long as theories evolve. This is also why I fail to see that quantitative and qualitative strategies for gaining insight into a set (or configuration) of properties are different paradigms. To me they are different approaches for describing the object of interest. Some of them may be used in tandem (e.g., within mixed-method designs) and some of them may just be steps in measuring completely different properties. Both strategies have to follow high standards and neither of them should be (but sometimes are) used to excuse a lack of skill or effort in research procedures. With the right instruments at hand we can apply them to the research questions, which will hopefully lead to a better understanding of the theory.

Experiments and Operationalization

The experiment is a specific procedure for investigating components of theories and their effects. The basic idea is to separate an effect from the “disturbing noise” of its common surroundings to make it easier to “listen” to its internal workings (internal validity). A very comprehensive overview about the different modes and applications of experimental research is provided by Campbell and Stanley (1966). When researchers talk about learning and instruction, they mainly talk about cognitive constructs, their interaction with constructed environments, and how both of them progress over time (see Seel & Ifenthaler, 2005). The constructs are metaphors because they do not describe physical states but make assumptions about reasons for systematic behavior of a learner. For example, many of the cognitive assumptions are still inspired by the computer metaphor (e.g., memory and central processing). This holds true whether or not one follows a materialistic theory (or even identity theory) – as long as the physical boundaries are not completely described and measured, the descriptions of the behavior-causing systems remain on a metaphorical level. And thus their “ins” and “outs” have to be operationalized in order to measure their properties on the behavioral level. Only that which leads to a certain behavior can be observed. The operationalization is then the crucial a priori guarantor for quality – long before measures of validity, reliability, and objectivity are reported a posteriori. Of course, strong theories are more likely to found good operationalizations because of the strong and specifiable assumptions – which do not necessarily have to lead to reductionism. Experimental research on learning and instruction implies that each measurement which is required by the operationalization can be induced (treatments), tested (instruments), and controlled (statistics) separately.

Quasi-Experiments

A set of separate tests is then combined as evidence or counterevidence for a theory or a section of a theory. Since classical experiments are sometimes unpractical or even unethical to conduct, e.g. in longitudinal studies, quasi-experiments can be used instead. The lack of controlled criteria and confounding variables as well as the inability to randomize the data has to be compensated by higher methodological requirements. The reason for the appearance of quasi-experiments was not only the constraints from practicability but also the further development of theories, which became more complex and specified over time.

The Evolution and the Complexity of Theories

We can clearly follow the path from traditional theories about very simple metaphors for memory structures, e.g., working memory assumptions (see Miller, 1956) or long-term memory experiments (see Ebbinghaus, 1885), to more recent theories about the simulation of virtual and real environments (see Archer et al., 2006; Curiel et al., 2005; Endsley & Smolensky, 1998; Leuchter & Jürgensohn, 2001), anchored instruction and goal based scenarios (see Blumschein, 2003, Schoenfeld-Tacher, 2001), system dynamics (Spector, 1995), and mental model theory (see Al-Diban, 2002; Gibbons, 2001; Hanke, 2006; Ifenthaler, 2006; Pirnay-Dummer, 2006; Seel, 1991, 2003; Spector et al., 2005) – and many more could be named. Of course it took decades of research for these theories to evolve. The application of quasi-experiments led to a methodological development in which the multiple dimensions of interacting assumptions within (and against) the model of a theory could be traced (e.g., Breslow & Clayton, 1993; Campell & Stanley, 1966; Long, 1983; Raudenbush & Bryk, 2002). Thus, theories became much more than a list of five separate statements. They became complex models themselves and were thus more fit to describe the real world. Of course it takes longer to study them properly than it does to study the simple lists of assumptions, which explains why dated theoretical basics appear on the agenda of research every now and then – especially when research is mainly interested in application.

Cognitive Modeling

Almost in parallel, cognitive modeling became a strategy for forcing theories to be formally more precise (Cooper et al., 1996), which leads to a better understanding and to a better operationalization.

> Notwithstanding the many ambiguities of cognitive theory, the criteria by which experimental research is judged are public and largely uncontroversial. On the theoretical side the situation is less satisfactory. Most theoretical discussion in psychology is informal. Theories are often presented in terms of box and arrow diagrams …, or draw on natural metaphors …, but the box-arrow notation and its underlying assumptions are generally poorly specified, and it has been argued that “attractive metaphors raise more obscurities than they resolve” (Cooper et al., 1996, p. 3).

It also comes with an additional advantage: Cognitive modeling (e.g., Anderson, 1983) forces the terms and boundaries of a theory to be defined properly rather than to have a “halo” of connotative associations which everybody can accept easily in his own theoretical context.

Common Effect Metaphor

Classical experiments and quasi-experiments can elicit single and multiple effects when they are conducted using the right instruments. When these insights are transferred to practical applications (e.g., in classrooms), a very common effect metaphor is implicitly used. This metaphor can be described as an “engine metaphor”. A class is composed of individual learners. A specific treatment is applied when it shows promising effects on learning within experimental research, like an *engine* which expedites or facilitates the learning progress for the group.

Design Experiments and Feasibility

The Design Experiment (DE) as introduced by Brown (1992) is a methodological framework for implementing and testing learning theory in classrooms. It was developed mainly to solve the problems of external validity which occur in classical experiments: When we cut the effects out of the natural learning environment for experimental research we can not make sure whether the effects are strong enough to play a role in the real environments. Sometimes we can not even be sure whether the theoretical constructs and their properties are at all observable in practice. Although the insight had been secured for experimental research, the transfers into practice were more of a trial-and-error style and had to rely on subjective plausibility only. Hence, the process of design, development, implementation, and evaluation is too time consuming to conduct unsystematically. Brown states that a

> critical tension in our goals is that between contributing to a theory of learning, a theoretical aim that has always been a keystone of our work, and contributing to practice. This is intervention to inform practice. For this to be true, we must operate always under the constraint that an effective intervention should be able to migrate from our experimental classroom to average classrooms operated by and for average students and teachers, supported by realistic technological and personal support (Brown, 1992, p. 143).

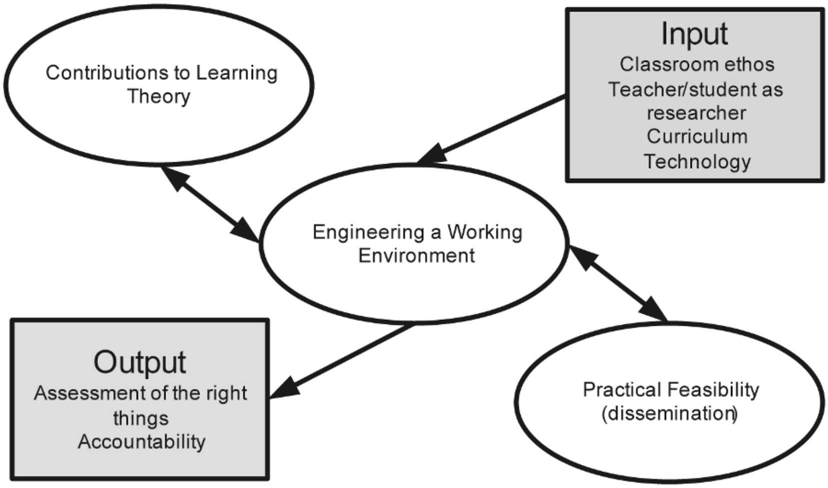

Fig. 1. Design Experiment (Brown, 1992, p. 142).

On the basis of these assumptions Brown introduced the methodological framework for Design Experiments (see Figure 1).

The main vision of this framework is that the contributions to learning theory and practical feasibility can be orchestrated by carefully engineering a working environment between the available input (things which are given to the environment) and the desired output. It is obvious that it requires for skilled researchers, instructional engineers, and teachers to work closely together in situ to achieve this high goal. But even great efforts will be very well justified because of the expected outcome. If a design experiment succeeds, three major benefits can be expected:

1. The learning theory hypothesis can be tested.
2. The implication from learning theory into technology will have to be clarified.
3. A documented prototype of the learning environment already exists, including evaluation and assessment of transfer.

In classical experimental research these three steps are usually sequential and spread out over different groups (research and design) – if they are carried out at all. But combining the steps within a single experiment may not only help just because they are conducted at the same time. They would also have to be realized by the same working group, thus forcing the researchers to think about everything from theory to implementation. So far, Design Experiments allow us to test designs and applications in living environments and to explore the boundaries of theories on learning and instruction. They help us to improve applications and to improve the design process itself. From design experiments we can get all the hints we need between practical feasibility and theory-based design to improve on developing theories and on implementing them right away. As in all good field research, design experiments have high demands for precise empirical methods because real life learning environments have a tremendous amount of factors and effects which have to be part of the theoretical model. We can often find common misunderstandings which say that design experiments can help us to ease the process of research. This is untrue and clearly leaves the path of Design Experiments as introduced by Brown.

Neither can we spare systematic research only because we are observing a real learning environment, nor is it possible to replace experimental research by just evaluating the success of an application. While the first assumption should be obvious for every serious researcher, I will go into more detail on the second statement in the following paragraph.

Retracing the Impacts of a Theory in Applications

In the following paragraph I will try to share some insights into how theories affect applications in general. I will do this in the text and in a more formal way. Both representations have their benefits, depending on what we need to use them for. Please feel free to use them separately or in combination, just as you need them for your studies, work, and thought.

Design as Deduction

As I pointed out, design is a reasoning process. It is a deduction of design constraints for application on the basis of the theory. The theory forms the premises for the conclusion. The results are the design constraints which basically form crucial parts of the application.

A_W is an application space (e.g., with a specific goal) within the domain of all empirically possible applications Ω_W. The space of applications Ω has to be selected so that all of its applications can be realized in the real world W: Ω_W. Constrained by learning goals, institutional limitations, etc., A_W is realized as a subset $A_W \subseteq \Omega_W$. If the single application $A^* \subseteq A_W$ which is constrained (designed) by the theoretically based deduction

$$Cn\left(T^{n-m}\right) = T^{n-m}$$

then T^{n-m} is the part of the theory which describes selected statements (between {n,m}) of the theory T. Cn is the deduction function. Since deduction always retains the truth values of the premises, each deduction necessarily produces a tautology. Every deduction from a theory can be realized in Ω. This is what always happens *during* design and *before* implementation. However, we can not show that deductions from a theory $Cn(T^{n-m})$ necessarily lead to applications which can be realized in Ω_W (see Figure 2).

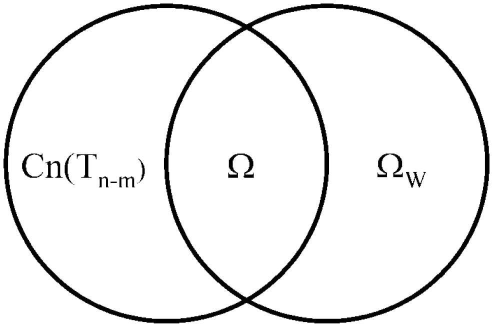

Fig. 2. Intersection of Theory-Based Deductions and Possible Applications.

Operationalization

Obviously the idea of instructional design demands that theories can be operationalized not only for empirical testing but also for practical applications. Both kinds of operationalizations differ in terms of what they aim at and what they can help us with. The *empirical operationalization* helps us to observe the testing of our theories in different ways while the *operationalization function f towards the application* $f(T^{n-m})$ helps us to come up with good designs. I will focus on the latter kind of operationalization.

Three Cases of the Theory-Design Relation

It can always happen that a change to a theory based on new empirical research falsifies prior deductions for design. This does not mean that the applications will suddenly be wrong, too (e.g., evaluation may still say otherwise). However, it will mean that the deduction was not the process which led to the application:

$$\exists x \rightarrow x \notin A_W \wedge x \in Cn\left(T^{n-m}\right)$$

In these cases, either the new theory can explain the success of the application or we just have a good application without knowing exactly why it works, which is still good but makes it harder to design similar applications in the future. All other implementations necessarily have to be one of the following cases (A fourth case could formally be described as $A^* \cap f(T^{n-m}) = \{\}$, which is left out because it would not be applicable for theory-based design – the application has got no intersection with operationalization of the theory):

Enclosed Design

If an application or a design is completely described in all possible aspects by a theory (and needs no further description at all), $A^* \subseteq f(T^{n-m})$ then the theory "generates" the design $f(T^{n-m}) \models A^*$ (Figure 3). Taking into account the deduction theorem, $\models A^* \Rightarrow f(T^{n-m})$ also holds true. Everything that can be observed for A* can be traced back to the theory. However, as long as the special case of a theory describing exactly one design $f(T^{n-m}) = A^*$, the design is over-specified compared to the theory. A complete matching may be possible for artificial learning (e.g., machine learning) but will presumably remain impossible for human learning processes.

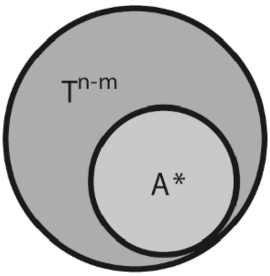

Fig. 3. Enclosed Design.

Enclosed Theory

If a theory makes assumptions within an application or design and consists of no models which are not contained in the application $f(T^{n-m}) \subseteq A^*$, nothing follows formally as long as the input and output conditions are not fully described for the operationalization $f(T^{n-m})$ (Figure 4). (Very long sentence with three negations. Suggestion try to split it.) The design experiment may follow such an approach. When it is properly arranged, it will always fall under this case. This is especially applicable for theories with a high level of generalization and a small number of variables, such as basic taxonomies of metacognitve strategies (cf. Brown, 1992).

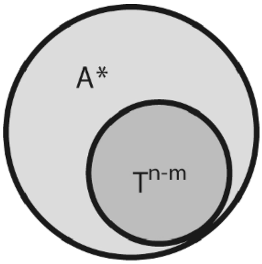

Fig. 4. Enclosed Theory.

Intersection of Theory and Design

The last case describes designs and applications which are to some extent described by theory and which have parts which are not described while the theory makes assumptions which are not applied into the design (Figure 5). This would be realized formally as

$$\exists x \rightarrow \left(\left(x \notin A^*\right) \wedge \left(x \in f\left(T^{n-m}\right)\right)\right)$$
$$\exists x \rightarrow \left(\left(x \in A^*\right) \wedge \left(x \notin f\left(T^{n-m}\right)\right)\right)$$
$$A^* \cap f(T^{n-m}) \neq \{\}$$

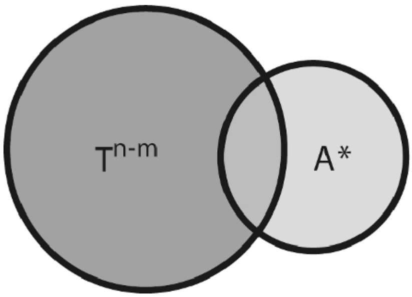

Fig. 5. Intersection of Theory and Design.

To describe the intersection properly it will be possible and necessary to postulate hypotheses on designs and applications which not only carry basic assumptions (which generally is sufficient in classical experimental research) but also structure from the corresponding theory. This can be done in different ways depending on the theory and will be discussed later on.

In research about learning and instruction we generally have to consider the third case. On the one hand we have the complexity which automatically arises when the assumptions are embedded in the design during operationalization (Einsiedler, 1997). The need for multiple causal references within a theory of learning and instruction can rarely be limited to the comparative functions of a classical experiment. This is not only a matter of external validity, as has often been discussed (Bracht & Glass, 1975), but also part of the models of complex and strong theories in general. Therefore, the same is true of internal validity (cf. Ifenthaler & Seel, 2005).

Thus, it may always be that crucial parts of the theoretical assumptions are not responsible for the good outcome of the learning environment although they still have an impact on the environment. This impact may, of course, not be the one the designers had in mind, even if the overall outcome of the environment is a success in learning. This can (among other strategies) be shown if we trace the navigation of computer-based learning environments and find out that the learners in fact follow different ways than theory would predict (cf. Dummer & Ifenthaler, 2005).

Conclusions and Constraints for Research

In accordance with the relations between theory and design described above, huge differences exist between interests of evaluation and research interests of theory testing. Good evaluation results undoubtedly indicate the success of an application and therefore also of the design. Campbell and Stanley state:

Much of what follows is an evolutionary perspective on knowledge (Campbell, 1959), in which applied practice and scientific knowledge are seen as the resultant of a cumulation of selectively retained tentatives, remaining from the hosts that have been weeded out by experience (Campbell & Stanley, 1966, p. 4).

However, these results can not simply be traced back to say *anything* about the theory. Doing so would lead to a modus tollens error.

The design process strictly uses deduction in the form of implications. Implications are directed and not necessarily reversible, even if the reversion may sometimes be a successful heuristic: a heuristic is not sufficient for theory testing. In other words: there can be many reasons why an application works fine; it does not have to be the impact of the theoretical implications within the design as long as

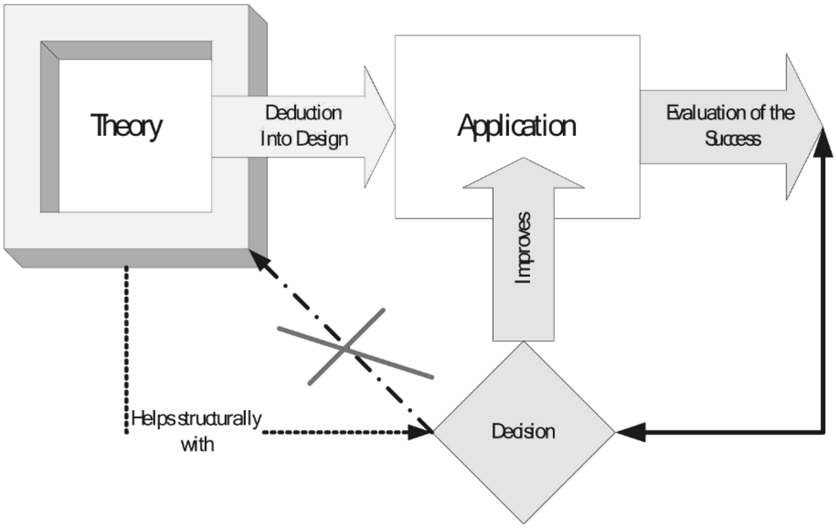

Fig. 6. Modus Tollens Error in Evaluation Based Decisions about a Theory.

we do not have the first case, where a theory describes every single aspect of an application (enclosed design) (Figure 6). However, as I pointed out above, within theories of learning and instruction we generally do have the third case (intersection of theory and design). Given this constraint, dilemmas arise within different philosophies of science. I will discuss three prominent schools of thought and the arising dilemmas based on the assumptions for each case separately.

Dilemmas within Positivistic Approaches

Logical positivism claims that a theory can be verified. A theory can be verified if its inherent logic comes from immaculate deductions (which are theorems) from comprehensible or congruently viewed premises (cf. Ayer, 1959). The premises emerge from inductive gain of insight. The hypothetical constructs of a theory are seen as being potentially true if they are applicable to at least one possible world (parallel world). It is very well possible to have contradictory assumptions which can exist in parallel. Contradictions can lead to a dispute (even to different "schools") or they can be combined to form a synthesis in a new set of premises.

A state of knowledge is derived from a significant set of different opinions which agree or disagree.

Empirical research in the social sciences has a tendency to present itself as having a rationalistic background but works mainly with positivistic premises, using the expression "finding evidence" to avoid the term "verification". This is closer to the classical thought experiment than the search for falsifying evidence. However, communities in specialized research areas (e.g., model centered learning and instruction, cognitive load, anchored instruction, system dynamics) reach the critical amount of researchers following a specific opinion much earlier than in the humanities, where methods of verification have an efficient tradition and where the methodology works with much larger groups. The discourse will thus be accepted as a standard far earlier. The Goodman Paradox (Goodman, 1965) shows how this can affect the observation process, which can become a significant factor. If verification is done empirically, it still uses all of the instruments which were originally invented for falsification. Hence, the process of verification of hypotheses within applications is quite easy to conduct. If some aspects of the operationalization can be observed coherently, this can be seen as a positive example of the working of the theory. Usually researchers from our field use probability values (for errors) of 5% in order to report effects. Statistically, every twentieth study will find effects just by chance. Given that studies with verification intent are very rarely published if they fail to find statistically significant effects, this may make us think twice about the idea of "finding evidence" even if we are considering systematic replication studies. In a falsification attempt two things can happen: If the corresponding tests turn out to be significant, the theory has to be modified to explain the alternative hypothesis. Otherwise the theory can be kept as it was. Both results can be interesting to discuss. In a verification attempt only one result can be of interest. If the tests are significant there is evidence. If not, nothing follows. Therefore, a methodology has to be designed for verification purposes. Its empirical testing mechanisms accordingly have to be different from tools used for falsification.

Dilemmas within Rationalistic Approaches

Critical rationalism claims that a theory needs to be falsifiable. This follows from the unsolved induction problem (cf. Popper, 1934, p. 3). A theory can not be verified with inductive methods because we would need every possible case (complete induction) to do so. A theory can never be proven true. Therefore, a theory has to come with at least one possible way of proving it wrong. Because of the induction problem, theories are always hypothetical. They differ in how well they can explain things that happen in the world. The easier it is to potentially falsify a theory (with the methods of experiments or observation) the higher is its scientific value and the more interesting it is for science and researchers, given that its domain allows valid predictions about the real world. This also leads to the assumption that there is a high probability that every theory *will* prove wrong eventually.

In principle, falsification (cf. Popper, 1934) is not a priori suitable for the further development of theories. There is no clear explanation, for example, as to whether the goal state of the process of progressive falsification lies in the refutation of all theories whose models become increasingly incomplete during abstraction and generalization (include too little information) and which thus must lead to falsification or whether this process leads to a copy of the world – which, although it would be an interesting endeavor, has little to do with the goal of an abstract description of the world. Every description is – as long as it is not an exact copy of that which is being described – a simplification. But since every simplification is incomplete and every incomplete explanation must at some point lead to falsification, the termination condition of critical rationalism will likely remain open in this matter. It might well be a facet of human nature that we are not eager to immediately deconstruct the models of the worlds around us, especially considering how much time and effort we put into constructing them and increasing their complexity. This, however, is exactly what the epistemological position of critical rationalism demands of us – and for good reason. It comes as no surprise that a replication which achieves less evident results than its experimental predecessor will attract little attention and, moreover, will even lead to a more meticulous search for errors in methods and conditions than twenty original studies. Theories and theorems which are widely known or advocated by leading research institutions sometimes experience something of a forced immunization (such as through the introduction of research standards). This became thoroughly evident in cognitive science in the heyday of the discussion over contradictions between theories of mental models and mental logic (Braine et al. eds., 1998), in which the opposing fractions accused each other's theories of having been subject to such an immunization. Almost all known scientific periodicals base their decision on whether to publish a research study or not primarily on the evidence it produced in support of a theorem. Only rarely are articles published which describe the downright refutation of a theorem, even though this still constitutes the epistemological basis of the methodology of empirical research. The tendency described above of referring to one's "own" theory is increased by this fact. This leads to the circumstance that practical research methods, even those which are situated within the epistemological field of influence of empirical research, often do not receive any "empirical" testing worth speaking of. Holzkamp (1996) illustrates this problem and expresses his doubts as to whether learning theories include real empirical hypotheses or whether they are not simply patterns for backing up hypotheses which require that one uses one's "good sense" between their if and then components.

Dilemmas within Constructivistic Approaches

Weak constructivism (cf. Ernest, 1996) – in comparison to radical constructivism – claims that knowing can be constructed by inter-subjective agreements due to processes of structured communication between individuals. In these processes, theories and parts thereof can be falsified *and* verified. Verification is an impulse of assimilation: A theory may claim its range as long as there is repeated evidence from examples which are somehow plausible or agreeable to all individuals. It has to be shown for each theory – *tertiam non datur* – that its constructs can not be reduced to a verified theorem of another theory which is older, more renowned, or which makes more extensive assumptions. Only if a theory has something significantly new may it be considered. A theory may assimilate new theorems and even new axioms, which is empirically equivalent to just finding significant evidence in samples, e.g., correlations, differences, etc., A theory may grow (almost) infinitely during this process. Therefore (but not only for this reason), a weaker form of falsification is necessary, which can be described as an impulse of accommodation: We can go on verifying the theorems of a theory as long as there is no falsification within the empirical predictions of each theorem. The "deconstruction" of assumptions (partial deconstruction of theories) needs to be sought continuously by the the scientific community. Like in model building processes, accommodation is harder to realize and will be conducted far less often than assimilation. However, the impact of a single falsification (even the attempt) is generally stronger than evidence from a single verifying experiment. If explorative verification and testing falsification are properly combined with good a priori work on the consistency of the theory, values of both insight and feasibility will be held at viable and constant levels. Within the pragmatic boundaries of modern research (e.g., limited funding, time and personnel, the pressure to publish frequently, well justified ethical constraints), this may by all means be seen as a satisfactory tradeoff. I will discuss this tradeoff later in more detail when I introduce the idea of an "enhanced design experiment".

A suitable combination of discovering verification and the verifying falsification principle serves to ensure that the status of the theoretical complexes, their constructs in the specification process, and their integration into the model in the generalization process will be as consistent as possible at any given point in time during research. The process of identifying the domains of intensified research interest is then primarily a matter of searching for existing inconsistencies, be they theoretical, formal, or methodological in nature. Accordingly, interest in a theory will lie in a combination of its ability as a relatively vague (but not arbitrary) construct to explain something which really is new as well as to produce evidence and test contradictions. Together, these factors result in the general ability of the theory to create an explaining and knowledge model of a describable world. The tools of *inductive logic* (Carnap, 1950; Carnap & Stegmüller, 1958) can be used to construct a probability model in which probabilities of error used for falsification can be placed directly on the relations. In this way, verification and falsification may be seen as logical extremes, as frames for every probabilistic observation during research. Provided that the empirical data (hard evidence) is present, every assumption can be tested for a conditional probability which emerges for the hypothesis h. This is an approach to knowledge construction which has not lost its significance in times of increasingly complex models, especially for dealing with causality references (Wainer, 1989; Pearl, 1995, 2000; Halpern & Pearl, 2001).

The solution to the dilemmas resulting from verification and falsification seems at first glance to be simple. A pragmatic synthesis of the advantages of these two vastly different epistemological approaches enables one to construct, by way of a largely inductive process, a world view which elegantly integrates the mutual control mechanisms of both methodological positions. But there are more than obvious limitations even within constructibility. One of them is a lack of determinability with contradictory models. Contradictory models are not discounted by way of a primary falsification process. The constructivist theories are, as it were, innately more immune to refutation, i.e., to accommodation through other approaches. A further problem with constructibility originates not in the epistemological approach itself, but rather in the fact that its methods are occasionally falsely assumed to be arbitrary. A lack of exactness cannot be justified by a change in epistemological foundations. It is not enough to use purely descriptive calculations to test the constructs for the quantitative-empirical part of research, just as it is not enough to use a naive theory of textual understanding as the standard for the qualitative parts. Our description of the constructivist methods has already made it clear that in the main they do not have their own testing and decision methods. This is admittedly not unusual for an epistemological line of theory whose modern approaches produce young offshoots. The difficulty, however, lies in the application of a synthesis of methods which incorporates widely contradictory methodological traditions. It stands to hope that researchers will succeed in using this synthesis in the best way possible. However, what still remains to be developed are standardized procedures and maxims for the decision-making process. Proponents of the construction approach are by and large in disagreement as to whether it is possible at all to decide between contradictory paradigms through logic and experiments (cf. Kuhn, 1976). If this is not the case, however, the consequence would be that the changes – and thus also the construction of knowledge itself – could not be explained through rational means, but rather only through phenomena in the history of science and sociology (cf. Lauth & Sareiter, 2002, p. 135). Thus, one searches in vain in Kuhn's article for a plausible explanation as to how it is possible for a new theory to displace an established theory; and the belief that big theories simply die out with their proponents is questionable, especially considering the nature of scientific careers, e.g., where students and also junior researchers depend very long on their advisors.

Extended Design Experiments (EDE)

Inspired by the idea of design experiments, we can use extended design experiments (EDE) to close the gap between evaluation and theory testing (Figure 7). An EDE is implemented like a classical experiment, except that we do not take away as much complexity as we can to control the confounding factors. We try instead to control the factors statistically according to the theory and as part of the complex model, which leads to a set of hypotheses. Thus, we do not look for single effects, e.g., between pre- and posttests, but for the realization of the whole model:

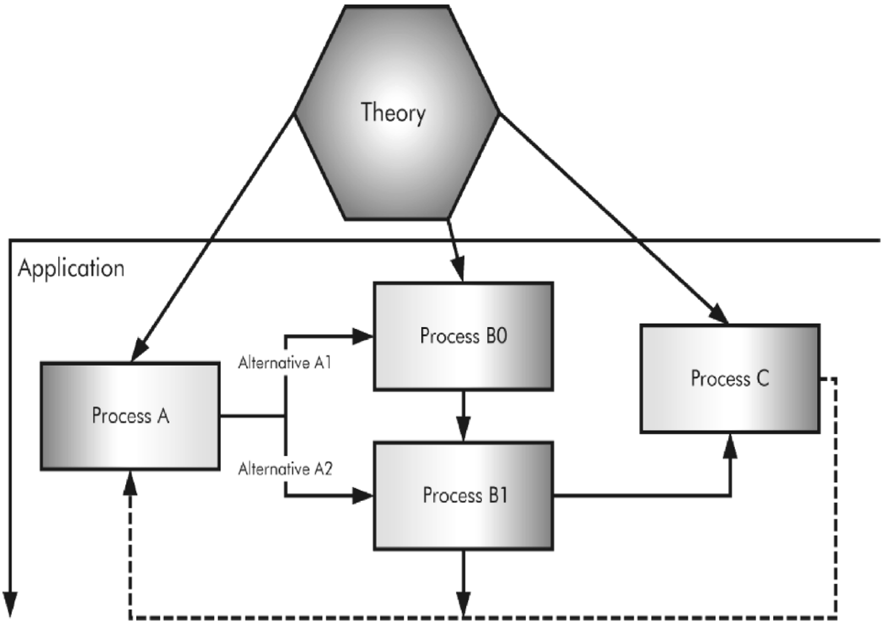

Fig. 7. Complex Predictions for Structural Support.

How well can our theoretical model predict what will happen within the application? How frequently will predicted actions take place? Are the predicted sequences of actions and processes systematically observable?

EDEs may still answer questions if something is happening, but they allow us to determine how things are proceeding and whether this is congruent with the assumptions of our theories. They may be constructed independently from the philosophy of science we are following. They can solve some of the dilemmas which I illustrated above: Because the structure of the theory leads to more complex hypotheses, the mapping of the structure on the application will be a valid indicator for the impact of the theory. In this way, the EDE may solve most of the problems which arise because of the *enclosed theory* or the *intersection between theory and* *design*. On the other hand, neither can EDEs replace good strategies of evaluation, which are still needed to test the success of applications, nor can they replace the classical design experiment, which is good for testing whether and how theories of learning are convertible to existing, living environments. We can, however, use the EDEs to test theories where existing environments are not (yet) suitable for bringing the operationalizations of the theories to life, e.g., if actual organizational structures hinder a complete realization of the design. EDEs can still be implemented without giving up the chance to test entire applications. In the following paragraph I will go into some more detail about the inner workings of structural support.

Predictions for Structural Support

To separate theory testing from evaluation we need hypotheses which postulate structure from the theory into the design. E.g., things have to proceed exactly in a specific way in order for us to believe that the impact of the theory did indeed produce the positive effects. It sometimes occurs that the application and the theoretical domain cannot be clearly separated from one another or mapped onto one another and one must nonetheless find a practical and valid possibility to construct and test the theory. I will give simplified examples later on (cf. 7.2.1 – 7.2.3). In such a case, it is necessary to dissolve the inner conditions of the design experiment and set them up again with the structure of the hypotheses in mind. Two static structural interfaces which are especially visible are sequence (cf. 7.2.1) and frequency prognoses (c.f. 7.2.2). Dynamics can interface to most hypothesis while measuring discrete change over time (cf. 7.2.3). For instance, model based learning and instruction postulates changes in understanding at key positions which are bound to certain specific interventions over time (cf. Seel, 2003) to reach, e.g., effects of accommodation, which are hard to reach in learning environments. Thus, the effects depend on a larger multi-causal network which is bound to the hypothesis stating the effect. With simple intervention designs, some effects may still be shown. However, they can not be traced back to the hypothesis nor to the underlying theory. The structure of the hypotheses should be tested more carefully for suitability during a single experiment. It is still possible to achieve a passable resolution of the deductive elements of the theory by separating them into static parts. It will probably still be possible to conduct multivariate statistical tests for some time depending on how far a theory has progressed. However, in a similar approach Opwis (1985, p. 39) shows that there is no adequate mathematical model for continuous or discrete dynamic systems, to say nothing of linearity. Although this objection is pertinent – and will doubtlessly need to be confronted by researchers engaged in complex research on learning and instruction sooner than it suits them – it is possible to translate even the complex processes of theories on learning and instruction into static nodes which may be calculated separately but must be observed as a whole, provided that one exercises enough deductive caution. The important thing is to avoid simplifying the methods too much by translating them into bundled statistical hypotheses. Rather, it must be possible to retrieve the model structure and the underlying relations directly in the methodically founded prognoses. On principle, *any* model structure may be used, even very inhomogeneous or theory-specific structures, when the aim is to verify the assumptions rather than to test them against the prognosis qualities of other theories. In this last case, it is again advisable to decide on a standardized structure which enables a comparison, taking care in doing so that the scales and processes still adhere to obvious standards. Generally speaking, the testing of theories should always include a precise conceptualization of the prognoses which goes beyond the simple variable relations and a test against plain chance (without determining whether the distributions can emerge at all without an ascertainable system) (cf. Holzkamp, 1994). Otherwise, one can not be certain that the experiment accommodates a comprehensive prognosis for structural support. This does not preclude the possibility that there are theories whose structure is situated precisely in this relational mapping, such as the classical experimental methods. One cannot assume a priori that influencing factors which are not or have not yet been explained by theories emerge by chance. Holzkamp (1994, p. 85), for example, warns that the only way that such factors can be shown to be above the suspicion of chance occurrences – and thus to possess empirical content – is to determine whether a result can come about by chance at all. If, however, one compares the results garnered from such an endeavor with the theoretical structure rather than with the assumption that they came about by chance (e.g., H_0 = "there is no correlation"), it is no longer necessary to prove that they appeared by chance because the only remaining chance lies within the probability of error (given good planning).

Kinds of Hypotheses – Three simplified Examples

A good place to begin when testing the verifiability of theories is with the theories themselves. In the simple case that the parts of the theory one is testing, namely the theorems, predict differences between groups or correlations directly and nothing more, the learning and instruction researcher can take recourse to classical experimental methods of observation without any further ado. As soon as a theorem exceeds this propositional spectrum of the assumptions, however, it will open the door to hypotheses which require not only for the variables to be operationalized but also the propositional structures of the assumptions themselves – at least in the case that they cannot be separated into smaller units which can be tested individually. But it is not always possible to theoretically found a reductionist assumption directly. The corresponding stipulations that less abstract theories should be generalized and that the others should be made more specific do not allow the assumption of a reductionist model alone. On the one hand, the less abstract theory can simply turn out to be false, and if it is based on the reductionist assumption alone, this would lead to its refutation. On the other hand, specification can also turn out to be a false conclusion resulting from axioms which are not concentrated as much as those in the less abstract theory. However, this should not lead one to call into question the possibility that causal correlations can be virtual or the merits of reductionist considerations concerning falsification through proof of pseudo-correlations and other causal co-occurrences. These assumptions could also enable one to connect individual components of a complex theorem by naming specific unification functions, which would mean that they would remain testable for the components in classical operationalization methods (only variables). However, this would only relocate the problem since setting up the unification function to conform to the model should be easier than making the propositional structures observable only in rare cases. The following examples, which are simplified for the purpose of clarity, should give a basic idea as to how these assumptions could work in research practice and for the development of theory. Real hypotheses on actual research would be more complex and would thus need an article of their own to be laid out properly.

Hypotheses on Sequence

Sequences are found all over theories in very different types. We find them, for instance, in theories on cognitive apprenticeship (cf. Collins, Brown, & Newman, 1989) and in theories about emotion (cf. Kuhl, 1983).

As we can see in Figure 8, the model states sequences which only make sense when they are integrated into the full model. In Kuhl's model it would not make sense to separate, e.g., effects from conceptual-semantic characteristics on autonomous reactions without regard to the processes from physical stimuli and schemas or

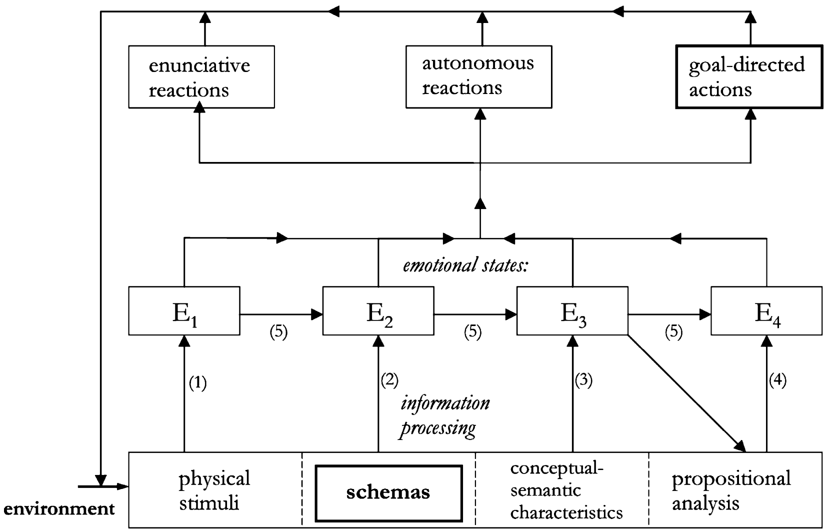

Fig. 8. Genesis of Emotions (Kuhl, 1983).

without regard to the activated (or generated) emotional states. Even in a theoretical framework which can still be represented in a graph like Figure 7, the sequences which have to be included by a hypothesis which really can test parts of a theory will become quite complex. It would take a book chapter of its own to lay out and show the details of the full experimental design for a theory to be tested. The following model is a simplified example which is implicitly part of many theories:

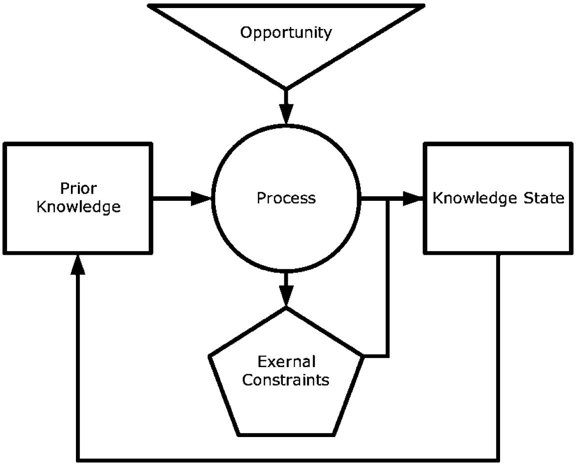

Fig. 9. Simple Example of a Model of Sequence.

Assume that a learner possesses a certain state of prior knowledge (K_n) (Figure 9). Opportunities for learning ($O_{n.m}$) which have environmental constraints (C_n) extend this prior knowledge to a knowledge state (S_n) in the transition (or mapping) function f_q. This state will then lead to a new state of "prior" knowledge (K_{n+1}) by a transition f_r (including, e.g., forgetting, assimilation/accommodation procedures, ongoing experience, evidence, counterevidence):

$$K_{n+1} := f_r(S_n) = f_q(O_{n,m}(K_n, C_n))$$

Under classical circumstances the sequential assumption would be reduced to a control group design, which could be realized as follows:

1. The vector of learning opportunities would be operationalized or modified for operationalization
2. The opportunities would be transferred into treatments which would be left out for the control group

3. The change in knowledge would be assessed in pre- and posttests within a 2x2 design

Using the formalism proposed by Campbell and Stanley (1966, p. 8), we would have a simple experimental matrix like this:

$$\begin{array}{ccccc} O_1 & \rightarrow & X_1 & \rightarrow & O_2 \\ O_1 & \rightarrow & & \rightarrow & O_2 \end{array}$$

Problems with this standard design as regards sequences occur when we apply the dilemmas which led to the idea of the EDE. The most important of them are:

1. The experiment can not test anything but *if* there is change and whether change *differs*. It can not be used to say anything about *why* and especially not *how* something changed. In order to describe the why and how, we would need to observe the change within the process of the treatment.
2. As for experiments within learning and instruction we do not have non-treatments, non-learning, or non-progress. Without a proper theory about the "standard" intervention (which is, for instance in schools, usually not standard at all), we can only compare theory-based design versus non-theoretical designs, which may be systematic or chaotic, based on belief, tradition, constraints, etc. In contrast to controlled psychological experiments which elicit single stimuli and responses for non- or only part-embedded effects and where confounding variables can be controlled for the sake of internal validity, the approach does not suffice to describe the whole process of learning. This is, of course, not a matter of esotericism: I still stand by the fact that these processes can be investigated in experiments. The experiments will only take some more effort in terms of methodology and implementation.
3. The third problem is that we almost automatically think of a between-group design. But when we think about learners in progress and in using this progress to test our theories, we may want to know more about the distance covered by the learners – and this is not only applicable as a contribution to learning theory. In the end we want to find out new insights which help us to develop tools and treatments, which in turn allow learners to progress faster or more easily or to reach goals which they could not reach before. Without monitoring the progress within a group, we are just not able to tell how our assumptions work in the application – no matter whether we are investigating real applications, laboratory experiments, or something in between.

If you consider the presumptions for the EDE there are some more problems with the experimental design which could be addressed. I chose the above list because it seems to summarize the most important traps. However, the problems for the experimental design are not too hard to solve.

For the *sequence prognosis* of this simple example, it is only necessary to make a methodological change which, although minor, can bear fruit for the consequences of the inferences. This begins with the search for instruments and a methodology which can track specific sequences over time to answer the following question: To what extent can the specific sequences which the theory assumes be observed within the design? For all possible tests of knowledge which gain in performance or transfer tests of operationalized learning and instruction goals, the model shows how the process changes depending on the "circumstances" (context, situation). Thus, although there are process attributes which are not influenced effectively by the circumstances, there are also others which change according to the circumstances. At this point the theory must show which way special causes of previous knowledge and occasions can go. This enables one to use the systematic change in the process and circumstance sequence as a general conditional variation to test each and every derived assumption. The hypotheses are specified to the degree of their resolution, and this specification can be refined further through quantitative specification or directedness depending on one's objectives. Whereas the appearance of two attributes together only affords insufficient insight into the theoretically predicted causal relations, the question as to the sequence in which they appear allows an additional interpretational step: The reaction to an observable action which was introduced as a stimulus will, provided that the action is also independent from the reaction (in all observations which include residue), strengthen the hypothesis concerning the causality within the model. There are still configurations in this case, primarily in the identification of suppressors in complex models, which can neutralize this causality. However, the modified assumption can always be adjusted by way of observable implication over simple co-occurrences. A random sample with too many or too few sequences implies for the examination of the data that the sample is not really random and that the sequences could be attributed to a systematic effect. The present example is, however, not comparable with the concerns of modern research on learning and instruction. It already shows deficits in its lack of a sound theoretical embedding (especially as far as the illustration model is concerned). It is, however, suitable for clarifying the advantages of prognoses for structural support. Sequence assumptions of recent theories on learning and instruction are generally more complex than older ones and possess many more descriptive levels. Even so, these descriptive levels are often connected causally with one another in the theories.

Hypotheses on Frequency

It is also possible to consider simple assumptions about frequency in this context. The influence of the frequency of repetition of learning tasks on the immediate memory, on memory persistence (against fading), or on operationable transfer performance (cf. Mandl & Reinmann-Rothmeier, 2000) has been investigated for transfers from descriptive geometry on special skills (cf. Gittler & Glück, 1998). These processes can be visualized using the following highly simplified model.

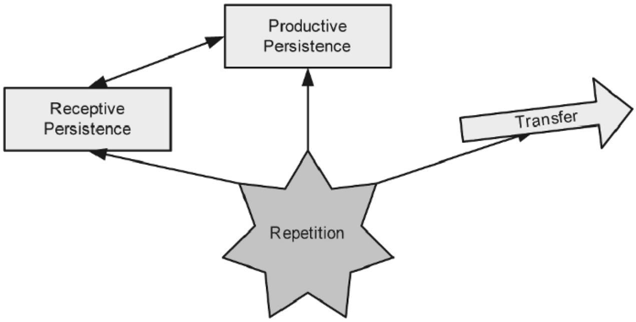

Fig. 10. Simple Example of a Model of Frequency.

Our example can also be used to illustrate a corresponding change for the *frequency prognosis*. The example shown in Figure 10 will be used to investigate the influence of the repetition frequencies on retention performance, productive and receptive persistence, and transfer performance. For classical experimental research, the same problems arise as for the elicitation of sequences. But over and above this, the use of discrete measurement points would not help either. Nominally divided group comparisons aided by variance-analytical methods of repeated content or correlation tables must be replaced by models of multiple regressions, in this case especially by the non-parametric testing method for polytomic ordinally scaled variables or for contingency tables with nominally scaled data or complete path analyses. Complete path analyses are especially interesting when the model structure is already pre-defined by the theoretical background and no special inconsistencies are expected in it. As with the simplified example within the sequence prognosis, the present example is a radical simplification of questions typically met with in research. Instruments for the observation of frequency are in general easy to construct. As long as the observed elementary part of the assumptions can be operationalized to be observed, the rest is tracking and counting. In the case of concrete research practice, the frequency prognoses are interesting especially when the theory not only predicts the presence of a special variable but also the frequency with which it takes effect, i.e., the time-dependent impulse which cannot be judged to be constant for all variables.

Hypotheses on Change

Hypotheses of change have a high potential for clarifying and testing many existing theoretical constructs. In most cases they are the consequent synthesis of sequence and frequency. This is what makes them complex on the theoretical side. On the methodological side they are complex because of the repeated errors and the high demand on instruments to be fast and as non-reactive as possible to trace the path of the learner without exerting too much influence on the learning process. When we are talking only about the application this may be negligible because we may want even the assessment intervention to have a positive influence on learning. If we want to test theories, however, we have to reduce the influence of an instrument to a minimum, which is especially hard to achieve when we need to have rapidly repeated measurements. However, theories about learning are very often theories of change, which simply results from the fact that learning is *by its nature* a specific kind of change. This is why hypotheses about change and Extended Design Experiments derived from them to rebuild the structure empirically may turn out to be the gold standard for theories of learning and instruction. Research practice still has to cover up on the (already existing) methodological standards of the measurement of change. While the theories themselves very often make assumptions about *how* things change, e.g., within a learner or within a learning environment, their empirical testing (especially within experiments) is only about *if* something changes (simple pre-post tests, some of them with a control group). Fortunately, this volume contains a chapter written by my colleague Dirk Ifenthaler (see Chapter 3) on methodological aspects of the measurement and application of learning dependent change and how experiments can be set up accordingly.

On the Danger of Arbitrary Methods

The question which could present itself in this matter is that of deciding which research instruments are appropriate for the operationalization assumptions one has made. In the long term, this requirement can lead in extreme cases to a situation in which every theorem and every problem connected with the theorems must be accompanied by a special methodology which can no longer be standardized, and this can lead to arbitrariness in the observation standards as well as to a lack of communicability. If the resolutions of one's research interests reach this level due to a high degree of specification, it will doubtlessly be necessary to reconsider how the research methods and results can be made communicable. From the perspective of current theories, the requirements can only result in a specification of the tests at appropriate points.

Controlled Laboratory Conditions

Whereas contemporary theories are becoming more and more complex as far as the phenomena they attempt to describe are concerned, leading to a situation in which it may be necessary to determine whether the infrastructure and organizational profiles of schools need to be changed radically to enable new, elaborate, theory-driven applications to be installed properly, the scientific questions connected with this matter remain open. From time to time there will be applications which cannot be realized in accompanying research due to the high amount of effort involved in "re-orchestrating" the school environment, or even because this endeavor seems impossible within the framework of the research project. Not all theory-based designs lead necessarily to only minor changes to the application in the classroom. If, for instance, a theory-based design which also has a good practical rationale would make major adjustments to the organizational procedures of a school (e.g., requiring a major rescheduling because of a change of learning times), we would certainly need evaluation studies and studies which show that the design is already feasible. Thus, the "orchestrability" which is needed to implement a Design Experiment has certain limits on both ends, for practice and for research. Once this point has been reached from only one side, it may be best to refrain from attempting the direct practical realization of theory-based application demanded by Brown (1992) and instead bring the design experiment back to the experimental laboratory. By no means must this lead to a radical simplification of the instrument or a drastic simplification of the experimental environment. A laboratory set up with the needs of the design experiment in mind can – provided that the theoretical problem requires it and the available means allow it – a complete teaching and learning environment which provides the necessary conditions for professionally designed instruction. Moreover, even teachers can, in addition to students and pupils of all ages, be counted among the test subjects of an experiment on location. In addition, the laboratory offers the dramatic advantage of enabling the permanent installation of any observation technology, room changes, or complete changes in time and infrastructure required by the theory one is testing. A well-equipped laboratory can, for example, use networks and bandwidths which cannot yet be installed in classrooms, and this equipment can also be used outside of the experiments by various target groups. Longitudinal studies in particular can be realized much more affordably and with less administrative efforts. Flexibility in room design and in the addition and removal of technical equipment is advisable for testing many different theories on learning and instruction. Basic equipment for presenting content in all central forms of perception is obligatory for the testing of the theories. Designed in this way, the laboratory can become a designed reality of instructional design in the truest sense. The only validity risk which needs to be taken into account is the novelty effect, especially when learners who are used to simple instructional conditions (such as pupils) are confronted for the first time with a learning environment outfitted for sophisticated theoretical purposes. This potential risk must be met with additional control mechanisms.

Conclusions

Enhanced Design Experiments (EDE) can help us to empirically trace complex processes in applications back to the theory used to design the applications. Thus, they allow us to develop and test more complex theories of learning and instruction. Conversely, they can also help us on the methodological side to keep up with our theories, since they of course develop to become more and more complex and specialized over time. We test them by carefully observing how well our theory predicts processes within the application. The difference from classical experiments is that we now not only test selected effects, but also whole models of a theory. Years of methodological development for experimental and field research have given us the tools to statistically test and reconstruct the models.

Common Processes of Research on Learning and Instruction

From the perspective of learning and instruction, we can trace the path of research as follows (see Figure 11): In the referring fields (e.g. cognitive science, computer science, psychology, etc.) components of learning are investigated mainly by classic experimental research. They have their own theories and research processes. Some but not all of these insights can be transferred and generalized into theories of learning and instruction, which are theories about environments of learning and about complete learning processes. Therefore, they have more complexity in their assumptions than theories from referring fields and they are on a generalized level. With these theories at hand, we can design instances of operationalizations which are applications or, more generally, technologies. This second theory-guided transfer is an integral part of instructional design. A good understanding of the referring fields helps to understand how theories of learning and instruction have developed and why they change. This leads to a simple working model of the whole process which can help to locate interests of research and design:

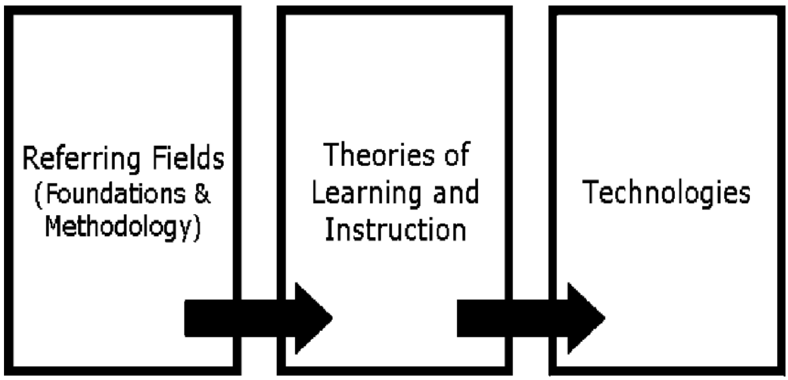

Fig. 11. Working Model of Research and Design.

The transitions are formed by deduction. One or more referring theories can imply a theory of learning and instruction. As discussed at the beginning, design is the second transfer and it is also done by deductive implications. The transfers need to be done separately. A transfer from the referring field directly into technology is not advisable for most cases: First, the external validity of the assumptions has to be tested, at least to make sure that their effects are not too small to play a role in a real learning situation or that they are too specific and thus appear too rarely. This can be tested by design experiments. Second, previously controlled variables may have moderating effects and can even have a negative effect on the learning process. In addition, the operationalization may have too little of an impact on the application. This can be tested by enhanced design experiments and hypotheses which postulate structure from the theory.

Instances of Learning and Alternative Effect Metaphors

Technologies and their applications should facilitate learning processes. Since they are deductions from operationalizations of theories, they are empirically based on tested effects. We now have quite a number of different learning theories which are all being used to build good learning technologies. Most theories of learning and instruction are not discussed antithetically, but rather describe different beneficial effects for different learning situations. We can see from many examples that the use of these effects generally helps to make things better for learners and for teachers, trainers and instructors. We trust an effect as soon as it helps the many. Some are more affected, others less, but all in all they are normally distributed most of the time. We see that what helps the many has to be feasible when we are teaching more than one person. We may still have the effect metaphor of an engine in mind, implicitly or explicitly. However, this is an axiom which is not discussed very much. As odds and endings of my article I would like to share some weird ideas on what may change in the whole model of learning and instruction if we use only very slightly different effect metaphors on the same domain of theories.

Extending the Engine Metaphor

I start with the learner. Most effects from treatments are not equally distributed within a group of learners. Some learners may benefit more from treatments than others. The individual benefit can depend on situational, personal, organizational, and many other factors. Using the engine metaphor from experimental research, we think of benefits from treatments. Based on this metaphor, we may assume that every learner has benefits from the treatments although they may vary. Everybody will be pushed or pulled (or encouraged to walk) from his or her starting point towards a given goal. When we think about learning and instruction we could extend this metaphor. We might ask: What does it take for the learner to "dock" to the engine? Since every learner has a different trait, this is no *on and off* question. The data on most effects are normally distributed. We could assume that every learner has a multiple compatibility vector to the treatment (the engine). The more elements match the better the effect on the learner is (the more she or he is "accelerated"). So far we need the learners to have a certain configuration to fit into the application based on the effects. If they do not have this configuration, we must first help bring them to a state they can start from. Or we have to search for a different theoretical basis which has different effects and different demands. This is what "learner orientation" is mainly about.

The Barrier and Rift Metaphor

The engine metaphor is still a metaphor and up to this point I have treated it like an axiom. Decisions about learning and instruction, Instructional Design, generate theorems (which are applications) on its basis, such as "model centered learning and instruction" (Seel, 2003). But as all models and metaphors have limits, and finding them out is the hard part, the engine metaphor has limits, too. We know that learning occurs naturally and without intent. In other words, learning can not be stopped. We use education and instruction and all of the associated techniques of coursework, assignment, teaching, and assessment to direct the learning towards certain goals, but then we find out that the learner is not always able or willing to follow this path – maybe other issues have more direct rewards, or maybe they are easier to grasp or more likely to raise curiosity. Thus, we design, develop, and implement environments on the basis of good research to help them concentrate on the given goals. But what if all of the environments, schools, classrooms, and instructional measures are more like barriers first than engines? And within this all of our efforts create rifts in the existing barriers? This metaphor could still explain every distribution, every effect, and every piece of data on control group designs, and it is not yet too different from the engine metaphor, except for the relation between the environment and the learner: The conditions for successful learning processes change completely. Within the range of the engine metaphor, the learner is the one who has to be compatible with the environment. We can change the basis for design but then other learners will be at the extremes of the expected distribution. Within the barrier and rifts metaphor it is the environment which has to fit the group of learners, as long as we have two or more learners. Up to now I have not said anything about how this could be achieved. After introducing the idea of learning instances, I will come back to the design of rifts, which differs from the design for engines in that it introduces the university metaphor. Before that I want to introduce some basic ideas which show a different perspective on the person we usually call "the learner."

Realizations

Everything that can be observed is one of many possible realizations consisting of causal influences, probabilistic interrelation, functional matrices, and maybe chaos to some unknown extent. While the boundaries between order and chaos may be investigated, chaos itself can by definition not be described by functions of order. Most realizations are not independent of other realizations. Those which depend on other realizations are special cases: Their dependency may be described. Maybe all realizations can be described. But even then, not all of them have yet been described. This is an inductive (or maybe abductive) conclusion. The conclusion is not independent of realizations. Deductive reasoning is *by design* independent of realizations. However, the results of the conclusions are not independent of realizations. Realizations can be influenced by the results (or the perception of the results) of deductive reasoning.

The knowledge about all of this is a set of realizations. This text is a realization. This text is a realization which is influenced by knowledge. We may not know the extent of this influence and which dependence this influence has within this (multi-dimensional) realization. I'm used to calling such realizations based on realizations of knowledge "language". On the other hand, "language" is also nothing other than realizations of realizations – containing all of the other ideas from set theory (see Bach, 1989; Thomason, 1974). The nesting allows us to describe sections of complex realizations (which is very useful). A section is a realization of a map of realizations on realizations. A map is a realization which is implemented between at least two realizations and which can interact with all other realizations involved. A section can have a simple, complex, concrete, or abstract impression. Realizations with which my senses (which are also realizations on realizations) can build direct mappings are typically called "concrete", while all other realizations are called "abstract". Because I can only reconstruct abstract realizations indirectly in mappings, I call them "constructs". These differ from all other realizations only in the way my senses interact with them. Realizations which can not hide the fact that they are integrated with many of my interrelated realizations are called "complex". Realizations which seem to be one or integrate only few realizations are called "simple". Simple and complex realizations differ only in the amount of single and interacting sub-realizations they present to my senses, which can be called the "surface of a realization". Realizations which do have significant sections which change differently over time can be described as being "dynamic". Whether realizations are seen as "concrete" or "abstract", as "complex", "simple", or "dynamic" depends on the level of integration of the realizations with which they interact. In my case this integration is very likely not constant even within short time spans. Something which is complex in one situation may be simple in another (e.g., "money"). Something which is abstract in one situation may be concrete in another (e.g., "ultrasound"). Systematic differences between systematically integrated realizations must be expected. We have to take this into consideration if we want to create environments which, like in our case, allow individuals to learn.

In a specific learning situation (time point), the set of realizations of realizations which is the learner can be described as an instance of himself. This instance differs from other instances at different time points. It should be clear by now that a good part of this difference can not easily or even statically be tracked. If tracking is not easy already, then prediction is even harder. The learner may have different instances at each point in time which are instances of learning.

Instances of Learning

The learner can be described as an entity of multiple and variable, potentially dynamic and complex realizations. Given a point in time, the learner is an instance of these realizations. Some of the realizations may be known, others may induce some (valid and false) clues on a well trained and emphatic teacher or instructor. Most of them will remain unknown for the time point due to their dynamic nature. The known attributes can lead to systematic instruction. The clues may lead to more or less working heuristics. The unknown parts will lead either to some chaos or to nothing. Unless we have a complete model of this specific instance of the learner, we may consider that there is more information about the realization which we do not know than information which we do know, regardless of how far experimental research has progressed in the area of common realizations. The learning instance is at least multivariate, if not something like “omnivariate”. Therefore, if we choose to follow the barrier and rifts metaphor, the rationale for design changes also. This decision leads to a model which I like to call the university metaphor.

A Quantum of Learning – Towards the University Metaphor

Each learning instance carries an opportunity to learn for the learner. Whether this opportunity can be taken by the learner depends on the availability of a compatible rift. The progress which a learner can make using a single opportunity can be described as a realization of a quantum of learning. Differences exist in learning potential and actual progress. If we have only a single learner the rift may be opened individually, given that we know a significant part of the learner’s learning instance. If we have two or more learners we might have to open multiple rifts in the barrier, given that not all learners in a group are alike as far as the properties of their learning instances are concerned. To do so we still would have to know a lot about each single instance. With larger groups at hand we may come to the conclusion that single treatments and interventions are not likely to sufficiently supply all learners with opportunities. Maybe a very old but still (somehow) applied metaphor can help to open the right rifts. University means “everything” – maybe not by accident. If the goal is to have a rift at hand for each individual learner, why not offer a diversity of opportunities to a group based on different designs and even on a set of theories? Even if an individual learner changes the properties within his or her instances, there may be another offer (open rift) which fits. Of course the navigation of the offers would have to be conducted by the learner and thus he or she must be trained for this complex task, but as soon as the learner performs well after a time, he or she may be able to conduct this task, given a specific learning goal. With current research at hand, we have a lot of opportunities to offer to learners. To step into applications based on the barriers and rifts metaphor and on the university metaphor we need basic experimental research on how learners handle and navigate firstly multiple content sources and secondly multiple learning strategies on their own. The first research can, for instance, be based on text and hypertext research (see Eigler, 1998). The second may be founded on research on metacognition (see Nelson & Narens, 1994). Over and above this, both strands would have to be synthesized to form a comprehensive model about learning navigation showing the interrelation between open rifts (different and available designs and representations) and a single quantum of learning. For model based learning and instruction, the research tasks which I have tried to sketch out here fit very well. If we have the right navigational mental models and the tools at hand to (self-) assess those models, we could gain great insight into the processes of the university based approach, which of course would not only fit with learning and instruction at universities.

Of course, these assumptions are only based on interpretations of the different effect metaphors. They are neither right nor wrong. But maybe they can help one way or another to construct better applications on the basis of the good theories we have at our disposal.

References

Al-Diban, S. (2002). *Diagnose mentaler Modelle*. Hamburg: Kovac.

Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Archer, R., Brockett, A.T., McDermott, P.L., Warwick, W., & Christ, R.E. (2006). *A simulation-based tool to train rapid decision-making skills for the digital battlefield*. Research Report 1859. Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Ayer, A. J. (1959). *Logical Positivism*. Glencoe, IL: Free Press.

Curiel, J.M., Anhalt, M.D., Tran, J.J., & Yao, K.T. (2005). Developing situation awareness metrics in a synthetic battlespace environment. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC).*

Bach, E. (1989). *Informal lectures on formal semantics*. New York, NY: State University of New York Press.

Blumschein, P. (2003). *Eine Metaanalyse zur Effektivität multimedialen Lernens am Beispiel der Anchored Instruction*. Freiburg: FreiDok.

Bracht, G., & Glass, G. (1975). Die externe Validität von Experimenten. In R. Schwarzer and K. Steinhagen, *Adaptiver Unterricht. Zur Wechselwirkung von Schülermerkmalen und Unterrichtsmethoden* (pp. 64-93). München: Kösel.

Braine, M., & O'Brein, D. (1998). *Mental Logic*. Mahwah, NJ: Erlbaum.

Breslow, N., & Clayton, D. (1993). Approximate inference in generalized linear mixed models, *Journal of the American Statistical Association, 88,* 925.

Brown, A. L. (1992). Design Experiments. Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings, *The Journal of the Learning Sciences, 2*(2), 141-178.

Campbell, D.T., & Stanley, J.C. (1966). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.

Carnap, R. (1950). *The logical foundations of probability*. Chicago: University of Chicago Press.

Carnap, R., & Stegmüller, W. (1958). *Induktive Logik und Wahrscheinlichkeit*. Wien: Springer.

Cooper, R., Fox, J., Farringdon, J., & Shallice, T. (1996). A systematic methodology for cognitive modelling, *Artificial Intelligence, 85,* 3-44.

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

Dummer, P., & Ifenthaler, D. (2005). Planning and assessing navigation in model-centered learning environments. Why learners often do not follow the path laid out for them. In G. Chiazzese, M. Allegra, A. Chifari & S. Ottaviano (Eds.), *Methods and technologies for learning* (pp. 327-334). Southampton, Boston: WIT Press.

Ebbinghaus, H. (1885). *Über das Gedächtnis*. Leipzig: Duckner und Humbold.

Ernest, P. (1996). Varieties of constructivism: A framework for Comparison. In L.P. Steffe, P. Nesher, P. Cobb, G.A. Goldin, and B. Greer (Eds.), *Theories of mathematical learning* (pp. 335-350). Mahwah, NJ: Lawrence Erlbaum.

Eigler, G. (1998). Textkommunikation als Wechselspiel von Textverarbeiten und Textproduzieren. *Unterrichtswissenschaft*, *26*(1), 51-66.

Einsiedler, W. (1997). Research on Instructional Methods. A European Perspective. In R. Tennyson, F. Schott, N.M. Seel, & S. Dijkstra, *Instructional Design. International Perspective, Volume 1. Theory, Research and Models* (pp. 269-291). Mahwah, NJ: Lawrence Erlbaum.

Endsley, M.R., & Smolensky, M.W. (1998). Situation awareness in air traffic control: The picture. In M. Smolensky & E. Stein (Eds.), *Human factors in air traffic control* (pp. 115-154). New York, NY: Academic Press.

Gibbons, A.S. (2001). Model-centered instruction. *Journal of Structural Learning and Intelligent Systems, 14,* 511-540.

Gittler, G., & Glück, J. (1998). Differential transfer of learning. Effects of instruction in descriptive geometry on spatial test performance. *Journal for Geometry and Graphics,* 71-84.

Goodman, N. (1956). *Fact, fiction and forecast*. Cambridge, MA: Harvard University Press.

Halpern, J.Y., & Pearl, J. (2001). Causes and explanations: A structural-model approach Part I: Causes. In *Proceedings of the 16th International Joint Conference on Artificial Intelligence* (pp. 27–34), Seattle.

Holzkamp, K. (1994). Am Problem vorbei. Zusammenhangsblindheit in der Variablenpsychologie. In K. Holzkamp (Ed.), *Problemblindheit in der Psychologie und Perspektiven ihrer Uberwindung*. Hamburg: Argument.

Holzkamp, K. (1996). *Texte aus dem Nachlass*. Hamburg: Argument.

Hanke, U. (2006). *Externale Modellbildung als Hilfe bei der Informationsverarbeitung und beim Lernen*. Freiburg: FreiDok.

Ifenthaler, D. (2006). *Diagnose lernabhängiger Veränderung mentaler Modelle. Entwicklung der SMD-Technologie als methodologisches Verfahren zur relationalen, strukturellen und semantischen Analyse individueller Modellkonstruktionen.* Freiburg: FreiDok.

Ifenthaler, D., & Seel, N.M. (2005). The measurement of change: Learning-dependent progression of mental models. *Technology, Instruction, Cognition and Learning, 2*(4), 317-336.

Kuipers, T.A.F. (2007). *General philosophy of science. Focal issues*. Amsterdam: Elsevier.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos, A. Musgrave (Eds.), *Criticism and the Growth of Knowledge,* Cambridge: Cambridge University Press.

Leuchter, S., & Jürgensohn, T. (2001). Situation awareness-training für Fluglotsenschüler. In H. Oberquelle, R. Oppermann, & J. Krause (Eds.), *Mensch & Computer 2001: 1. Fachübergreifende Konferenz* (pp. 437-438). Stuttgart: B.G. Teubner.

Miller, G. A. (1956). The magical number seven, plus minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Johnson-Laird, P. (1983). Mental models. toward a cognitive science of language, inference and language. Cambridge: Cambridge University Press.

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

Nelson, T.O., & Narens, L. (1994): Why investigate metacognition. In J. Metcalfe, & A.P. Shimamura (Eds.), *Metacognition. Knowing about knowing* (pp. 1-25). Cambridge, MA: Cambridge University Press.

Kuhn, T.S. (1976). *Die Struktur wissenschaftlicher Revolutionen*. Frankfurt: Suhrkamp.

Kuhl, J. (1983). Emotion, Kognition und Motivation. Auf dem Wege zu einer systemtheoretischen Betrachtung der Emotionsgenese. *Sprache und Kognition*, *2,* 1-27.

Long, J.S. (1983). Covariance structure models. An introduction to LISREL. Newbury Park, Ca.: Sage.

Opwis, K. (1985). *Mentale Modelle dynamischer Systeme. Analyse und Weiterführung von psychologischen Experimenten zum Umgang von Personen mit Systemen*, Freiburg: Forschungsberichte des Psychologischen Instituts der Albert-Ludwigs-Universität Freiburg i. Br.

Pearl, J. (1995). Causal diagrams for empirical research, *Biometrika*, *82*(4), 669-710.

Pearl, J. (2000). *Causality, models, reasoning, and inference*, Cambridge, MA: Cambridge University Press.

Pirnay-Dummer (2006). *Expertise und Modellbildung – MITOCAR*. Freiburg: FreiDok.

Popper, K. (1934). *Logik der Forschung*, Tübingen: Mohr.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

Schoenfeld-Tacher, R., Jones, L.L., & Persichitte, K.A. (2001). Differential effects of a multimedia goal-based scenario to teach introductory biochemistry – Who benefits most? *Journal of Science*, *Education and Technology*, *10*(4), 305-317.

Seel, N.M. (1991). *Weltwissen und mentale Modelle*. Göttingen: Hogrefe.

Seel, N.M. (2003). Model-centered learning and instruction. *Technology, Instruction, Cognition and Learning, 1*(1), 59-85.

Seel, N.M., Al-Diban, S., & Blumschein, P. (2000). Mental models and instructional planning. In J. Spector & T. Anderson (Eds.), *Integrated and holistic perspectives on learning, instruction and technology: Understanding complexity* (pp. 129-158). Dordrecht: Kluwer Academic Publishers.

Spector, J.M. (1995). Using system dynamics to model courseware development: the project dynamics of complex problem-solving. In K.M. George, J. Carroll, & D. Oppenheim (Eds.), *Proceedings of the 1995 ACM Symposium on Applied Computing, Nashville, Tennessee, United States, February 26-28, 1995* (pp. 32-35). New York, NY: ACM Press.

Spector, J.M., Dennen, V.P., & Koszalka, T.A. (2005). Causal maps, mental models and assessing acquisition of expertise. *Technology, Instruction, Cognition and Learning, 3*(1-2).

Thomason, R.H.E. (1974). *Formal philosophy: Selected papers of Richard Montague*. New Haven: Yale University Press.

Wainer, H. (1989). Eelworms, bullet holes, and Geraldine Ferraro. Some problems with statistical adjustment and some solutions. *Journal of Educational Statistics, 14*, 121-140.

7. Externalizing Mental Models with Mindtools

David Jonassen and Young Hoan Cho

University of Missouri

Abstract: Mental models are complex and multi-faceted, so they cannot be adequately represented using any single form of assessment. After reviewing traditional methods for manifesting and representing mental models, we describe how Mindtools can be used by learners to externalize their mental models using different tools that represent different kinds of knowledge.

Keywords: Mental models; knowledge representation; cognitive tools; mindtools; procedural knowledge; structural knowledge; episodic knowledge.

Mental Models

Mental models are the internal constructs that humans construct to represent phenomena in the world. Through interaction with the environment, with others, and with the artifacts of technology, people construct mental models of the world with which they interact (Norman, 1983, p. 7). Based on the mental models, people describe why a system exists and what a system looks like, explain how a system operates and what a current system state is, and predict what a future system state is (Rouse & Morris, 1986). For instance, when driving a car, people construct and manipulate their mental models of how a car operates, what traffic laws are, where they are on a map, and so on. These mental models are different depending on the way people have interacted with the system. Because of a different traffic system, English drivers have a mental model for driving on the left side of the road, whereas the U.S. drivers are uncomfortable applying that model.

Mental models have been researched by cognitive scientists, psychologists, human-computer interaction specialists, and educators, each constructing their own interpretations of mental models. Johnson-Laird (1983), a psychologist, describes mental models as structural analogues of the world as perceived and conceptualized, which enable people to make inferences and predictions, to understand phenomena, to decide and control actions, and to experience events by proxy. Gentner and Stevens (1983) state that mental models are concerned with human knowledge of the world and of how it works. Wilson and Rutherford (1989) reflect a human factors orientation by conceiving of a mental model as a representation formed by a user of a system on the basis of previous experience and current observation. Mental models provide most of the subsequent system understanding and dictate the level of task performance. Vosniadou and Brewer (1994) define mental models as the kinds of mental representations individuals construct when they reason about the physical world. They also assume that mental models are dynamic structures usually created on the spot to meet the demands of specific problem-solving situations. From these explanations, we conclude that a mental model is a representation or a structural analogue of the world that is constructed and manipulated by a person, and it is the basis of cognitive activities such as understanding, predicting, and reasoning which are necessary for specific task performance.

Norman (1983) distinguishes users' mental models from conceptual models of teachers, designers, scientists, and engineers, and he argues that mental models are incomplete, limited, unstable, confused, unscientific, and parsimonious. For instance, Norman (1983) found that people tended to excessively push the clear button of a calculator when they wanted to restart it, whereas they exhibited reluctance to use the clear button during problem solving for the fear of clearing too much. This shows that doubts and superstitions govern users' behavior and enforce extra caution when they operate a machine.

Assessing Mental Models

Just as theorists have differed in their conceptions of mental models, researchers have also differed in the methods they have used to assess mental models. There is no single agreed-upon measurement tool for mental models. Rowe and Cooke (1995) compared several mental model measures in terms of their correlation with troubleshooting performance. They found that laddering interviews, relatedness ratings, and diagramming techniques were predictive of troubleshooting performance, but the think-aloud measure had low correlation with the performance. However, think-aloud protocols are effective in identifying the sequence of states people progress through (Chi, 2006) and in assessing concepts-in-use because participants think aloud while they solve a problem (Jonassen, 2006). Different measurement techniques focus on different aspects of mental models (Royer, Cisero, & Carlo, 1993), so it is hard to assess all aspects of mental models by a single method. Mental models have been measured extensively by five methods: problem solving, verbal report, drawing, categorization, and conceptual pattern representation.

First, problem-solving performance manifests the features of mental models. Mental models are the basis of understanding phenomena and making inferences and predictions (Johnson-Laird, 1983; Norman, 1983), so if people have different mental models, they will understand a problem differently and create different solutions. Gentner and Gentner (1983) showed that the patterns of solving electrical circuit problems were different depending on the mental model the subject had. Based on this assumption, researchers inferred the characteristics of mental models from problem solving outcomes. McCloskey, Caramazza, and Green (1980) asked students to predict the path a metal ball would follow after it came out of a curved tube in order to assess their mental models of physical motion. Thirty six percent of the pathways drawn were curved lines rather than straight lines. They inferred that people who predicted curved pathways had different mental models of physical motion from those who predicted straight pathways. In addition, Azzarello and Wood (2006) recommended unfolding case studies for assessing situational mental models that develop while a person is actively engaged in solving a problem in a specific situation. Unfolding case studies present scenario data in stages, so students' mental model of the case may change with additional information at each stage. The examination of task performance after each stage can reveal how mental models of the situation are evolving. Although problem-solving performance produces only indirect information about mental models, it provides objective evidences and it can be used with other methods effectively.

A second method for assessing mental models, verbal reports, is a direct method for eliciting mental models. Assessment of mental models depends on what people say about their mental models and verbal reports can be done as interviews, explanations or think-aloud protocols (Chi, 2006). This method is based on the notion that individuals had privileged access to their mental models and their report can reveal their cognitive process and thoughts (Ericsson & Simon, 1984). Southerland, Smith, and Cummins (2000) suggested structured interviews as a method of investigating students' conceptual frameworks. Structured interviews have the advantage of allowing students to express what they know and how they can apply the knowledge in their own words. The use of generative questions that require construction and manipulation of mental models are effective in measuring mental models. For example, Vosniadou and Brewer (1992) used such generative questions as "If you were to walk for many days in a straight line, where would you end?" in order to examine children's mental models of the earth. In addition, explanation of observed phenomena is one of the methods measuring mental models (Sabers, Cushing, & Berliner, 1991).

Another form of verbal reports is the think-aloud method (Ericsson & Simon, 1984). For think-aloud, subjects are asked to simply verbalize their thoughts they attend to while performing a task rather than describe or explain what they are doing. Think-aloud protocols have been used for assessing the difference of mental models between experts and novices and the processes in which mental models are constructed and developed for problem solving (Anzai & Yokoyama, 1984; Hong & O'Neil, 1992; Simon & Simon, 1978). Think-aloud protocols are useful data for analyzing mental models because they provide direct information about ongoing thinking processes rather than the outcome of thinking.

Third, drawings can be a complementary method of verbal reports because verbalization of a nonverbal image leads to a biased model. Whitfield and Jackson (1982) found that air traffic controllers had difficulty in verbalizing the image of the system's states and Rouse and Morris (1986) argue that verbal reports have limitation because mental models are frequently pictorial. For this reason, drawings have been used with verbal reports in several mental model research studies. For example, Butcher (2006) asked participants to draw a picture of what they know about the heart and circulatory system and to explain their drawings before and after learning. He categorized drawings according to the mental model of the heart and circulator system and compared students' drawings and verbal explanations in order to examine whether the mental model is improved by learning. In addition, Vosniadou and Brewer (1994) asked children to explain the day and night cycle not only by a verbal response but also by making a drawing. They provided the drawing of a person living on the earth and asked children to draw a picture that made the earth day or night for the person. The drawings represented children's different mental models of the day and night cycle. Drawings can provide information about pictorial aspects of mental models, which are difficult to be measured by verbal reports.

Fourth, categorization of instances reveals how mental models are developed and organized. Categorizing problems based on their similarity has been frequently used for identifying the cognitive difference between experts and novices (Chi, Feltovich, & Glaser, 1981; Hardiman, Dufresne, & Mestre, 1989; Schoenfeld & Herrmann, 1982; Silver, 1979). For example, Chi et al. (1981) asked participants to sort physics problems and to explain the reasons for their categorization. Novices sorted the problems based on the surface features such as the presence of blocks and inclined planes, whereas experts tended to sort them according to the major physics principles that were critical to solutions. This result shows that the mental models of novices are different from those of experts because mental models are the basis of perceiving and sorting problems. Moreover, novices who judged problems based on principles tended to categorize problems similarly to experts and solved problems better than other novices who relied on surface features (Hardiman et al., 1989). Thus, categorization can be used for assessing whether mental models are constructed based on principles or surface features of problems.

Finally, mental models have been represented in the form of concept maps. Concept maps spatially represent concepts and their relationships and they have been used extensively to assess learning outcomes (Jonassen, 2000; 2006). Students can easily create concept maps without statistical analysis and they provide extensive information of conceptual patterns. In addition, multidimensional scaling (MDS, Kruskal, 1964) and Pathfinder (Schvaneveldt, 1990) scaling algorithms have been used for visualizing structural knowledge. The outcomes of MDS and Pathfinder can be assessed both qualitatively and quantitatively. Conceptual patterns of MDS are qualitatively assessed by examining the clusters of concepts and the meaning of dimensions, whereas networks of Pathfinder are qualitatively assessed by analyzing the location and links of concepts and hierarchical features. For instance, Jonassen (1987) used MDS to assess student's conception of Newtonian mechanics. Wilson (1994) used both MDS and Pathfinder to examine the variation of knowledge representation about chemical equilibrium and qualitatively compared conceptual patterns between higher achievers and lower achievers. For quantitative analysis, similarity scores between each person and an expert are frequently used. That is, the higher the similarity score of conceptual patterns is, the closer the mental model of an individual is to that of the expert. The similarity of networks has been reported to predict domain performance highly effectively (Goldsmith, Johnson, & Acton, 1991; Gomez, Hadfield, & Housner, 1996).

Mental Models are Multi-Dimensional

In most of the research on mental models, scholars have attempted to define mental models uni-dimensionally, that is, to identify a single descriptor for mental models. We assume that the construct, mental model, is too complex to be described using any single measure or form of assessment. Mental models are more than structural maps of components. They are dynamic constructions that can be manipulated and tried out. They are multimodal as well as multi-dimensional. Mental models are complex and inherently epistemic, that is, they form the basis for expressing how we know what we know. Because mental models are epistemic, they are not readily known to others and, in fact, not necessarily comprehended by the knower. Jonassen and Henning (1999) showed that troubleshooting performance (a manifestation of procedural knowledge) was positively related to a variety of mental model measures, including structural knowledge, as represented by Pathfinder Networks (Schvaneveldt, 1990), a verbal recollection of their visual image of the system, metaphors that students generated about the system, and retrospective debriefings. That is, students who were better troubleshooters produced better structural knowledge, metaphors, and images of the system they were troubleshooting. That is, the larners had constructed more robust mental models. Each measure of each construct was highly related. They concluded that mental models possess multiple forms of representation. Research to identify all of the relevant components in mental models needs to be conducted. It is likely that mental models also possess executive control or strategic knowledge as well as episodic memories. The latter construct will be described later. In this paper, we briefly describe some of these cognitive dimensions of mental models and also suggest computer-based tools for externalizing representations of those cognitive dimensions. These are tools for externalizing mental models.

Limitation

In this paper, we address only cognitive dimensions (mental models in the head) of mental models. While we accept the existence of social mediated or team mental models derived from the intersection of different individuals' mental models, that discussion is beyond the scope of this chapter.

Modeling Mental Models: Alternatives for Facilitating the Construction and Assessment of Mental Models

The premise of this chapter is that externalizing mental models improves the utility, coherence, and cogency of mental models as wel as providing external representations of those mental models. That is, building external models of internal mental models improves the mental models and the learner's understanding of those models and provides evidence about theor coherence and completeness.

The primary purpose of modeling, from our perspective, is the articulation of mental models. Building explicit models externalizes or reifies mental models. These models are separate fromfrom their referent mental models. Perhaps the most important characteristic is the evaluation of competing alternative models, that is, the comparison of two or more models for their relative fit to the world (Lehrer & Schauble, 2003). Which model better reflects the external world? Comparing and evaluating models require understanding that alternative models are possible and that the activity of modeling can be used for testing rival models.

Modeling is fundamental to human cognition and scientific inquiry (Frederiksen, White, & Gutwill, 1999). Modeling helps learners to express and externalize their thinking; visualize and test components of their theories; and make materials more interesting. Models function as epistemic resources (Morrison & Morgan, 1999). We must first understand what we can demonstrate in the model before we can ask questions about the real system.

If we agree that mental models are multi-faceted and multi-modal, then in order to externally represent mental models, we need to employ multiple representational formalisms. Jonassen (2006) describes a variety of computer-based Mindtools for representing domain knowledge, systems, problems, experiences, and thinking processes. Using computer-based tools, such as concept maps, databases, expert systems, spreadsheets, hypermedia, and teachable agents, to construct models fosters mental model development. That is, there are models in the mind (mental models), and there are external models that represent the models in the mind. The relationship between internal and external models is not well understood. We believe that there is a dynamic and reciprocal relationship between internal mental models and the external models that we construct. The mental models that we construct in the mind provide the material for building external models. The external models in turn regulate the internal models that we build, providing the means for conceptual change (Nersessian, 1999). In this paper, we argue that the construction of models using different computer-based modeling tools (Mindtools) enables learners to tune their internal models.

In the reminder of this chapter, we describe how different Mindtools can be used to construct models of different kinds of knowledge that represent some of the facets of mental models.

Representing Structural Knowledge

Many psychologists equate mental models with concept-map-like representations. Concept maps are representations of structural knowledge (Jonassen, Beissner & Yacci, 1993), knowledge of the semantic relationships among the schemas comprising the model. Structural knowledge is also known as a cognitive structure, the pattern of relationships among concepts in memory (Preece, 1976). For example, Pathfinder nets generated from relatedness data were created to depict mental models (Kraiger & Salas, 1993). Carley and Palmquist (1992) use their own software for constructing interlinked concept circles (maps) based upon text analysis or interviews.

Structural knowledge may be modeled with semantically sensitive software such as concept maps and databases. Figure 1 illustrates a structural model of the molar conversion process. Each concept (node) represents a concept, while each of the lines represents a semantic relationship between the concepts (a proposition). The larger these concept maps are, the more useful they are in supporting mental model construction. A student's concept map for any course, we believe, should probably include more than 2,000 nodes that are inter-connected.

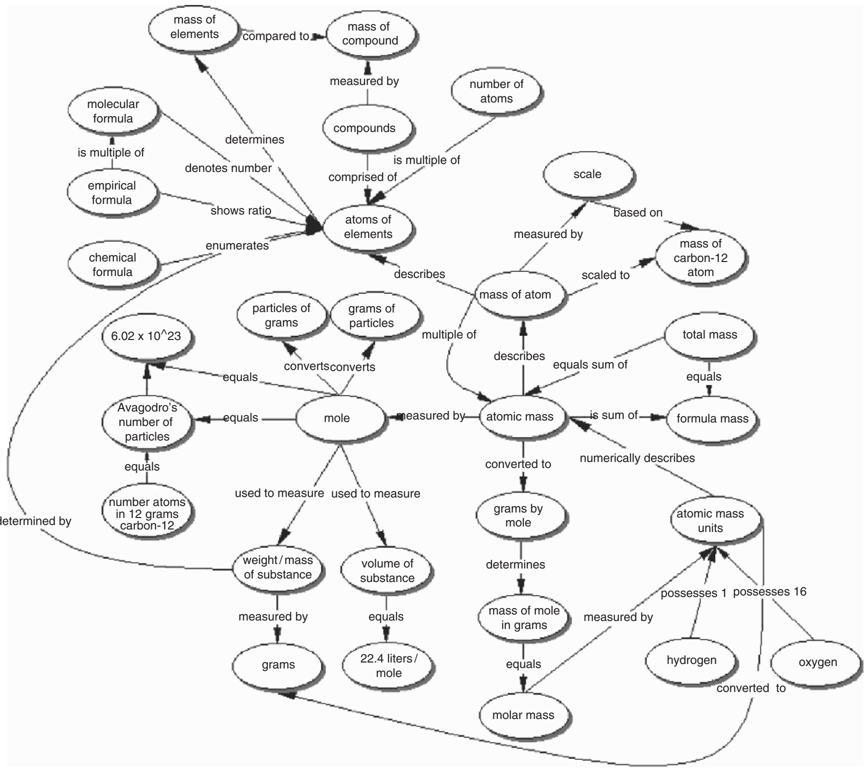

Fig. 1. Concept map representing structural knowledge of molar conversion process.

Databases support more integrated structural models of content. Each cell in a database represents a node. The links are represented by the relationships between records and field. Therefore, databases constrain the kinds of relationships that can be depicted in the model. That characteristic also provides for a more integrated model, with records more tightly yoked to each other. However, both tools enable learners to construct semantic models of the concepts that are integral to a domain or discipline. Too often, teachers attempt to achieve this effect by having students memorize definitions of domain concepts, a process that is much weaker.

Representing Visual Knowledge

Jonassen and Henning (1999) found that mental models also contain a spatial or pictorial representation. As mentioned before, Whitfield and Jackson (1982) and Rouse and Morris (1986) identified pictorial images as an important component of mental models. Images are perhaps the most important dimensional representation of mental models. Wittgenstein (1922) described propositions as imaginal models of reality. Most humans generate mental images of verbal representations. The statement, "The stone gained speed as it rolled down the steep slope" is meaningful only when an image of a mountain with a stone descending along its side is generated. Mental models definitely include mental images of the application of domain knowledge. So, it is important to elicit the learner's mental image of a prototype of the system s/he is constructing.

Generating visual presentations of mental images is very problematic. Mentel images are private and cannot readily be externalized. There are no tools for converting mental images into pixels. Rather, the normal method is to use some sort of graphics program to generate a mental image. That process requires both computer skills and drawing or painting skills. Most of us lack such skills, so our presentations will be inexact models of what we imagine. What humans need are visual prostheses for helping them to visualize ideas and to share those images with others.

Within scientific domains, there are computer-based tools for generating visualizations. In mathematics, tools such as MatLab and Mathematica readily convert mathematical formulas into dynamic visual representations, resulting in better understanding. For example, engineering mechanics students who used Mathematica solved problems requiring calculus more conceptually when compared to traditional students who focused only on the procedures (Roddick, 1995). Being able to interrelate numeric and symbolic representations with their graphical output helps learners to understand mathematics more conceptually.

A number of visualization tools have been developed for the sciences, most especially chemistry. For example, Figure 2 illustrates a molecule of androsterone. Not only does the McSpartan program enable the learners to visualize molecules using five different representations (wire, ball and wire, tube, ball and spoke, and space filling) but it also enables the student to test different bonds and create ions and new molecules. Understanding molecular chemistry is greatly facilitated by visualizing these complex processes. There has been a bit of research on these tools. For example, high school students used eChem to build molecular models and view multiple representations of molecules. Students using the visualization tool were able to generate better mental images of a substance that aided their understanding (Wu, Krajcik, & Soloway, 2001), confirming our belief that there is a reciprocal relationship between mental models and the external models learners construct to represent them.

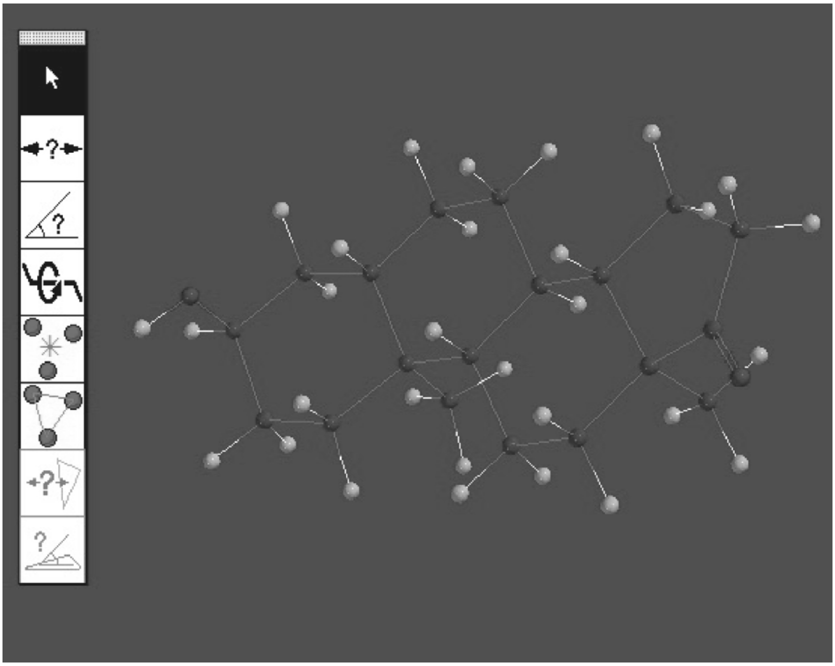

Fig. 2. Visualizing molecules.

Representing Procedural Knowledge

Jonassen and Henning (1999) also showed that the procedural knowledge of effective troubleshooters exceed that of the poorer troubleshooters. Procedural knowledge includes not only a description of the process but also a causal model describing and predicting the performance of the system. The best mental models are runnable, that is, they can be used to model and test how the system functions (Gott, Benett & Gillet, 1988). Assessing procedural knowledge is difficult. Most commonly, researchers ask performers to think aloud while performing a process.

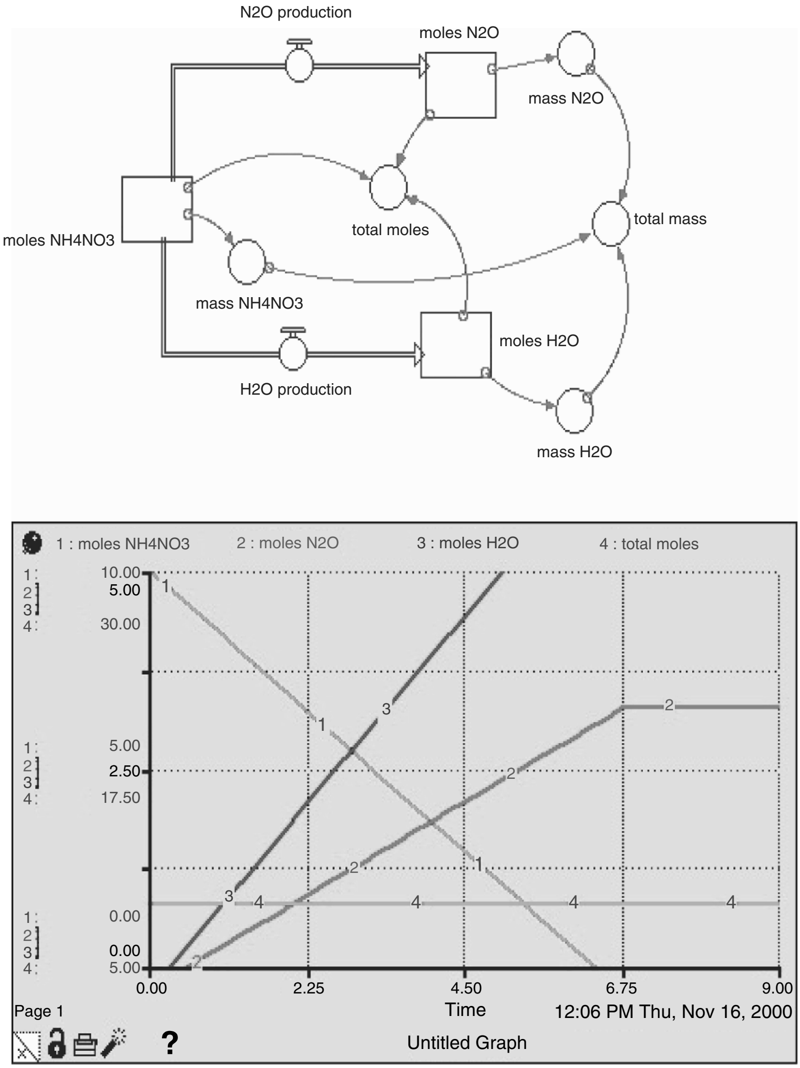

Fig. 3. Stella model of molar conversi on problem

Retrospective debriefing involves asking the performer for explanations of their actions after the performance. These data are difficult to analyze.

A powerful tool for building models of procedural knowledge is the expert system. An expert system is a computer program that attempts to simulate the way human experts solve problems—an artificial decision maker. For example, when you consult an expert (e.g., doctor, lawyer, teacher) about a problem, the expert asks for current information about your condition, searches his or her knowledge base (memory) for existing knowledge to which elements of the current situation can be related, processes the information (thinks), arrives at a decision, and presents a decision or solution. Like a human expert, an expert system is approached by an individual (novice) with a problem. The system queries the individual about the current status of the problem, searches its knowledge base (which contains previously stored expert knowledge) for pertinent facts and rules, processes the information, arrives at a decision, and reports the solution to the user. When used to model procedural knowledge, learners assume the role of an expert and construct a set of IF-THEN rules using an expert system editor. In those rules, they embed the causal and procedural relationships that experts use when making a diagnosis, for instance. The rule base can be tested by running the model and seeing whether the advice that is provided by the rule base, is viable. Building expert systems is technologically easy and intellectually compelling.

Another tool for building external models of procedural knowledge is the systems dynamics tool. The more common kind of tool, the aggregate modeling tool (such as Stella, PowerSim, VenSIm) uses a set of building blocks (stocks, flows, converters, and connectors) to construct a visual model of the components of a system. The systems model in Figure 3 applies the molar conversion process. Students then embed mathematic formulas in the connectors. Students test their models by running them and observing the graphic output in Figure 3. Building systems models is an engaging and powerful process for representing mental models. They probably provide the most complete externalization of mental models that is possible. They can be used to model social, psychological and other processes as well as scientific.

Episodic Knowledge

The strongest kind of memory is episodic. People often remember their experiences with accuracy decades after they occurred. The most common form of external representation of experience is the story. When faced with a problem, the most natural problem-solving process is to first try to recall a similar problem that you have experienced, what you did about it, and how effective that solution was.

Failing that, we tend to communicate with friends or colleagues, tell our problem, and ask if they have experienced a similar one. Frequently they have, and they are usually very willing to share with you a story of their experience along with the lessons learned.

Students can model their own or other people's experiences by collecting stories, indexing them, and entering them into a database to make them accessible. In order to collect stories, it is productive to tell a story to experienced folks about a problem you have. Then ask them if they are reminded of a similar experience. Usually they are. Having collected stories, we must decide what the stories teach us. We tell stories with some point in mind, so the indexing process tries to elucidate what that point is, given a situation. Schank (1990) believes that indexes should include the experience and the themes, goals, plans, results, and lessons from the story. Themes are the subjects that people talk about. Goals motivated the experience. Plans are personal approaches to accomplishing those goals. Results describe the outcome of the experience. The lesson is the moral of the story — the principle that we should take away from the case. Indexing is an engaging analytical process, the primary goal of which is to make the stories accessible.While indexing, we must continually ask ourselves under what situations we would be reminded of this story.

Indexed stories are then entered into a database. The indexes that we construct to describe the stories become the fields of the database. Each field describes the elements of the story on which we may want to retrieve a story. So indexes (fields) may include context, actor, learned lesson, result, or similarity. This process is more fully described by Jonassen and Hernandez-Serrano (2003). It is the process of semantically organizing the story for inclusion in the database that requires the model and conceptual understanding.

Conclusion

The concept of mental models is a powerful construct for describing the meaning that learners make. After describing the various methods that have been used to manifest or assess mental models, we argued that mental models are multi-dimensional, so no single form of assessment can be used effectively to describe mental models. In order to manifest mental models, learners need to use computer-based modeling tools to externalize their mental models in the form of computer-based models. Because mental models are multi-dimensional, no single modeling tool can manifest the complexity of mental models. So, we suggest that learners use Mindtools to construct models of structural knowledge, procedural knowledge, visual knowledge, and experiential knowledge. Research is needed to identify the most effective combination of modeling tools for representing the underlying complexity of mental models.

References

Anzai, Y., & Yokoyama, T. (1984). Internal models in physics problem solving. *Cognition and Instruction, 1*(4), 397-450.

Azzarello, J., & Wood, D. E. (2006). Assessing dynamic mental models. *Nurse Educator, 31*(1), 10-14.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology, 98*(1), 182-197.

Carley, K., & Palmquist, M. (1992). Extracting, representing, and analyzing mental models. *Social Forces, 70* (3), 601-636.

Chi, M. T. H. (2006). Laboratory methods for assessing experts' and novices' knowledge. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 167-184). Cambridge, MA: Cambridge University Press.

Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis*. Cambridge, MA: MIT Press.

Frederiksen, J. R., White, B.Y. & Gutwill, J. (1999). Dynamic mental models in learning science: The importance of constructing derivational linkages among models. *Journal of Research in Science Teaching, 36* (7), 806-836.

Gentner, D., & Stevens, A. L. (Eds.). (1983). *Mental models*. Hillsdale, NJ: Erlbaum.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 99-129). Hillsdale, NJ: Erlbaum.

Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology, 83* (1), 88-96.

Gomez, R. L., Hadfield, O. D., & Housner, L. D. (1996). Conceptual maps and simulated teaching episodes as indicators of competence in teaching elementary mathematics. *Journal of Educational Psychology, 88* (3), 572-585.

Gott, S. P., Bennett, W., & Gillet, A. (1986). Models of technical competence for intelligent tutoring systems. *Journal of Computer Based Instruction, 13* (2), 43-46.

Hardiman, P. T., Dufresne, R., & Mestre, J. P. (1989). The relationship between problem categorization and problem solving among experts and novices. *Memory and Cognition, 17*(5), 627-638.

Hong, E., & O'Neil, H. F. Jr. (1992). Instructional strategies to help learners build relevant mental models in inferential statistics. *Journal of Educational Psychology, 84* (2), 150-159.

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

Jonassen, D. H. (1987). Verifying a method for assessing cognitive structure using pattern notes. *Journal of Research and Development in Education, 20* (3), 1-14.

Jonassen, D. H. (2000). *Computers as mindtools for schools: Engaging critical thinking.* Columbus, OH: Prentice-Hall.

Jonassen, D. H. (2006). On the role of concepts in learning and instructional design. *Educational Technology Research and Development, 54* (2), 177-196.

Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural knowledge*: *Techniques for assessing, conveying, and acquiring structural knowledge.* Hillsdale, NJ: Lawrence Erlbaum.

Jonassen, D. H., & Henning, P. (1999). Mental models: Knowledge in the head and knowledge in the world. *Educational Technology*, *39* (3), 37-42.

Jonassen, D. H., & Hernandez-Serrano, J. (2002). Case-based reasoning and instructional design: Using stories to support problem solving. *Educational Technology: Research and Development, 50* (2), 65-77.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of applied psychology, 78* (2) 311-328.

Kruskal, J. B. (1964). Nonmetric multidimensional scaling: A numerical method. *Psychometrika, 29*, 115-129.

Lehrer, R., & Schauble, L. (2003). Origins and evolution of model-based reasoning in mathematics and science. In R. Lesh & H.M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, teaching, and learning* (pp. 59-70). Mahwah, NJ: Lawrence Erlbaum. McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naïve beliefs about the motion of objects. *Science, 210,* 1139-1141.

Morrison, M., & Morgan, M. S. (1999). Models as mediating instruments. In M. S. Morgan & M . Morrison (Eds.), *Models as mediators: Perspectives on natural and social science* (pp. 10-37). Cambridge: Cambridge University Press.

Nersessian, N. J. (1999). Model-based reasoning in conceptual change. In L. Magnani, N. J. Nersessian, & P. Thagard (Eds.), *Models are used to represent reality.* New York: Kluwer Academic/Plenum Publishers.

Newell, A., & Simon, H. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 7-14). Hillsdale, NJ: Erlbaum.

Preece, P. F. W. (1976). Mapping cognitive structure: A comparison of methods. *Journal of Educational Psychology, 68*, 1-8.

Roddick, C. S. (1995). *How students use their knowledge of calculus in an engineering mechanics course.* ERIC Document Reproduction Service No. ED 389 546.

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin, 100*(3), 349-363.

Rowe, A. L., & Cooke, N. J. (1995). Measuring mental models: Choosing the right tools for the job. *Human Resource Development Quarterly, 6*(3), 243-255.

Royer, J. M., Cisero, C. A., & Carlo, M. S. (1993). Techniques and procedures for assessing cognitive skills. *Review of Educational Research, 63*(2), 201-243.

Sabers, D. S., Cushing, K. S., & Berliner, D. C. (1991). Differences among teachers in a task characterized by simultaneity, multidimensionality, and immediacy. *American Educational Research Journal, 28,* 63-88.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(5), 484-494.

Schvaneveldt, R. W. (Ed.). (1990). *Pathfinder associative networks: Studies in knowledge organization*. Norwood, NJ: Ablex.

Schank, R. C. (1990). *Tell me a story: Narrative and intelligence*. Evanston, IL: Northwestern University Press.

Silver, E. A. (1979). Student perception of relatedness among mathematics verbal problems. *Journal for Research in Mathematics Education, 10*(3), 195-210.

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. Siegler (Ed.), *Children's thinking: What develops?* (pp. 325-348). Hillsdale, NJ: Erlbaum.

Southerland, S. A., Smith, M. U., & Cummins, C. L. (2000). "What do you mean by that?": Using structured interviews to assess science understanding. In J. J. Mintzes, J. H. Wandersee, & J. D. Novak (Eds.), *Assessing science understanding: A human constructivist view* (pp. 72-95). San Diego, CA: Academic Press.

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.

Whitfield, D., & Jackson, A. (1982). The air traffic controller's "picture" as an example of a mental model. In G. Johannsen & J. E. Rijnsdorp (Eds.), *Analysis, design, and evaluation of man-machine systems* (pp. 45-52). London: Pergamon Press.

Wilson, J. M. (1994). Network representations of knowledge about chemical equilibrium: Variations with achievement. *Journal of Research in Science Teaching, 31*(10), 1133-1147.

Wilson, J. R., & Rutherford, A. (1989). Mental models: Theory and application in human factors. *Human Factors, 31*(6), 617-634.

Wittgenstein, L. (1922). *Tractatus logico-philosophicus*. London: Routledge.

Wu, H. K., Krajcik, J. S., & Soloway, E. (2001). Promoting understanding of chemical representations: Students' use of a visualization tool in the classroom. *Journal of Research in Science Teaching, 38* (7), 821-842.

8. Model-Centered Instruction, the Design, and the Designer

Andrew S. Gibbons

Brigham Young University

Abstract: A model of instruction described by Wenger (1987) identifies three elements that are active during instruction: the mental model the instructor wishes to share with the learner, the external experience used to communicate the mental model, and the evolving mental model of the learner. Gibbons (2003a), writing in response to Seel (2003), noted this three-part description as a bridge concept relating learning and instruction. This view has important practical implications for designers of instruction. For example, Gibbons and Rogers (in press) propose that there exists a natural layered architecture within instructional designs that corresponds with instructional functions. Among these layers is the content layer, which determines the structural form in which learnable subject-matter is stored and supplied to the learner. This may include the expression of the content in terms of tasks, semantic networks, rules, or other structures. The designer's commitment at the content layer strongly constrains all other parts of the design, making some future decisions imperative, some irrelevant, and defining the range of possibilities for still others. One possible content layer commitment is to select the *model* structure as the basic unit of analysis. Having made the model the primary content structure commitment influences designer choices within other layers. This chapter describes the implications for designers of a model content commitment. It describes the constraints automatically placed on other layers of the design.

Keywords: Model-centered instruction; instructional design; design layers; design languages.

Introduction

This chapter discusses implications for the structures included in an instructional design when the subject-matter consists of a model of a dynamic system. Wenger (1987) identifies three elements that are active in such instruction: the mental model the instructor wishes to share with the learner, the external experience used to communicate the mental model, and the evolving mental model of the learner: "Of central importance…is the notion of model: the model of the domain, model of the student, and model of the communication processes…" (p. 7). According to Wenger, this model of instruction leads to a "radical" methodological shift for designers: "…The primary goal becomes computational understanding of processes of knowledge communication rather than an optimization of the production of systems." As a result, he says, the designer's problem becomes creating a representation of knowledge, which eventually results in "a mapping of knowledge into a physical medium" (p. 312).

This mapping of conceptual content which exists only in the mind of the designer or the subject-matter expert onto a physical medium creates a subtle distinction, which Wenger feels has much practical significance.

> It is useful to be able to speak about the knowledge that is the source of this mapping, and we will use the adjective *epistemic* to refer to this 'disembodied' level. Whether such an epistemic level really exists in some Platonic sense is not the point here. The claim is, rather, that the distinction between the epistemic and the representational levels is *useful* in designing and evaluating models of communicable knowledge. (p. 312, emphasis in the original).

Wenger's isolation of the representational (mediated) model has further implications when we consider that there are two ways in which the computer medium can represent the knowledge: on the one hand *invisibly* within an information-processing engine, and on the other hand as *sensations* at a sensory surface where the learner can experience the model.

> In fact it is a distinction between two forms of representation of the knowledge to be conveyed; we will simply call these the *internal* and *external* representations, respectively. …This perspective is a useful one for tutoring systems. Indeed, not only does the interface language map meaning onto a system of symbols, but this 'external representation' can actually compete with the internal representation as a vessel for domain knowledge (p. 312, emphasis in the original).

Model-Centered Instruction

Gibbons (2003a), commenting on Seel (2003), supported Wenger's distinction between the expert's and the learner's mental models and the "experience used to communicate the model". This distinction presents a challenging design problem for which designers have few formal design concepts. According to Gibbons, "the question of interest…is design structure. How do we harness and focus the structuring principle of the state-changing model—not only as an influence on the type of knowledge learned, but as an influence on the nature and structure of the design itself"(p. 296). This influence is felt in many parts of the design:

> Designers need (but do not currently have) the ability to describe classes of design architecture and discuss them in professional discourse. Design architecture here does not refer to software architecture or instructional strategy architecture: it refers to the architecture of the entire design, but most importantly to the central structural element of the design: those structures of the design that are decided first, that determine and shape the rest of the design, and that do not change as details and finish are added to a design (p. 296).

Gibbons developed a design theory of model-centered instruction (2001) for the purpose of exploring the design implications of dynamic-model content. Model-centered instruction is instruction that is carried out through interaction with dynamic models, and the experience with the model is supplemented by the activities of a learning companion that may supply a variety of coaching, feedback, and other learning support services. Varieties of model-centered instruction are created by considering all of the variations of this basic configuration.

In a model-centered design the initial commitment to the dynamic model as the "central structural element" of a design places constraints on those design decisions that follow it. These constraints remain in force as long as that original commitment to dynamic modeling is maintained. Among these constraints are:

- Constraints on the type and execution of instructional strategy employed as an augmentation of the learner's experience interacting with the model.
- Constraints on the types and actions of controls given to the learner for managing the model experience.
- Constraints the kinds of message that can pass from the model and its augmentations to the learner.
- Constraints on the representation of the model to the learner (what Wenger would call the "external" representation).
- Constraints on the media logic used to execute the model and its augmentations.
- Constraints on the collection and use of data generated during the model experience.

"Constraint" is used here in the sense that Stokes (2005) described, in which a constraint is both a limitation (a closing off of certain options) and an opportunity for innovation (an opening of new options).

Model-Centered Instruction and Simulation

To this point, the reader may have assumed that the terms "simulation" and "model-centered instruction" are synonymous. However, I do not believe that these terms should be used interchangeably. Each represents a way of viewing the assumptions of an instructional design from a particular perspective. *Simulation* (of the instructional variety) usually refers to an external representation and a type of experience afforded to a learner. A learner is said to "use" a simulation. *Model-centered instruction* refers to a product and experience *architecture* that involves many layers of organization, some visible and some completely invisible. Implicit in a model-centered architecture is a commitment to one or more abstract dynamic models (Wenger's *internal* model) which must be represented to a learner (by Wenger's *external* model) in a way that communicates the essential aspects of the expert's model to a learner, who uses the communication in the construction of personal knowledge.

An example of this is an instructional methodology described by Brown and Palincsar (1989) as reciprocal teaching. The core activity of reciprocal teaching is the use of pre-assigned questions asked by learners as a means of mining a text reading or a shared observational experience in order to comprehend its meaning. The details of how this is accomplished are not as important to the present purpose as is the statement of the principle by which reciprocal teaching works. In describing this, Brown and Palincsar name the activities and then identify the operational principle of reciprocal teaching:

> …These…strategic activities…structure intrapersonal as well as social dialogues. Reviewing content (summarizing), attempting to resolve misunderstandings (clarifying), anticipating possible future text development (predicting), and assessing the state of one's gradually accumulating knowledge (questioning) are all activities that the experienced learner engages in while studying independently, by means of an internal dialogue. The reciprocal teaching procedure renders such internal attempts at understanding *external*. Reciprocal teaching provides social support during the inchoate stages of the development of internal dialogues. In the course of repeated practice such meaning-extending activities, first practiced socially, are gradually adopted as part of the learner's personal repertoire of learning strategies (p. 415, emphasis in the original).

Reciprocal teaching relies on the choreographed joint activities of several learners to produce a visible *model* of comprehension which, if observed and experienced repeatedly by a learner may be internalized and used as a personal comprehension process. Since this dynamic model of that comprehension process constitutes the key subject-matter to be learned, reciprocal teaching is a design configuration that can be termed "model-centered". All design decisions that follow after the model decision revolve around this central commitment and are conditioned by it. Should some other design factor be given higher priority in the design, it would become strategy-centered, message-centered, or media-centered instead (see Gibbons, 2003b). Gibbons and Fairweather (2000) describe several innovative designs emerging from research which are quite diverse in their surface features but that share an underlying model-centered design architecture. Contrasting simulation and model-centered instruction clarifies the important point that model-centered instruction focuses attention on the entire architecture of the design, placing priority on the model as a structural foundation. The remainder of this chapter describes the impact of the commitment to model-centering on the remaining layers of a design. In this discussion, the reader should keep in mind that though true simulations have a model-centered architecture, many members of the class of model-centered designs (reciprocal teaching being an example) do not look on the surface like simulations and would be construed by many as not being simulations.

Design Layers

Gibbons and Rogers (in press) describe instructional designs structurally, providing a way to consider the "remaining layers of design". A layered instructional design sees the total design problem in terms of many individual sub-problems:

- Every instructional design must solve a *representation* problem by providing specifications for the part of the instructional artifact that can be sensed (through sight, sound, touch, smell, etc.) by exposure to media or realia. (This is Wenger's *external* model).
- Every instructional design must solve a *strategy* problem by describing the patterns of tutorial conversational exchanges that can be engaged in between learner and instruction source, the setting in which they take place, and the social configuration and roles of all participants.
- Every instructional design must solve a *messaging* problem by providing a description of the system of individual messages that can be communicated from the instruction source, in service of the instructional strategy, and for the purpose of driving the selection or construction of representations to the learner. The solution to the messaging problem supplies a bridge between abstractions in the strategy layer and concretions in the representation layer.
- Every instructional design must solve a *control* problem by specifying the communication controls and related symbols through which the learner will be able to communicate choices, responses, and strategic control over the instructional source.
- Every instructional design must solve a *media-logic* problem by describing the manner in which the functions of all of the other layers will be enacted by humans, instructional media, or some combination of both.
- Every instructional design must solve a *data management* problem by describing the elements of data from the instructional interaction that will be captured, recorded, stored, analyzed, reported, and used to influence the ongoing course of the instructional interaction.

Content Layer Constraints in Model-Centered Designs

The only layer of the design problem that is addressed by a commitment to model-centering is the *content* problem: in model-centered instruction the content (which includes model state data, subject-matter information, and dynamic change information) is supplied through computation of changing model states. Model-centered instruction assumes that learners will be enabled to observe and interact with three types of dynamic model: (1) models of cause-effect systems, (2) models of performance with respect to those systems, and (3) models of environments that influence either the performance or the cause-effect systems. It assumes that learners will either observe the operation of models or perform operations on models to observe the effects. Selection of the appropriate model or combination of models is critically important. Bransford et al., (2000), asserts that “one is struck by the complexity of selecting appropriate models for particular mathematical ideas and processes” (p. 168). A designer must avoid unnecessarily complex models which have variables that are of no consequence to the learner and must be careful to select a model that leads to the desired processing by the learner.

The commitment to the model at the content layer of the design imposes limitations and provides opportunities (both of which can be considered constraints) at all other layers of the design. The sections that follow describe some of these.

Strategy Layer Constraints in Model-Centered Designs

Models themselves can only supply information on changing model states. The model itself produces no commentary on its own actions, no insight into the play of forces in its inner workings, and no scaffolds to support incremental learning (see, for instance, Clancey, 1984a). It is possible to learn from an unaugmented model by observing and experimenting with it, but the efficiency of such learning is low and can lead to misconceptions. Therefore, most instructional model experiences are accompanied by supports that assist the learner during observation and operation of the model (Gibbons, 2001; Gibbons et al., 1997). The strategic design principles described in this section do not comprise a complete list of scaffolding augmentations for model-centered instruction. The ones included have been chosen to illustrate the important structural differences implied by a decision to use a model-centered design architecture. Additional types of augmentation during model experience are described by Gibbons et al. (1997).

One type of augmentation includes supplying one or more problems to frame the learner’s exposure to and interaction with the model (Gibbons, 2001). The problem serves as a lens and a mask during the model experience. As a lens, problems stimulate learner interaction with and observation of model details. As a mask, problems focus interaction on only selected relationships within the complete model, allowing relationships of immediate interest to be foregrounded for consideration. The learner can either solve a problem or observe a problem being solved as a means of exposing key relationship information of momentary interest. Problems may take a number of forms, but a model-centered commitment implies that problems of some type will be used. Model experience in the absence of problems is a possibility, but in such unaugmented explorations it can be seen that the learner becomes the problem poser and that actions toward the model in service of learning are evidence of self-posed problems (or questions), even if they consist of something as simple as "What happens when I press this button?". The design of the model can be influenced by the types of problem the designer intends to use.

A second type of strategic model augmentation consists of initiative sharing in the selection roles and goals. Gibbons (2001) describes several strategic decisions learners may share or fully control:

- Role negotiation (observer, participant, main agent, exclusive agent)
- Initiative negotiation (learner, instruction source, shared)
- Performance goal selection (at any of several levels of granularity)
- Problem selection (at any of several levels of granularity)
- Strategic goal selection (for problem solving approach)
- Means selection/design (for strategic goals selected)
- Means execution
- Evaluation of goal achievement.

A third type of augmentation used to supplement model experience consists of conversational tutorial messaging support during model interaction. Messaging is discussed in more detail in a later section. Strategically, however, it is important to note that familiar structures of exposition and information-delivery that are used in the design of more traditional forms of instruction are subordinated in model-centered instruction. Model-centering does not encourage the use of long information presentations, so the designer must think more in terms of the conversation the learner is having with the model, expressed through choices and interactions.

Control Layer Constraints in Model-Centered Designs

The design of control systems takes on special importance in model-centered designs. Controls of several types are required in the less-structured environment created by model-centered instruction. Gibbons et al. (in press) names them:

> …Sets of special-purpose controls that serve needs related to several simulation functions: (1) controls that allow the learner to act upon the model, (2) controls that adjust patterns of augmentation, (3) controls that adjust the representation of the model or the viewpoint from which the learner can observe the presentation, and (4) controls over personal data reporting for monitoring outcomes, performance, progress, trends, history, and scheduling.

Learners use controls to convey messages to the instructional source. In combination, the control and messaging systems provide the two-way communication channel through which learner and instructional source communicate. Controls and messaging are thus the medium through which interactions proceed. In traditional instructional forms, control systems are so standard that they tend to fuse with other aspects of the design. In a model-centered design, control systems must be invented which are related to the characteristics of the content model(s), the support functions, and conversational patterns of the strategic augmentations, so they tend to be more customized. Crawford (2003) suggests that the beginning point of the design of such control systems is to define the "verbs" that represent actions the learner can take during interactions.

Message Layer Constraints in Model-Centered Designs

The instructional conversation referred to earlier takes its structure from the strategic instructional augmentations of the strategy, but the expression of a strategy as a conversation sometimes entails a complex pattern of learner-to-instruction exchanges. Message layer structures provide an intermediate mapping entity that allows the larger intentions of the strategy to be expressed in terms of smaller particles of expression. There are many examples of message structuring systems in the literature, including Merrill's Component Display Theory (1994), Horn's information mapping (1997), systems for the recording and analysis of classroom conversation (Simon & Boyer, 1974), and more recent ones for analysis of peer-to-instructor and peer-to-peer conversation (Sawyer, 2006).

Message structures provide the possibility of flexible, unfolding patterns of communication for a single strategic intention. Moreover, they allow a single communication intention to map onto multiple representation forms. Other kinds of communication must also be provided for in addition to the strategic ones. Messages must be conveyed to the learner about control availability, access to personal learning data (progress, scores, etc.), and access to alternative representation modes, such as different perspective or representation style (schematized, literalized, abstracted, etc.).

The design for a messaging system centers around an ontology of message types and a catalogue of messaging patterns. The first concern of message layer design for model-centered instruction is to enumerate the basic patterns of moment-to-moment exchange that the designer feels will permit the expression of the full range of strategic, control, data, and representation information. The designer also must define the rules for interpreting messages, such as those described by Fox (1993) for technology-based media:

- Interruptions by the learner should be possible
- Thing reference (pointing) and object sharing should be more formalized
- Silences should be flagged with intent

- Communication of backchannel cues (emotional states, body language, attitude) should be facilitated
- Multiple sequential messages should be possible from the same speaker without a break (e.g., musings “aloud”)
- Short delays in correction might be deliberately used to signal to the student the need to think and respond again
- Ways should be found to make the learner’s process actions (thinking) known to the tutor.

When message design has been executed, the designer has the core of a mechanism by which instructional messages can be generated dynamically using computed data from the model, the strategic function, and other sources. This was an early goal of some intelligent tutoring systems (Carbonell, 1970; Collins et al., 1975; Clancey, 1984b). More recently, Drake et al. (1998) have used a messaging design to allow a simulation to generate messages and representations from a combination of computed data and primitive message fragments during presentation, demonstration, and practice stages of instruction.

Representation Layer Constraints in Model-Centered Designs

Up to the present, the representations—the sensed surface elements of instruction—for both live and technology-based instruction have tended to be static and unchanging, with a relatively small seasoning of dynamic ones. Since model-centered instruction is grounded in the principle of making it possible for the learner to sense state changes, forces, and trends of change, this typical balance between static and dynamic representations is reversed. Moreover, what is represented changes as well. The most common (and affordable) tradition has been to show static 2-dimensional opaque surfaces superimposed with static symbolic enhancements (arrows, auras, etc.) intended to illustrate flow and dynamism. Model-centered instruction favors dynamic 4-dimensional effects incorporating integral dynamic symbolic elements that illustrate changes in multiple invisible forces at once. This constraint is important because it is the dynamic operation of invisible forces that most often constitutes the basis for understanding dynamic models.

Model-centering introduces new terms into the representation lexicon for designers used to traditional and low-cost approaches to representation. Designers must consider refresh rates, strict synchronization of multimedia events, multi-perspective views, intelligent display assembly and coordination, storage and controlled replay of representation event sequences, correlation and synchronization of stylistic modes (schematic, literal, metaphorical, etc.), time and space warping (slow-down, speed-up magnification, diminution, zooming), navigation, time-trace representation, and multiple message-to-representation mappings. Rather than thinking of representation resources as stored, pre-composed, static elements, the model-centered designer thinks in terms of data-driven, generated media experiences where possible and families of well-crafted animation sequences where it is not.

Media-Logic Layer Constraints in Model-Centered Designs

Media-logic consists of the set of rules necessary to stage the events incorporated in a design. Media-logic is an essential element of live instruction as well as technology-based instruction; it generates the sequence of events during instruction. For model-centered instruction of all kinds it consists of algorithms, heuristics, or explicit instructions used by a computer or a human instructor that direct how to make instructional events occur. Media-logic should not be confused with the strategic decision-making function (within the strategy layer) that determines which events might take place. Gibbons et al. (2001) describes media-logic as "the place where the designer's abstract instructional constructs and concrete logic constructs [of] the development tool come together".

Media-logic executes models, executes augmentation computations, executes message formation, executes representations, accepts control input, and executes data-related functions. Moreover, it integrates and coordinates the order of computation for these functions. (Keep in mind that "computation" here includes human instructor decision making and judgment.) This integration and coordination most frequently takes the form of a continuous cycle of activities, many of which can occur in a parallel sequence, where that is possible. Baldwin and Clark (2000) describe the economics of functional modularization with respect to logic functions. Munro et al. (2006) demonstrates that this principle applies to model-centered forms of instruction.

Data Management Layer Constraint in Model-Centered Designs

Model-centered designs can make much different use of data management functions than non-model-centered instructional forms. Because model-centered instruction entails learner interactions within a dynamic context, it is possible for model-centered instruction to generate much larger volumes of data from an instructional encounter. Moreover, that data can be interpreted with respect to the momentary state changes within that context. An action at Time A can be interpreted as having a particular meaning in terms of the learner's knowledge; the same action at Time B may have a much different interpretation.

Because the data generated during such interactions is interpretable and can be used in future instructional decisions, much more data can be captured. In some cases, the volume of this data prohibits immediate processing, so provision for data storage, eventual analysis, and visualization becomes an important design factor. Following the analysis of this volume of data, it may not be possible to describe the performance within the environment in terms of a few simple indices. Because of this, model-centered instructional designs can challenge the most widely-held metaphors of instructional management systems as simply score repositories. Where the volume of data does allow immediate processing, the results of processing can be reported to strategic functions that use the results to make real-time modifications to the model and its augmenting processes.

These considerations make the design of data management more involved. The designer must consider when and how often data will be collected, the granularity of data recorded, the interpretation rules (whether the data is processed immediately or after a delay), and the use of the results of interpretation, both by the learner, and by the instructional source.

Conclusion

The purpose of this chapter is to demonstrate the effects that ripple through the many layers of a design when a specific commitment is made within one layer of the design. Using proposals by Gibbons (2003a) that content can be described in terms of dynamic models, I have traced the implications of a commitment to model content within the strategy, control, message, representation, media-logic, and data management layers of the design. This analysis has highlighted the many differences within each layer of the design attendant to the content decision. A similar analysis based on a different content commitment or a similar commitment within any of the other layers would demonstrate the same result: any layer of a design is sensitive to decisions made in other layers, and decisions made within one layer constrain decisions within other layers, either by eliminating or creating design possibilities.

This finding should stimulate designers to examine more carefully the assumptions that are often built into their designs. It should also lead to more detailed examination of classes of design, of which model-centered is but one example. I propose that doing so may result in the recognition of distinct classes of design that are based on underlying structural differences rather than on surface appearances. Such a perspective encourages thinking about designs and the creation of designs in a new light—one that sees the abstract operational principle of a design as being a tool for generating not just individual designs but whole new families of designs that may appear much different on the surface but owe their genesis to a similar underlying architecture.

References

Baldwin, C. Y., & Clark, K. B. (2000). *Design rules, vol. 1: The power of modularity*. Cambridge, MA: MIT Press.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, D. C.: National Academy Press.

Brown, A. L., & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Carbonell, J. (1970). AI in CAI: An artificial intelligence approach to computer-assisted instruction. *IEEE transactions on man-machine systems*, 11, 190-202.

Clancey, W. J. (1984a). Extensions to rules for explanation and tutoring. In B. G. Buchanan & E. H. Shortliffe (Eds.), *Rule-based expert systems: The MYCIN experiments of the Stanford heuristic programming project.* Reading, MA: Addison-Wesley.

Clancey, W. J. (1984b). Use of MYCIN's rules for tutoring. In B. G. Buchanan & E. H. Shortliffe (Eds.), *Rule-based expert systems: The MYCIN experiments of the Stanford heuristic programming project.* Reading, MA: Addison-Wesley.

Collins, A., Warnock, E., & Passafiume, J (1975). Analysis and synthesis of tutorial dialogues. In G. Bower (Ed.), *The psychology of learning and motivation (Vol. IX).* New York: Academic Press.

Crawford, C. (2003). *The art of interactive design*. San Francisco, CA: No Starch Press.

Drake, L., Mills, R., Lawless, K., Curry, J., & Merrill, M. D. (1998). *The role of explanations in learning environments.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Fox, B. A. (1993). *The human tutorial dialogue project: Issues in the design of instructional systems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gibbons, A. (2001). Model-centered instruction. *Journal of structural learning and intelligent systems, 14*(4), 511-540.

Gibbons, A. S. (2003a). Model-centered learning and instruction: Comments on Seel (2003). *Technology, instruction, cognition, and learning*, 1(3), 291-9.

Gibbons, A. S. (2003b). What and how do designers design?: A theory of design structure. *Tech trends, 47*(5), 22-27.

Gibbons, A. S., & Fairweather, P. G. (2000). Computer-based instruction. In S. Tobias & J. D. Fletcher (Eds.), *Training and retraining: A handbook for business, industry, government, and the military*. New York: Macmillan Reference USA.

Gibbons, A. S., Fairweather, P. G., Anderson, T., & Merrill, M. D. (1997). Simulation and computer-based instruction: A future view. In C. R. Dills & A. J. Romiszowski (Eds.), *Instructional development paradigms* (pp. 769-805). Englewood Cliffs, NJ: Educational Technology Publications.

Gibbons, A. S., Lawless, K. A., Anderson, T. A., & Duffin, J. R. (2001). The web and model-centered instruction. In B. R. Khan (Ed.), *Web-based training*. Englewood Cliffs, NJ: Educational Technology Publications.

Gibbons, A. S., McConkie, M., Seo, K. K., & Wiley, D. (in press). Theory for the design of instructional simulations and microworlds. In C. M. Reigeluth & A. Carr-Chellman (Eds.), *instructional-design theories and models, Volume III*. Mahwah, NJ: Erlbaum.

Gibbons, A. S., & Rogers, P. C. (in press). The architecture of instructional theory. In C. M. Reigeluth & A. Carr-Chellman (Eds.), *Instructional-design theories and models, Volume III.* Mahwah, NJ: Erlbaum.

Horn, R. E. (1997). Structured writing as a paradigm. In C. R. Dills & A. J. Romiszowski (Eds.), *Instructional development paradigms*. Englewood Cliffs, NJ: Educational Technology Publications.

Merrill, M. D. (1994). The descriptive component display theory. In M. D. Merrill & D. G. Twitchell (Eds.), *Instructional design theory*. Englewood Cliffs, NJ: Educational Technology Publications.

Munro, A., Surmon, D., & Pizzini, Q. (2006). Teaching procedural knowledge in distance learning environments. In R. Perez & H. O'Neal (Eds.), *Web-based learning: theory, research, and practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sawyer, R. K. (2006). Analyzing collaborative discourse. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences*. Cambridge, UK: Cambridge University Press.

Seel. (2003). Model-centered learning and instruction. *Technology, instruction, cognition, and learning*, 1(1), 59-85.

Simon, A., & Boyer, E. G. (1974). *Mirrors for behavior III: An anthology of observation instruments*. Wyncote, PA: Communication Materials Center in Cooperation with Humanizing Learning Program, Research For Better Schools, Inc.

Stokes, P. (2005). *Creativity from constraints: The psychology of breakthrough*. New York: Springer.

Wenger, E. (1987). *Artificial intelligence and tutoring systems*. Los Altos, CA: Morgan Kauffmann Publishers.

9. Realizing Model-Based Instruction

The Model of Model-Based Instruction

Ulrike Hanke

University Freiburg, Germany

Abstract: The theory of mental models assumes that learning is a form of information processing. When individuals are confronted with a new learning subject or task, they have to construct a mental model integrating their already existing knowledge and the new information from the learning environment. This mental model is not stored immediately, but rather has to be reconstructed several times to become a schema, i.e. to be learnt. In this sense, learning consists of different subprocesses which have to be supported by the five teaching interventions of the model of model-based instruction (MOMBI). (1) In order to initiate mental model construction, teachers have to ask real questions and present problems or conflicting information that makes students think ("provocation" teaching intervention). (2) Then they have to activate the students' preconceptions ("activation" teaching intervention) and (3) present information in order to enable the students to answer the question/solve the problem/explain the conflicting information ("presentation" teaching intervention). In order to make sure that students construct mental models that are similar to expert models, i.e. conceptual models, (4) teachers have to ask students to externalize their models and give them feedback ("scaffolding" teaching intervention). (5) Finally, they have to give them the opportunity to reconstruct their models several times in order to store and schematize them ("practice" teaching intervention).

Keywords: Mental model; schema; teaching; instruction; learning.

Introduction

The overall goal of teaching is to support and to facilitate the learning of students. In this sense, teaching is successful when the students have learnt the things the teacher intended them to learn (more quickly or more reliably than they would have alone) (Schulz, 1969). Therefore, teaching has to consider the learning process (Seel, 2003), and every decision a teacher makes has to be based on the learning process and its subprocesses. This is the only way teaching can be an intervention and successful in the sense that it facilitates and supports learning.

Following this line of argumentation, teaching may be described as any intervention by a teacher that supports the learning process of students.

In order to support and facilitate the learning process of their students, successful teachers have to possess deep knowledge about the learning process (Seel, 1998). This knowledge enables them to choose the teaching intervention that is most appropriate in a specific teaching situation.

Therefore, the first part of this article considers the learning process and tries to describe it in as much detail as possible. It will be referred to as the assumptions of model-based learning.

The second part describes the teaching interventions of the model of model-based instruction (MOMBI), which is based on the assumptions and findings of model-based learning. This model contains various teaching interventions that correspond directly to the subprocesses of learning.

The third part describes a design for a class of 15-year-old pupils on job contracts in order to give an example of a possible realization of the model of model-based instruction (MOMBI).

Model-Based Learning

The assumptions of model-based learning are based on the theory of mental models (Seel, 1991, Johnson-Laird, 1983; Gentner & Gentner, 1983). This theory describes how individuals are able to understand and explain phenomena/information of the world that they have never experienced before in spite of the restriction that their knowledge is incomplete and fragmentary. It is assumed that individuals construct so-called mental models with the help of their available knowledge (preconceptions) in order to overcome these restrictions. These models have predictive and explanatory power for understanding the phenomena of the world (Norman, 1983).

Mental models are a form of knowledge representation (Seel, 2001, 1991, 1986; Hegarty & Just, 1993) that is constructed in order to make new information plausible/understandable/useful and to allow inferences (Ifenthaler, 2006; Pirnay-Dummer, 2006; Seel, 1991, 2003). They are defined as constructs of the human mind (Seel, 1986, 1991, 2001; Hegarty & Just 1993), are based on existing knowledge, and are constructed intentionally when an individual is not able to assimilate, i.e. is not able to understand, new information immediately (Piaget, 1976). They may be wrong in a scientific sense, i.e. not similar to the models of experts in the field (Hillen, Berendes, & Breuer, 2000; Park & Gittelman, 1995; Halford, 1993; Johnson-Laird 1989; Dörr, Seel, & Strittmatter, 1986; Norman, 1983), but this does not matter because for the model-constructing individual the value of a model is defined by its plausibility (Seel, 1991, 2000; Ifenthaler, 2006) and its usefulness (Pirnay-Dummer, 2006). A model is plausible and useful for the model-constructing individual if it is able to explain the phenomena/the information that was not understood at first. Therefore, mental models are per definition correct for the model-constructing individuals.

Since mental models integrate new information with already existing knowledge, they serve the function of facilitating information processing visualizing information, simulating processes that can not be observed, and enabling analogical transfer (Seel, 1991).

In this sense, constructing a mental model is the basis for learning a new concept because before something can be learnt, it has to be understood (Seel, 2003; Hilgard & Bower, 1975), and therefore a student first has to construct a mental model. However, learning is more than information processing. Learning requires for the new information not only to be processed but also stored in such a way that it can be retrieved in the future. Therefore, it has to be related to the entities and their changes in the real world.

Learning does not have to be an intentional act; sometimes learning happens unintentionally when an individual is confronted with, processes, and stores new information that often conflicts with existing knowledge. This non-intentional learning takes place because the new information provokes a mental state of disequilibrium (Piaget, 1976). Since people generally attempt to maintain mental equilibrium, they immediately try to understand the new information (Buggle, 1997; Trautner, 1991; Piaget, 1976) and integrate it with their existing knowledge. In order to understand the new information and to retain equilibrium, they have to activate their existing knowledge, compare this knowledge with the new information, gain new knowledge in order to reconstruct it, and compare it again with the new information and their existing knowledge.

In this way, the individual constructs a mental model that explains the new information and enlarges and/or restructures their existing knowledge (Seel, Dörr, & Dinter, 1992; Seel, 1991). The mental model is therefore the result of information processing that retains an individual's mental equilibrium. Therefore, this process of information processing is also called mental model construction/building (Seel, 1991) and should be understood as an accommodative process in the sense described by Piaget (1976).

It can be concluded that this process of mental model construction consists of different subprocesses: At first the construction of a mental model is initiated by new, often conflicting information in the environment because it causes a mental state of disequilibrium (Piaget, 1976). In order to process and make sense of this conflicting information, the individual activates relevant prior knowledge (Seel, 1991; Hilgard & Bower, 1975). Then the individual searches for further information in the environment and in his or her existing knowledge and integrates all of the information and the existing knowledge in a mental model (Seel, 2005, 2000, 1991). This mental model is then fleshed out (Johnson-Laird, 1983) until it makes the new information plausible, i.e. until the individual understands the new information (Ifenthaler, 2006; Seel, 1991) and/or until it is useful for the model-constructing individual (Pirnay-Dummer, 2006). At this point, the process of mental model construction is stopped.

As stated before, the construction of a mental model thus enlarges or restructures existing knowledge (Seel, Dörr, & Dinter, 1992; Seel, 1991), but this enlargement or reconstruction is not permanent because mental models are not stored (Seel, 1991; Gick & Holyoak, 1983, 1980). They are constructed to explain new information, i.e. to make information plausible and understandable, but they have to be reconstructed every time they are needed. In order to be able to retrieve or understand this kind of information later on without constructing a new mental model, the individual has to process similar information several times. This subprocess of reconstruction is the subprocess of schematization. This is a process of abstraction. In contrast to the mental model at the start of the process, the schema is a generalized knowledge structure that is able to make sure that the same or similar information does not provoke mental disequilibrium the next time and that it is understood immediately, i.e. automatically without the construction of a mental model. As soon as a mental model has been schematized, i.e. constructed several times and stored in an abstracted and generalized way as a schema, we can say that the new information and its handling has been learnt.

Therefore, learning can be defined as the construction of a mental model and its schematization. This form of learning consists of five subprocesses (cf. table 1):

Table 1. Subprocesses of learning

Subprocess of learning	Description
Subprocess 1	provocation of mental disequilibrium
Subprocess 2	activation of prior knowledge
Subprocess 3	search for further information
Subprocess 4	integration into a mental model
Subprocess 5	schematization

Learning is initiated by new information from the environment (subprocess 1) that provokes a state of mental disequilibrium (Piaget, 1976). In order to understand this information and retain equilibrium, the individual activates his or her prior knowledge (subprocess 2), searches for further information in the environment (subprocess 3), and integrates all information and prior knowledge into a mental model (subprocess 4) (Seel, 1991; Johnson-Laird, 1983). This construction process is stopped as soon as the individual succeeds in constructing a mental model that makes the new information plausible for them (Ifenthaler, 2006; Pirnay-Dummer, 2006; Seel, 2003, 1991). In order to understand similar information immediately in the future, the individual has to construct the model several times and store it as a schema (subprocess 5) (Hanke, 2006; Seel, 1991).

In order to facilitate learning, these subprocesses of learning have to be supported by teaching, i.e. teachers have to support these five subprocesses.

Each subprocess of learning can be supported by at least one teaching intervention. These teaching interventions are described in the following chapter.

Model of Model-Based Instruction (MOMBI)

Teaching is successful if it facilitates learning (Schulz, 1969). In order to facilitate learning, each teaching intervention should consider the subprocesses of learning introduced above. Since the learning process can be divided into five subprocesses, we can distinguish between five different teaching interventions. Each intervention supports one subprocess of learning and can be realized in different ways.

Since learning is initiated by new, often conflicting information (subprocess 1), a teacher can provoke learning with the help of new or conflicting information which brings about a conceptual change (Duit, 1999). Provoking learning is therefore a teaching intervention ("provocation").

This intervention is seldom used because most teachers seem to forget that they have to explain to their students why they should learn or why they should reflect on the learning subject. This does not cause problems for students who are interested in the subject, i.e. who are intrinsically motivated. But students who are not intrinsically motivated will not learn or even think about the learning subject. Such students do not have any reason to do so as long as they are not at least extrinsically motivated. One way to try to motivate them is to provoke mental disequilibrium (Piaget, 1976).

This can be realized by confronting students with conflicting information that they do not understand immediately. If teachers succeed in this, the students will try to balance the new conflicting information with their available knowledge. The mental imbalance therefore fosters learning. In this way, "provocation" can be understood as a teaching intervention: the first teaching intervention of the model of model-based instruction (MOMBI).

After the learning process has been initiated (subprocess 1/"provocation" teaching intervention), students have to activate their prior knowledge (subprocess 2) in order to construct a mental model that can explain the new conflicting information. This subprocess can also be supported by teaching. As a teaching intervention teachers can instruct their students to activate their prior knowledge, which should support the mental model construction because mental models are based on prior knowledge. Activating prior knowledge therefore enables individuals to find initial explanations for the new, conflicting information and shows them connections between the new information and prior knowledge. In short, it facilitates the process of mental model construction.

Whereas the first teaching intervention – "provocation" – is not widely used, the usefulness of the "activation" teaching intervention is widely accepted (Seel, 2000; Ausubel, 1968).

In most organized learning situations, students do not have enough prior knowledge to explain the new conflicting information. The teacher will thus have to present more information to them. This presentation of further information can have various forms: Information can, for instance, be presented in the form of a speech by teachers. Alternatively, teachers can provide their students with texts or other sources which they can use to get the information they need for constructing their mental models. Such sources can be pictures, concept maps, tables, various online sources, or even experts or whole libraries. The only important thing about this teaching intervention is for the students to get enough information to construct a mental model that explains the new information/makes the new information plausible for them.

A topic often discussed in the context of model-based instruction is the presentation of external models such as concept maps for structural knowledge or visual representation for the entities that are supposed to be stored. Since mental models are assumed to have an image component, the presentation of concept maps or visual representations is thought to be effective in supporting the construction of mental models and thus also learning. There is a lot of research showing this effect (Seel, 2000, 1986; Mayer, 1989; Dörr, Seel, & Strittmatter, 1986).

Nevertheless, presenting external models is not the only way to realize the "presentation" teaching intervention because there are other possibilities for supporting mental model construction.

One of these possibilities which has often been discussed in the context of model-based instruction is for the teacher to act like a model in showing the learners how to do something. This is often done in apprenticeship situations when an expert models a special skill to the novice in order to make the novice learn. As the Cognitive Apprenticeship (Collins, Brown, & Newman, 1989) shows, this can also be realized with cognitive elaborations (cf. method "modeling" of the Cognitive Apprenticeship, Collins, Brown, & Newman, 1989). This kind of modeling is supposed to support the students in constructing their own mental models about a special skill.

However, the intervention of presentation does not only intend to present the information to be learnt but also to make the students realize their misconceptions. For the acquisition (or development, construction) of knowledge the teacher has to show why a certain concept or theory is not valid for the phenomena that the students observe and why the depiction should be changed for the entity that is under discussion.

As has been shown, there are various possibilities for realizing the "presentation" teaching intervention because there are various possibilities for preventing students from getting lost in searching for relevant information and offering them the relevant information they need, i.e. for supporting learning subprocess 3.

This teaching intervention – "presentation" – is often seen as the most typical activity of teachers. When people are asked what teaching means for them, most of them think of teachers in front of classes presenting new topics and explaining new things.

Although "presentation" is often seen as the most typical activity of teachers and although it is the teaching intervention that has inspired the most research, "presentation" is not more important than the other interventions because without them learning only takes place if students are interested in the subject, i.e. if they are motivated intrinsically and want to understand the subject. Students who do not have this intrinsic motivation need more interventions, i.e. they need more support in order to learn effectively.

The function of the next teaching intervention of the MOMBI – “scaffolding” – is to make the students’ models fulfill the features and descriptions of the scientifically developed models. Since the construction of a mental model is stopped as soon as the model makes the new information plausible and/or is useful for the model constructor (Ifenthaler, 2006; Pirnay-Dummer, 2006; Seel, 2000, 1991), it is very important for teachers to make sure that the students’ models are not only plausible and useful for *them* but also correct in a scientific sense, i.e. similar to a conceptual model (scientifically correct model). For this reason, teachers have to ask questions and give hints in order to make their students think and continue constructing their mental models (Riedel, 1973). This teaching intervention can be compared to the method of scaffolding in the cognitive apprenticeship approach (Collins, Brown, & Newman, 1989), which is where it got its name (“scaffolding”). It accounts for the individual differences of the students, which is important because the construction of a mental model is a very idiosyncratic process.

Unfortunately, “scaffolding” can often only be realized inadequately or even not at all because there are so many students in a class that the teacher does not have the time to work with every one of them individually. As a result, at least some of the students come up with incorrect mental models.

The four teaching interventions described above (xxx) support the first four subprocesses of learning, i.e. the subprocesses of information processing.

But since mental models are not stored, a fifth subprocess of learning has to take place: the subprocess of schematization. During this subprocess the mental model has to be reconstructed several times to be stored as a schema. This subprocess can also be supported by teaching. In order to do this, teachers have to give students the possibility to practice so that they can reconstruct similar mental models several times and finally store them as a schema. Therefore, the teaching intervention which supports this learning subprocess is called “practice”.

Since “practice” aims at the whole process of model construction, it means that all of the other four teaching interventions have to be realized again in a more or less strict way. This time the intervention of “presentation”, for example, is less important because the students should already have the information necessary for constructing the mental model. Otherwise, they can use the sources they used before or ask the teacher or the other students. With ongoing practice sessions and, hopefully, ongoing schematization of the mental models, teachers should also fade out the “scaffolding” intervention because after this intervention students are supposed to be capable of explaining the new information without the help of a teacher. The “practice” teaching intervention aims at schematization.

To sum up, it is possible to distinguish between five different teaching interventions that support the five subprocesses of learning (see table 2):

Table 2. Subprocesses of learning an teaching interventions of the MOMBI

Subprocess of learning		Teaching intervention
Subprocess 1	provocation of mental disequilibrium	"provocation"
Subprocess 2	activation of prior knowledge	"activation"
Subprocess 3	search for further information	"presentation"
Subprocess 4	integration into a mental model	"scaffolding"
Subprocess 5	schematization	"practice"

1. To initiate the learning process, teachers have to provoke their students by asking a question, presenting new, conflicting information or problems, or giving them a task to solve ("provocation").
2. As mental models can only explain new information on the basis of existing knowledge, teachers then ask their students to activate their prior knowledge ("activation").
3. In most cases the prior knowledge of students is not sufficient for constructing mental models that are not only plausible and useful for the students themselves but are also scientifically correct. For this reason, teachers have to present further information ("presentation").
4. Then, teachers have to make sure that their students construct scientifically correct models ("scaffolding"). This involves supporting them individually by giving them hints and asking and answering questions.
5. In order to schematize mental models, students have to be given a chance to practice ("practice").

As already mentioned, all of these teaching interventions can be realized in different ways.

"Provocation", e.g., can be realized by asking a provocative question, showing a provocative picture, presenting new and conflicting information, or presenting a problem or case that has to be solved.

"Activation" can be realized in the form of classical brainstorming. It is also possible to ask the students to brainstorm alone or in small groups or to write their ideas on posters or paper. The results are then passed to the other students so that they can read their classmates' ideas.

To realize "presentation", teachers can present new information in the form of a presentation/speech. As discussed above, they can also realize modeling or provide external models or texts or give access to books or even a library or the internet. The most important thing is for the information presented to be limited in a certain respect so that the students do not get lost and are able to find the relevant information in order to be able to construct their mental model.

"Scaffolding" can be realized by asking questions, giving hints, and answering the questions of the students. In order to identify scientifically incorrect models (misconceptions), teachers can also ask their students to externalize their models in the form of maps or texts. These externalizations facilitate communication about the models under construction by enabling students and teachers to refer to specific aspects of the model.

Since “practice” is a repetition of the other four interventions, it can be realized with the help of the other methods. It could, however, be advantageous to vary class arrangements, i.e. let the students sometimes work in groups, sometimes in pairs, and sometimes alone. This makes the learning situation more diversified.

Realizing all of these teaching interventions should facilitate and support the learning of students. In order to show whether the realization of this model of instruction is really effective, several lessons would have to be designed and then implemented and evaluated. Studies designed to test the effectiveness of this model are currently being implemented. Unfortunately, there are not yet any results, and it is thus only possible to show how lessons can be realized using the MOMBI. The next paragraph therefore describes two lessons for 15-year-old pupils.

Realizing the Model of Model-Based Instruction (MOMBI)

In the following, two lessons for 15-year-old pupils about job contracts will be described in order to show how MOMBI can be realized. These two lessons are part of a unit on working life planning that consists of ten lessons. The pupils are supposed to learn essentials that they need for their working life. Two lessons cover job contracts, two lessons wage accounting, four lessons how to behave towards one’s boss and colleagues, and two lessons unemployment. Instruction in all four subjects is planned with the model of model-based instruction (MOMBI). The two lessons concerning job contracts are described in the following paragraphs.

The learning objective of these two lessons is for pupils to be able to read and understand job contracts and to identify whether the contract contains all of the necessary elements or whether something is missing.

The lessons are planned for 20 15-year-old pupils attending a class on preparation for their working life. The pupils come from various backgrounds: Some finished school regularly, some finished special school, some have not succeeded in finishing school. Many of the pupils come from immigrant families and have problems with the German language.

The lessons about job contracts start with a little story that the teacher reads aloud to the pupils. This story is supposed to “provoke” the pupils to engage in mental model construction by showing them why the learning objectives of the two following lessons are important, i.e. why it is important to be able to read and understand job contracts and to know the necessary points of job contracts. As described above, this teaching intervention is supposed to motivate the pupils to start mental model construction.

The story is an example of a youth who is fired because he did not read his job contract carefully and therefore behaved wrongly.

After this “provocation” with the help of the story, the teacher realizes the “activation” teaching intervention by asking the pupils the following questions: What is the story about? What mistake did the youth in the story make? What is important to pay attention to when signing a job contract? These questions as well as the pupils’ answers are written on the blackboard without comment. This brainstorming helps the students to systematically activate their prior knowledge about job contracts. Activating prior knowledge is extremely important for mental model construction because mental models are based on prior knowledge. Mental models are constructed when new information is integrated with prior knowledge (preconceptions). For this reason no mental model construction can take place without the activation of prior knowledge.

The third step is the “presentation” teaching intervention. In this step, the teacher presents the information the pupils need in order to construct a mental model. In the case of this realization of MOMBI, the teacher presents a checklist which contains the most important aspects of a job contract. The pupils are supposed to integrate this new information with their preconceptions or rather change their preconceptions.

After this teaching intervention, the pupils should have started to construct their mental models about job contracts, but it can not be assumed that they have come up with a complete mental model. Therefore, the teacher has to support the pupils in constructing their models and make sure that they construct correct mental models. As soon as the model is plausible for the model constructing individuals, they stop the construction. The teacher therefore has to make sure that the pupils do not stop construction as long as the models are not correct. This is the function of the “scaffolding” teaching intervention.

In order to realize “scaffolding”, the teacher hands out two incomplete job contracts and asks the pupils to read them carefully and find out which of the necessary aspects are missing. While working on this task, the pupils are allowed to use the checklist and discuss with a partner. In this way, they can go over the important aspects of job contracts again and again in order to integrate them into their mental models about job contracts.

While the pupils try to complete the incomplete contracts, the teacher is available to answer questions, supports the pupils, and tries to identify which pupils need help. After the pupils have completed this task, the teacher asks them to present their solutions as a second step of the “scaffolding” teaching intervention. This gives the teacher the chance to correct incorrect models, i.e. to recover misconceptions and make sure that the pupils have correct solutions, i.e. that they have constructed correct mental models.

At this point in the learning process, all of the pupils should have finished constructing their mental models.

The “practice” teaching intervention is realized in order to help the pupils to schematize their models. The teacher initiates this intervention by handing out further incorrect and incomplete job contracts. Since the pupils were asked to work in pairs for the first two contracts ("scaffolding" teaching intervention), they are now supposed to work alone. This means that support is reduced since they do not have a partner they can work with and share their ideas with.

When working on the first of the two contracts of the "practice" teaching intervention, they still have the possibility of using the checklist the teacher presented before. When working on the second contract, they do not have this possibility. In addition, the teacher fades out her scaffolding continuously. This does not mean that she does not pay attention to the pupils' solutions, but rather that she reduces her hints, the goal being for the pupils to be able to construct the models alone.

By this point, the pupils have reconstructed their models at least four times It can thus be assumed that the models are about to schematize because schematization is a process of reconstructing mental models. This repeated reconstruction is needed in order to generalize the models and abstract from the details of the models.

After these two lessons, the pupils should be able to read job contracts and recognize whether necessary elements of a contract are missing or even incorrect because they have learned how to construct mental models of job contracts and maybe even have a schema of job contracts.

These two lessons give an example of how the teaching interventions of the model of model-based instruction (MOMBI) can be realized in schools.

This model is not only appropriate for individual lessons in schools. It can also be realized in university courses or in continuing education.

Conclusion

The model of model-based instruction (MOMBI) is a systematic realization of the ideas of the theory of mental models and the model-based approach to learning. It realizes all of the assumptions concerning the functioning of learning in order to facilitate and support learning with the help of different teaching interventions.

These teaching interventions consider the subprocesses of learning and attempt to optimize them. As each of the interventions is designed to optimize exactly one of the processes, it can be assumed that the interventions facilitate and support learning by accelerating it or improving understanding.

None of the interventions has been systematically evaluated so far. In order to do so, more lessons and courses would have to be designed and implemented systematically. It would be important to ask whether the interventions are effective as well as whether they are accepted by the students.

These questions will be explored in various small-scale projects in the coming year in order to concretize the description of the teaching interventions and shed light on the strengths and weaknesses of the MOMBI.

References

Ausubel, D. P. (1968). *Educational psychology: A cognitive view.* New York, Chicago: Holt, Rinehart and Winston.

Buggle, F. (1997). *Die Entwicklungspsychologie Jean Piagets.* Stuttgart: Kohlhammer. 3. edition.

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

Dörr, G., Seel, N. M., & Strittmatter, P. (1986). Mentale Modelle: Alter Wein in neuen Schläuchen? *Unterrichtswissenschaft*, 2, 168-189.

Duit, R. (1999). Conceptual change approaches in science education. In W. Schnotz, S. Vosniadou, & M. Carretero (Eds.), *New perspectives on conceptual change* (pp. 263-282). Amsterdam: Elsevier Science.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner, & A. L. Stevens (Eds.), *Mental models* (pp. 99-129). Hillsdale, NJ: Lawrence Erlbaum.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*, 306-355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*, 1-38.

Hanke, U. (2006). *Externale Modellbildung als Hilfe bei der Informationsverarbeitung und beim Lernen.* Freiburg: FreiDok.

Hegarty, M., & Just, M. A. (1993). Construction of mental models of machines from text and diagrams. *Journal of Memory and Language*, *32*, 717-742.

Hilgard, E. R., & Bower, G. H. (1975). *Theories of learning.* Englewood Cliffs, N.J.: Prentice-Hall.

Ifenthaler, D. (2006). *Diagnose lernabhängiger Veränderung mentaler Modelle*: *Entwicklung der SMD-Technologie als methodologisches Verfahren zur relationalen, strukturellen und semantischen Analyse individueller Modellkonstruktionen.* Freiburg: FreiDok.

Johnson-Laird, P. N. (1983). *Mental models. Towards a cognitive science of language, inference, and consciousness.* Cambridge, UK: Cambridge University Press.

Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen.* Stuttgart: Klett.

Pirnay-Dummer, P. N. (2006). *Expertise und Modellbildung – MITOCAR*. Freiburg: FreiDok.

Schulz, W. (1969). Umriss einer didaktischen Theorie der Schule. In P. Fürstenau (Eds.): *Zur Theorie der Schule* (pp. 27-45). Weinheim: Beltz.

Seel, N. M. (1991). *Weltwissen und mentale Modelle.* Göttingen: Hogrefe.

Seel, N. M. (1998). Instruktionspsychologiesche Rezeption der didaktischen Reduktion. In H. Ahlborn, & J. -P. (Eds.), *Didaktische Vereinfachung – eine kritische Reprise des Werkes von Dietrich Hering* (pp. 237-260). Seelze-Velber: Kallmeyer.

Seel, N. M. (2000). *Psychologie des Lernens.* München und Basel: Reinhardt.

Seel, N. M. (2001). Epistemology, situated cognition, and mental models: 'Like a bridge over troubled water'. *Instructional Science*, *29*, 403-427.

Seel, N. M. (2003). *Stochastic models of change.* Unpublished manuscript.

Seel, N. M. (2005). Modellzentriertes Lernen und Lehren. In: B. Lehmann, & E. Bloh (Eds.), *Methodik und Content-Management* (pp. 198-223). Hohengehren: Schneider-Verlag.

Seel, Norbert M., Dörr, G., & Dinter, F. R. (1992). Lernen im Medienverbund – Eine Untersuchung zum Erwerb neuen Wissens. *Unterrichtswissenschaft*, *20*, 275-288.

Seel, Norbert M. (1986). Wissenserwerb durch Medien und "mentale Modelle". *Unterrichtswissenschaft*, *14*, 384-401.

Trautner, H. M. (1991). Lehrbuch der Entwicklungspsychologie. Band 2: Theorien und Befunde. Göttingen: Hogrefe.

Engineering the Model Environment

Pablo Pirnay-Dummer and Dirk Ifenthaler

University of Freiburg, Germany

Preparing the Luscious Meals of Technology and Applications

With the discussion of the first transfer at hand (see Preface, figure 1), the second transfer comes in sight: using the theories of learning and instruction as the basis to design various applications and environments for all kinds of subject domains. This part contains different perspectives on the transfer from theory into practice as well as concrete technologies and applications for several different target audiences. It contains contributions from Sanne Dijkstra & Henny Leemkuil, Andrei I. Podolskij, Gerhard Weber, Patrick Blumschein, Jim Pellegrino & Sean Brophy, and Tristan E. Johnson & Wen-hao D. Huang.

Dijsktra and Leemkuil start with the design of complex model-based electronic learning environments which follow the systems approach to education. Revising general models of instructional design and the ADDIE model, the authors discuss the specifics of mental models and their role as a tool for knowledge construction in gaming and simulation. They then provide a study which uses simulation to support model based learning for the domain of knowledge management. One key focus lies on the support of the learner's knowledge construction process.

Podolskij provides a framework to support researches in constructing technological theories to help designers construct proper applications. The chapter thus covers both transfers and all three fields addressed in this book. The framework, which includes several stages, is created by deriving models of transfer on the basis of mental actions. The author concludes his chapter with a three-model framework which shows a valid path for a stepwise transfer of fundamental research to design and application.

Experience from the use of intelligent tutoring systems and simple adaptive learning environments are the basis for a synthesis towards model-based learning environments in Weber's chapter. The architecture, program code, and user modeling capabilities are presented within field-tested and running environments. The computational and technological aspects of design are discussed for practicability and feasibility in the context of learning goals. Weber shows how the chances and demands for model-based learning environments have changed in recent years due to both aspects.

Blumschein maps model centered learning to a generic process of instructional design, starting with an overview about the strengths and weaknesses of common instructional design approaches. He exemplifies his vision with examples from problem based learning, anchored instruction, and model based learning, both for technological theory and concrete application. Blumschein concludes by suggesting a paradigm shift towards a more product- and at the same time learner-oriented educational science, which would grant immediate access to the design process by means of a model-centered instructional design.

Pellegrino and Brophy extend the well-known principles of anchored instruction in two subsequent frameworks to fulfill their specified principles of instructional design, thus aiming at powerful technology-supported learning environments. The first framework implements a process-oriented progression cycle of work and revision and provides tested and standardized tools to support learning with high effects. The second framework provides a generalized inquiry model and a technology support tool to apply the process cycle domain independently. Thus, both students and teachers are able to use the same technology to organize and structure their content technologically in the same way. The authors generalize the frameworks further, with special remarks for K12, adult, and further education.

Johnson and Huang show in this chapter how games can be used to generate learning environments which are especially suitable for supporting model-centered learning. Games which use mental model theory to solve major problems for complex skill development are considered to be a bridge between computer-based and lively learning environments and thus aim at experience. Since games are necessarily based on models which are under full control of the game designer, they can be mapped to learning goals. The authors discuss the experience game environments afford as well as their interactivity, complexity, interrelatedness, and realism, all of which are necessary and sufficient attributes for model-based learning environments. Johnson and Huang show significant game characteristics which lead to concrete design constraints for games to support transfers from learning theory, instructional system design, and modeling. The applicability and feasibility are discussed with a broad focus on workforce learning.

10. Developments in the Design of Instruction

From Simple Models to Complex Electronic Learning Environments

Sanne Dijkstra and Henny Leemkuil

University of Twente, The Netherlands

Abstract: This chapter addresses the relationship between the psychology of cognition and learning and the design of instruction in its development during the last half of the former century and the following years in this century. The findings of research on cognition and learning and the applications for the design of instruction become integrated with the developments in information and communication technologies. The chapter shows both the development of the generic ADDIE model and the design of highly specific situated simulation and games.

Keywords: Adaptive learning; ADDIE model; complex learning environments; instructional design; learning support; simulation-based learning.

The Growth of Information and its Influence on Instructional Design

The ever increasing amount of information and problem-solving procedures and the regular changes of both led and still lead to the question how to pass on the content and sequences of operations to members of future generations in such a way that they can use this as knowledge and skills. That question concerns both the education and training in schools and in other organizations such as governments and industry. Of course, this question is not new and answers have been given in the last three centuries. However the increasing amount and complexity of the information and methods content and the need to pass on much of this to as many human beings as possible, made it necessary to continue the study for answers. That study is part of instructional design (ID), an applied science that became established in the second half of the former century. The study and research on ID led to a substantial body of design knowledge and methods and to many useful instructional programs. The purpose of this chapter is to provide a concise overview of the developments of ID and to show how these developments have led to rather diverse results: to the general ADDIE model on the one hand to situated instructional games on the other. This will be explained and illustrated.

The field of Instructional design (ID). The organizations and situations of education and training comprise the field of ID. The goals of these organizations are to pass on the information and problem-solving methods to members of future generations or to the members of the organizations. ID will at least support these goals. The design of instruction is the design of the communication between an expert (teacher) and a novice (student) in such a way that the student will acquire the knowledge, skills and attitudes that are the goal of the curriculum. Instruction is any intended activity to make that acquisition possible or easier to accomplish. The acquisition of knowledge and skills is a psychological process. Instruction should facilitate that process in the best possible way. The students have to cognize the content of the instructional communication and to practice the procedures in order to become skilled performers. The result of the design of instruction can be delivered as an instructional program in printed or electronic format in order to either be used by individual students for self-study or to be used with all kinds of help of an expert.

Foundations of Instructional Design

There have been many developments in the design of instruction. The authors cannot do justice to all developments here. In Europe the concept of didactics is used. In a recent publication Seel and Dijkstra (2004a) provided an overview of studies into this concept and its relationship with the concept of ID. Instructional design started and became established as an applied science in the last half of the former century. Seel and Dijkstra mention three sources for the development of a theoretical basis for the design of instruction, which will be shortly outlined in the next paragraphs. These are (a) the psychology of cognition and learning, (b) the engineering or systems approach to education, and, (c) the information and communication technology. A fourth source is the epistemology of knowledge acquisition.

The Psychology of Cognition and Learning and Instructional Technology

There will be no doubt that learning is a psychological process, though actually the whole human organism is involved during learning (Dijkstra, 2004a). Before the 1950s the foundation of instruction in the psychology of learning was often made, but a direct relationship between the science of learning and an instructional technology was missing. Skinner (1954, 1958) was the first to state the rules for solving an instructional-design problem and to construct a device — the teaching machine — to apply the rules that make learning possible. Though the idea of a science of learning and a technology for teaching and learning was admired, the interpretation of all learning as instrumental conditioning was soon abandoned. Moreover the rules for programming instruction and the use of teaching machines did rarely lead to the promised results. As Seel and Dijkstra concluded: 'The application of the rules led to a splitting up of the subject matter into a huge number of instructional frames that were even able to prevent the integration of the concepts and methods involved. The teaching machines were not able to adapt the instruction to the students' learning flexibly' (p. 5). A new theory, a new technology and better devices were needed. These will be discussed in the next section. Nevertheless, Skinner realized the elsewhere productive relationship between science and technology for the psychology of learning and marked the beginning of instructional design as an applied science. Moreover he developed technical equipment for the application of the design rules.

The Systems Approach to Education

The whole process of teaching and learning can be described in a sequence of components. For this process Glaser (1964) used the technical term instructional system that consisted of five components: (a) the instructional goals or system objectives, (b) the students' entering behavior of system input, (c) the instructional procedures or system operations, (d) the performance assessment or output monitor, and, (e) the research and development logistics. The objectives were formulated in observable behavior; the instructional procedures and the assessment of results were founded in the theory of learning and educational measurement. The systems approach became influential for ID. An instructional system provides the instructional designer a soft technology to hold on (see next paragraph).

Technology and Technical Equipment

A technology is the whole of the science (theory and research methods) that is valid for a domain and of the rules for solving a design problem in that domain in order to realize a public or individual goal. For example, chemical technology comprises the theory and research methods that are valid for the molecules of substances (their structure and their change) and the rules to construct devices and installations for producing substances that will be used for a public or individual goal, such as refineries for producing gas for transport and heating (Dijkstra, 2004a, p. 17). If the rules are well-defined the authors suggest the label hard technology. If they are ill-defined, the label soft technology is preferred. Soon after it became available, the technical equipment that was based on laws of physics was used in education and training. As Seel and Dijkstra stated, in the 1950s and 1960s the teaching machine, television and the language laboratory were used in schools. Nearly all this equipment became replaced by microcomputers. By the end of the 1970s the information and communication technology entered the field education and training definitely. Though minicomputers were already used for instructional purposes in the early 1960s, the invention of the microcomputer by the end of the 1970s had a tremendous influence on education, both in the use of the hardware as well as in software. The increasing qualities of the computer equipment were: (a) smaller equipment, but more information storage and (b) faster in processing information. The digitalization of all information, including large amounts of visual information and the developments of communication technology led to a pervasion of all sectors of the industrial and everyday world (Seel and Dijkstra (2004b). For instructional communication the computer equipment did replace nearly all other devices that were used previously, such as the slide, the overhead and the film projector. For providing information a computer and a beamer produce at least the same quality of text and pictures. Streaming video, which is a sequence of movie parts that is sent in compressed form over the Internet, makes it possible for the students to see a movie for instructional purposes, without the need to first download the whole movie. Thus the student does not waste time. Today, for an appropriate computer-supported instructional communication the student can interact (a) with the subject matter content, both with the objects that are mediated and with the text, and can get the best possible feedback immediately at any distance from school, (b) with the teacher at any distance, (c) with peer students at any distance, and (d) with the Internet as a library. Besides the hardware qualities the software is still becoming more “intelligent” in initializing and supporting instruction and learning. Special applications, such as animations, simulations and games may help the students’ understanding of the subject matter. All the features of the computer, both hardware and software, became integrated in education and training.

The Epistemology and Knowledge Acquisition

If the goal of education and training is described as the acquisition of knowledge and skills the instructional designer needs to have at least some idea of what knowledge is and what this means for the acquisition of knowledge. Situated cognition (Brown, Collins and Duguid, 1989) and constructivism (Glasersfeld von, 1996; Jonassen, 1992; Piaget, 1937) will get some attention in one of the next sections.

The Developments in the Psychology of Cognition and Learning

About 1960 and afterwards the model of instrumental conditioning as the only model for the description and interpretation of human learning was rejected. The change of the theories of learning became known as the cognitive shift that led to several new concepts and interpretations of the human actions. The authors of this chapter first make a general assumption. They assume the existence of cognitive entities, such as a semantic network and cognitive processes, such as the mental operations in problem solving. The conceptions of the psychology of cognition that influenced the development of ID as an applied science encompassed the existence of categories of human learning, of cognitive processes and of memory structures. It has led to many new conceptions about cognition and to several applications. For this chapter the categories (conditions) of learning and the conceptions of semantic network, cognitive structure and mental model will be discussed concisely, more or less in chronological order.

Bruner, Goodnow, and Austin (1956) showed the existence of cognitive processes for the formation of concepts. The participants in their studies constructed identification algorithms in order to categorize objects. Based on tentative categorizations of artificial objects, such as rectangles and crosses, they constructed provisional concepts. As soon as their categorizations proved to be always correct either conjunctive or disjunctive concepts were formed. Thus concept learning as a category of learning could be described and the process of acquisition could be shown in detail. A few years later Melton (1964) published a book on categories of human learning, soon followed by a comparable book on the conditions of learning (Gagné, 1965). The latter seminal publication is often considered as a basis for the design of instruction. Gagné distinguished different types of learning that could be realized if special conditions for that type were met. Among these types he mentioned stimulus-response learning, chaining, verbal association, multiple-discrimination learning, concept learning, principle learning and problem-solving. Of these types of learning he provided examples all taken from the curricula of elementary and secondary education. All types of learning had an outcome that is described in observable behavior. The types of learning were ordered from simple to complex. Gagné showed some learning structures. The ultimate objective of a learning structure was the acquisition of complex behavior. A learning structure is a sequence of different types of learning that are organized in a hierarchy. One example given showed the structure for learning to read larger units of text in English language (p. 201). The content of this structure was ordered from simple to complex types of learning. The instructions to support the learning of the content of these types should follow the order given. For structuring more complex content, consisting of concepts and principles, the label learning hierarchy was introduced. For determining the most plausible structure of a learning hierarchy the final learning outcome must be broken down into outcomes on a prerequisite level that is one step lower in the hierarchy, and so on, until the level of concepts and principles, the content of which the students already know is reached (p. 188). In order to be able to acquire the concepts and principles at a certain level, those of the lower level have to be mastered. The results of research by Gagné and his co-workers (e.g. Gagné et al., 1962) did support the description of the conditions of learning and the concept of hierarchies of learning.

In the following decades the idea of different categories or conditions of learning was influential. For example Merrill (1983) and Reigeluth (1983) used the same categories. The research on concept learning and problem solving did strongly support the assumption of the existence of these categories of learning. The learning structure and learning hierarchy, which were constructed after the final learning outcome was analysed into the lower levels of concepts and principles of a domain, could be compared with the idea of cognitive network or cognitive structure. These are assumed cognitive patterns of organizations of knowledge components. Such patterns are represented in graphs. Strong support for the idea of a hierarchical network was provided by Collins and Quillian (1969). The Collins and Quillian semantic network model, which was represented in a graph, organized concepts in a hierarchy of subcategories in such a way that the most general concept was represented at the top of the hierarchy and the most specific concept was represented at the bottom. All concepts (categories) at a lower level were subcategories of those at a higher level. Collins and Quillian (1969) supposed that the properties of the instances of a category are stored only once at the highest possible level in the hierarchy. For this principle they found evidence in the reaction times on determining the correctness of statements. For example, the duration of the reaction times to verify the correctness of the following three statements about properties (a) a canary can sing, (b) a canary can fly, and, (c) a canary has skin, showed a significant increase. Property (b) is stored at the represented category birds and (c) at the category animal. Starting at the category canary to check for the relevant property, the traversing of the hierarchical network takes time. A semantic cognitive network can take on other structures. For the representation of the causal structure of a domain Collins and Stevens (1983) used an and/or graph. Such representations are useful for the design of a detailed instructional communication. They showed how different rules of inquiry teaching can be used to enable students to extend their fragmentary knowledge of such a structure into the complete knowledge network. Each inquiry rule provides the students with a problem that can be solved by reasoning with the knowledge they have. If the students don't know or seriously hesitate the teacher can provide the information.

Though the assumption of a pattern of knowledge components as a cognitive structure is useful for the design of instruction, it does not interpret the dynamics of knowledge use to reach a goal. For that the concept of mental model will be discussed.

Mental models and goals. Human behavior is motivated, either for doing a job or for other goals. If a motive is active human beings will execute the steps to reach the desired goal. They will work from the actual situation to the desired goal state in a goal-directed interaction with the environment. For appropriate actions they need a model of the complex reality in which they work or function. Only understanding the processes and procedures to change the given situation make the fulfillment of the motive possible. Therefore understanding the processes of change and the procedures to design, develop and use artifacts are such important goals of education. For the goal-directed interaction with the environment it is supposed that human beings construct mental models. As Norman (1983) described it: "In interacting with the environment, with others, and with the artifacts of technology, people form internal, mental models of themselves and of the things with which they are interacting. These models provide predictive and explanatory power for understanding the interaction." (p. 7). Mental models are representations of real or imagery situations. As Seel (2004) wrote: "..models are always constructed in accordance with specific intentions of the model building person. Models are always representations of something; they represent natural or artificial objects, so-called originals, which can be again models of something." (p. 54). Mental models are the result of reflection about objects and how these change. Everyday human beings make use of mental models or are constructing and changing models. For example the development of a mental model of the road plan in a new city area. Sometimes international agreements are made about the representations of objects. For example the particle model in chemistry to designate atoms and molecules and their change is represented all over the world in the same way, both in icons and in symbols (see the chapter of Jonassen and Young in this volume). Human beings are creative in constructing mental models that are used as metaphors. The reader is referred to the container model, an example discussed by Lakoff (1987).

How can the concept of mental model be useful for the design of instruction? The authors suggest that the instructional designers make problems in such a way that the students need to observe and manipulate concrete objects or representations of these. The information given must lead the students to observe defining and characteristic features of the objects. The goal of the manipulation of the objects is that the students can find regularities in the behavior of those objects. These objects must be available or represented in such a way that the students can make predictions about what will happen over time. Information and communication technology can be of substantial help, especially if the objects are complex and situated. An elaborate example of a company as an object will be given later in this chapter.

The new conceptions of the psychology of cognition and learning formed a rich foundation for instructional design. Nevertheless it did not lead to full agreement about the rules to apply among instructional designers as will be shown in the following sections.

General Models of Instructional Design

The label design is used by Gagné (1965) when he outlines the predesign of the conditions of learning and outlines the principles of design for those conditions. A few years later the label instructional design is used (Gagné & Briggs, 1974). In the following decade more and different instructional-design models were published. Nearly all models clearly show the phases of solving a design problem, comparable to solving practical problems in the engineering sciences (Dijkstra, 2000).

In his first attempts to solve instructional-design problems, Gagné integrated the at that time relevant conceptions of the psychology of learning, both from behaviorist psychology and from the first studies of cognition into the design of instruction. This led to emphasis on the description of learning outcomes as capabilities and to the categorization of the learning content as mentioned afore. In the following decades the ideas became worked out (Dick & Carey, 1978; Gagné & Briggs, 1979) into instructional-design models. Gagné and Briggs (1979) extended the categories of educational goals to intellectual skills, cognitive strategies, verbal information, motor skills and attitudes. The original categories of concepts and principles became subsets of intellectual skills: concrete and defined concepts, rules, higher order rules and problem solving. The phases or series of steps of the instructional-design models start with a description educational goals (see Gagné & Briggs, 1979) in terms of the students' behavior or capabilities. The more general labels knowledge and skills were not used in those years. These conceptions were directly useful as design rules, such as the analysis of the content of subject matter and to the analysis of educational goals into hierarchies of concepts and principles. These rules were helpful for the overall instructional design. For separate conditions of learning the design rules could be worked out into nine events of instruction (see Gagné & Briggs, 1979, for a detailed description). Other scholars emphasized different aspects in their instructional-design models. In an anthology, edited by Reigeluth (1983), different instructional-design theories and models were presented. These models show the phases or clusters of steps how to solve instructional-design problems. It was supposed that the models did apply for all domains and fields. However those who invented the models nearly always used isolated examples of domain knowledge and problem-solving procedures from mathematics and physics, in any case well-defined procedures. Nearly all models structure the content of the subject matter into the categories that Gagné had provided. The models specify the concepts that can be used to describe the content of the curriculum and the main variables that should be considered for the design of the instructional messages. They further provided rules to describe the educational goals and often give examples of achievement test items that can be administered to the students in order to evaluate their performance. The instructionaldesign models show much overlap. By the end of the 90's of the former century, a generic approach, already recognizable in the Principles of Instructional Design by Gagné and Briggs (1979), was referred to as the ADDIE model.

The ADDIE Model of Instructional Design

The characters mean analysis, design, development, implementation and evaluation. A clear origin of the acronym is not found (Molenda, 2003). The model became a pragmatic tool for instructional-design projects. The results of each phase are evaluated and used to alter or reinforce the steps in former phases (feedback function).

Analysis

Depending on the purpose of the project, the analysis phase becomes worked out into needs analysis, subject matter content analysis and job analysis. This results in a description of the learning objectives for a certain group of students. The description will profit from an assessment of the students' knowledge and skills that are conditional to understand the course content. The analysis phase results in the program of requirements described as knowledge and skills to be acquired and in the design of a prototype of an achievement test to be administered by the end of the course, together with the description of a criterion score that marks the level of knowledge and skills that the students should reach.

Design

In the design phase the instructional-designer makes a general plan for the arrangement of the content of the instruction. It contains the categories of problems to be solved by the students, the procedures how to solve these, and the concepts, hypotheses and theories that the students should understand, remember and use.

Development

The development phase results in the course materials that are ready for use. During the development the evaluation of the first release leads to corrections and to a second release of the course materials. This process can be repeated until a satisfactory product is constructed. If the materials will be presented online, corrections are based on the students' errors. Regular formative evaluation with a group of students will support the development of a useful product. The developers will require an expert evaluation as well.

Implementation

The implementation phase may comprise different jobs, such as training the trainers, scheduling the courses, preparing a time table, scheduling evaluation sessions and so on.

Evaluation

Finally in the evaluation phase different assessments can be made, such as: (a) the students' affective reception of the course, and (b) a measurement of their achievement. Do the knowledge and skills that are acquired meet the learning objectives and the criterion that they were told? Do they apply what is learned in their jobs? The results of the evaluation serve the purpose of feedback, both for the instructional designers who can improve the design and the course materials. Moreover they provide feedback for the boards of executives of companies and of schools whether their goals became realized.

At first sight the generic model seems useful for those who are responsible for education and training in school systems and in business and industry, certainly if they don't have much knowledge of the psychology of cognition and learning. Most probably the thorough analysis of the subject matter and the direct approach of instruction to which many students are used, underlines their success. In many training institutes the model was accepted and became leading for the design of instruction. Today many universities show the model on their websites, the purpose of which is to help their professors with the design of their instructions. And many training institutes advertise the model to their clients to convince them of their expertise in course design. In spite of the growing use of the model as a generic model for instructional design criticism remains, both on practical and theoretical grounds.

A first practical criticism refers to the often too detailed prescriptions for the five phases, as a result of which the designers don't see the wood for the trees. It makes the use of the model inefficient and ineffective (Gordon & Zemke, 2000). Secondly, the model leads to a linear way of working, which can easily result in a rigid course program that is unable to resemble the flexible communication between a teacher and a student. The model should use the advantages of digital technologies, such as rapid prototyping (Tripp & Bichelmeyer, 1991) and should leave room for individualization.

The models were seriously criticized on epistemological grounds (e.g. Jonassen, 1992). Too often the students were provided information about objects, instead of providing opportunities for the construction of knowledge from the real objects or from those that are mediated (Seel & Winn, 1997). A separate task of media selection, which became part of the set of design rules, for example selection of classroom instruction, written materials or e-learning is at best pragmatic.

It ignores the essential meaning of the concept medium as a way to represent the reality. The authors' students of instructional design had serious difficulties to analyze the information of a textbook into concrete and defined concepts, rules (principles) and higher-order rules. This was one of the reason's to abandon the use of those categories and start from problem solving as the students' activity to construct knowledge from the reality. The reality is specified by the concept of object of a domain. An object can be categorized, its change can be interpreted and it can be designed as an artifact. Information about the objects can be provided to the students and questions can be asked. These questions are the problems of categorization, of interpretation and of design. The label object can mean any entity that is perceived or imagined. Instructions are about objects and what to do with them (Dijkstra, 2004b). These include (a) real objects such as plants, birds, houses, cars, the earth; (b) inferred objects, such as atoms, the psyche, a group; and (c) systems, such as the solar system, a taxonomy, a machine, a company, a political system. For the design of instruction the object can be real or needs to be mediated for and during instruction (Dijkstra, 1997, 2000, 2004b). The mediated object should help the student to develop a mental model of it.

These criticisms about the ADDIE model make sense. The designers are suggested to start with a general description of the final goal and then work with broad outlines of the phases of the model. The designer should prevent to get swamped into a too detailed set of steps at the start of the procedure. Further a pretest and regular formative evaluations during the development phase can make clear what the students already know and what their learning needs are. Though the ADDIE model is a useful model for instructional design the user should be aware that it only is a general heuristic.

Mental Models and the Design of Instruction

Human beings have an ability of modeling the world. They anticipate a new situation that they expect or predict to happen from their actions or from the events they observe. The expectations are based on experiential knowledge and the theories they know about the change of objects. (Seel, Ifenthaler, & Pirnay-Dummer, in press). It is supposed that the development of mental models is a long term process. The models can develop to a high level of abstraction, for example a model of a refinery. The assumptions on the development of mental models lead to a design of a learning environment and a choice of real or represented objects that are needed for manipulation by the students, both for categorizing the object and for eliciting change. Moreover the objects are needed for the reflection on their behavior. During instruction for the acquisition of knowledge students try to develop a mental model of the objects that are used to clarify and explain the knowledge. And they manipulate the objects for studying their features, their change and for practicing their skills. The mental model is used to plan future actions and to predict the results of an action. During the whole period of elementary and secondary education the models of the reality and of domains of that reality change and are becoming increasingly complex. The models of real and mediated objects play a crucial role in solving both well- and ill-defined problems. For example, the value of assets at the stock market can be calculated precisely, based on the values of the variables given, but the prediction of the amount of change of value over a certain time lapse is very difficult if not impossible.

For functioning in a complex environment the use of complex ill-defined procedures is needed, such as logistics for building construction, rules for leading a department, a company, an army, and so on. In such situations different variables and their interactions do influence the outcome of human actions. How can instruction help the students to develop a mental model of the organization, its structure and the effects of possible actions on the output of the system or organization? Cognitive flexibility theory and mental model theory were used and did support to design the instructions for learning complex ill-defined procedures that can be applied in complex environments (e.g. industries).

Virtual learning environments. How can students learn to construct a mental model of a complex reality such as an engine or an industrial enterprise? This is possible in a long time apprenticeship that can be supported by a simulation of the object (e.g. engine, organization, and so on). Achtenhagen (2004) provides an example of an industrial enterprise. It is supposed that the mental models change as knowledge and skills increase and if making predictions with a model fail.

The instructional design for the construction of mental models of complex artifacts of technology and of complex structures in an educational setting meets some difficulty. An interaction with the environment is needed, thus on the job training prevails. However, unskilled personnel can easily make mistakes, which involve a risk for damage and personal safety. Moreover the complexity of a system may evoke a model that is only partly useful and will easily lead to errors. To prevent these, the use of a complex electronic learning environment may be designed and turn out to be helpful.

If a complex reality such as an industrial enterprise is simulated for the purpose of learning, its structure should be depicted in such a way that the effect of an intervention and the consecutive process can be illustrated. As Achtenhagen makes clear, the design of a virtual enterprise for the purpose of learning consists of two steps (a) modeling the reality, and, (b) modeling models of reality from a didactic perspective. For the second step the designer should balance the information given and the questions asked. The instructional communication starts with the presentation of information, both the description of the features of the system or organization and their structure, the concepts and the illustration of the method how to manipulate the objects. It is the explanatory part of the instruction. This part is needed to coach the manipulation of real objects and the study of the meaning of the depictions. Worked examples of problems will be presented and discussed. These will be followed by real problems, which the students should solve by using the knowledge and the methods to reach the goals. In the next section a complex virtual environment is outlined. These environments can also be used by employees who start working at a company and for web-based education and training

(Perez, Gray, & Reynolds, 2006). The use of simulations and games for learning to operate in a complex company is discussed and shown in the next sections.

Simulations and Games for Learning to Model Complex Realities

The Construction of Knowledge

Simulations and games can be powerful tools to help develop concepts, principles and schemata of complex systems. They can have a role in education and training in putting learning into a context. Furthermore, they are environments in which students are invited to actively solve problems and thus construct their knowledge. Games and simulations provide students with a framework of rules and roles through which they can learn interactively through a live experience. They can tackle situations they might not be prepared to risk in reality and they can experiment with new ideas and strategies. They involve individual and group interpretations of given information, the capacity to suspend disbelief and a willingness to play with the components of a situation in making new patterns and generating new problems (Jacques, 1995).

Information and Guidance

Van Merriënboer (1997) stated that constructivistic and instructivistic approaches of instruction and learning need not be seen as distinct alternatives, but merely as two aspects of instruction that can, and often should, complement each other. In order to make the training process more efficient, it is sometimes necessary to provide the learners with pre-specified, general knowledge that may be helpful and offer guidance to solve the problems in a particular domain. Recently, Kirschner, Sweller and Clark (2006) also pointed to the fact that some form of guidance is needed in rich problem based experiential learning environments to prevent that learners miss essential information (see also Mayer, 2004), experience a cognitive overload, and are not able to construct adequate mental representations. De Jong and van Joolingen (1998) have argued that in simulation based inquiry environments learners often experience problems. They stated that cognitive scaffolds should be integrated into simulation based environments to support learners. Cognitive scaffolds may structure a task, take over parts of a task, or give hints and supporting information for the task. If this is true for simulations than this is also true for games, since games and simulations have a lot of elements in common.

Games and Simulations

Games are competitive, situated, interactive (learning-) environments based upon a set of rules and/or an underlying model, in which, under certain constraints and uncertain circumstances a challenging goal has to be reached. In games players (sometimes in cooperation with others) are actively solving challenging situated problems. Simulations are environments that are also based on a model of a (natural or artificial) system or process. In a simulation learners can change certain input variables and can observe what happens to the output variables. The main distinctions between games and simulations are that games contain elements of competition, chance, surprise, and fantasy that are not found in simulations. Furthermore, the goals are different. In simulations the goal is to discover the underlying principles of the simulation model, while in a game a person tries to win the game, get the highest score or beat the system or other players. In a simulation the learners have more freedom to act and experiment and in most cases they do not have to cope with limited resources. Finally, in a simulation it is relatively easy to recover from wrong choices. In games participants have to think about the tradeoff between costs and profits of actions and most often it is not possible to "undo" the actions. One has to face the consequences of one's actions, while in a simulation it is easy to restart and experiment in the same situation.

Support for Learning

One of the elements that could be added to the didactical context to support players is a debriefing activity. This is advocated by several authors (Garris, Ahlers, & Driskell, 2002; Klawe & Philips, 1995; Peters & Vissers, 2004). Debriefing supports reflective thought. Although there is consensus on the role of debriefing, in the literature about learning with games there are only a few studies that present data about the role of other supports. Stark, Graf, Renkl, Gruber, and Mandl (1995) focused on a problem solving scheme, Leutner (1993) on just-in-time information and advice, and Halttunen and Sormunen (2000) on feedback.

Leutner's study showed that different types of support could lead to different results. He found that advice (provided by means of warnings if decisions are likely to lead to problems) increased verbal domain knowledge, but decreased game performance. Furthermore, his data indicated that system-initiated adaptive advice had short-term effects (measured directly after game play), while learner requested non-adaptive background information had long-term effects (measured by a test that was administered a week after game play). This raises the question which combination of scaffolds is most powerful? To get a first clue about this, three studies were performed in which a game was used that contained a set of supports. The game and the scaffolds that are implemented are described below.

KM Quest: A Simulation Game about Knowledge Management

KM Quest is an Internet based simulation game about knowledge management. It was used to study the effectiveness of a combination of scaffolds (Leemkuil et al., 2003). Several universities and institutions for higher education in the Netherlands have used and still use the KM Quest learning environment in courses on knowledge management. The goal of the game is to learn basic knowledge management concepts and actions and the steps of a systematic approach to solve knowledge management problems. Furthermore, the goal is to learn to assess the KM situation of an organization and to advise/implement appropriate interventions.

In the simulation game KM Quest the player takes the role of a knowledge manager in a fictitious large product leadership organization named Coltec. The task of the player is to improve the efficacy of the company's knowledge household. More specifically, the goal of the game is to optimize the level of a set of general organizational effectiveness variables (or indicators): market share, profit, and the customer satisfaction index, by influencing the effectiveness and efficiency of knowledge management processes (knowledge gaining, development, retention, transfer and utilization). These processes can be influenced by choosing and implementing interventions from a pool of 57 possible interventions. The game is driven by an underlying simulation model that combines the organizational and knowledge management variables (see Shostak & de Hoog, 2004). Most of the indicators in the simulation model are characterized by a decay factor. This means that the value of the indicators decreases over time when no interventions are implemented.

In the game, players can use several resources while performing their task. They can inspect the status of business process indicators and knowledge process indicators that are incorporated in the simulation model, and they can inspect additional information about interventions, indicators etc. The implementation of interventions involves costs, as well as several other activities that the players can perform. Players receive a limited budget that they can use to implement interventions and buy information.

A three year period is simulated (divided into 12 quarters). Changes in the status of the business indicators are only computed at the end of each quarter. At the beginning of each quarter an (unexpected) event is introduced that could affect the knowledge household of the company. Players have to decide if and how they want to react on these events. Events are generated from a pool of 50 events. Different types of events can be distinguished based on two dimensions: the locus of the event (internal or external), and the effect of the event (direct, delayed, or no effect). Effects either can be positive or negative. For instance the following event will have a negative influence on market share: “Gluco has bought the company STIK, which has a strong position in industrial glues. It intends to expand the R&D department in STIK in order to strengthen its position in the Do-It-Yourself (DIY) household glues”.

Fig. 1. Virtual office interface.

There is no time limit to playing the simulation game. Players set their own pace. When players think they know enough to solve the problem, they indicate that they want to implement the proposed interventions. After the implementation the simulation game proceeds to the end of the quarter and the business simulation will calculate new values for each of the business indicators. The game ends after players have indicated that they have implemented the last intervention(s) in the fourth quarter of the third year in the life span of the company.

Players can interact with the environment by using tools and resources that are presented in an Internet environment, based on a “virtual office metaphor” (see figure 1). Clicking on a specific element in the “office” will open a window with additional resources or tools. To support the learners in performing their task and to support learning while playing the game several features have been implemented in the environment: a knowledge management model with shared worksheets, a help functionality, just-in-time background information, feedback, advice, visualization tools, and monitoring tools. The knowledge management model and process worksheets describe a systematic approach to solving KM problems and provide support by structuring the task and dividing it in phases and steps. Just-in-time background information supports players by giving access to domain and task relevant knowledge at any time needed. It is available by means of what and how files attached to the worksheets and by means of books (like the intervention and indicator handbook) that are placed at the bookshelves of the virtual office, the organigram (a link to static information about the company), and a help functionality in the task bar (green circle with question mark).

Feedback supports players in evaluating their actions. There are two types of feedback: dynamic feedback consisting of data generated by the simulation model, and pre-canned conceptual knowledge about knowledge management that is based on the experiences from KM experts and is coupled to certain events. The latter contains information (reference data) about the type of event, the knowledge domains and the knowledge processes that it is related to. Furthermore, it contains a list of interventions that are considered to be relevant to react upon this specific event. Players can compare their own interpretation of the event with the description given and can compare their actions with the suggested interventions.

Furthermore, the environment contains advice that supports players by giving warnings and hints. The advice is only available when certain values in the business model are below a fixed threshold value. The advisor icon in the status bar (a triangle with a! in it, see figure 1) normally is passive but starts blinking when advice is available. When the player clicks on this icon pre-canned text will be displayed that warns that there is a problem and that gives hints about what one can do about this problem by means of a reference list to suitable classes of interventions.

To help the players with interpreting the values of the large set of indicators and with seeing trends in the data, several types of visualizations are implemented like line or bar charts (available by means of the icons on the whiteboard). The last type of support consists of monitoring tools consisting of 12 quarterly reports that are available on the top two bookshelves (see figure 1). These reports give information about the players' actions and about data generated by the system in the quarters that are completed. This supports reflection by giving players the opportunity to go back in time without having the opportunity to reverse activities and/or actions that they have chosen.

Results with the KM Quest Simulation Game

The results of the experiments with the simulation game showed that the learning environment was effective (Leemkuil, 2006). The differences between post-test and pre-test scores were significant. The data show no relationship between game performance and post-test scores or knowledge increase. On the one hand this means that the learner does not have to be successful in the game to learn, which indicates that it is important to make a distinction between the goal of the game and the learning goal. In a game these two need not be the same. On the other hand this also means that in some cases students can be successful in the game while a learning test does not reveal an increase in knowledge.

Support Tools

In the game that was used in our studies several support tools were implemented. The assumption was that in learning environments like games all kinds of barriers to game play and learning could occur that would lead to ineffective learning, and to prevent this, cognitive scaffolds should be integrated that may structure a task, take over parts of a task, or give hints or supporting information. The data indicate that the tools with domain related background information, feedback with reference data and advice were frequently used by the players. The use of advice has a significant relationship with game performance as indicated by the level of a set of indicators in the business simulation model, but does not have any relationship with knowledge increase. There are some indications that the use of background information and process feedback have a positive effect on learn ing and transfer. This is in line with research with simulations (de Jong & van Joolingen, 1998) which also emphasizes the importance of direct access to domain information. This finding also emphasizes that it is important to make a distinction between scaffolds that support game play and scaffolds that support learning from the game.

Our studies focused on scaffolds that could be incorporated in the game itself. Probably the acquisition of new knowledge will profit from support that is not in the game itself but in the setting in which the game is used like a debriefing session after game play. In the past years several authors, like Dawes and Dumbleton (2001), Gee (2003), Kirriemuir and McFarlane (2004) and Jansz and Martens (2005), have stressed the importance of the social aspect of game play. When people think about children playing computer games the prevailing image is that of a boy sitting alone behind a computer screen. This image is too short-sighted because in many (internet) games players play together with others and furthermore after game play much discussion is going on with others about the game experiences and (during or after game play) knowledge and strategies are exchanged between players. Kirriemuir and McFarlane (2004, p. 27) stated that there are indications that interaction in (online) communities could contribute significantly to learning related to games play.

Thus, supports that focus on the social aspect of learning like collaboration in teams, classroom discussions during the period the game is played and a debriefing session after the game has ended could be powerful. These supports can foster a reflective strategy during the game and reflection after the game is played because players have to make their ideas explicit to be able to discuss with others and to exchange experiences. Furthermore, such supports make it possible to compare strategies and their results, to discuss the role of good or bad luck and could enhance the transfer of knowledge gained while playing the game to "real" life.

Concluding Remarks

During the second half of the former century instructional design became established as an applied science. The design knowledge and rules have firm grounds in the psychology of motivation, cognition and learning, in systems theory, in information and communication technologies and in epistemology. However the design knowledge and rules did not lead to uniformity of instructional designs. Of course, the results of a design in whatever field differ, because of the designer's creativity. Teaching has its own rules, but how these are applied is an art. For the design of instructions the design knowledge and rules have led to both general models and to highly specific situated electronic simulations and games. This chapter illustrates how this could happen. The features of the general models showed much overlap. They finally became combined in one generic model, labeled the ADDIE model. It is a general heuristic for the design of instruction. The model is used frequently and strongly supports the designers. They can fall back on it, which means that all the necessary steps to reach the goal will be taken. The use of the model is no guarantee that the instruction will be successful. The main shortcoming of the model is its lack of prescriptions how to design the mediated content of the instructional communication. Both the information and the representation of the object (s) involved and what the students must do with these. The acquisition of domain knowledge and skills to operate on and with objects requires that the students have to manipulate the objects for answering questions about their features and their change. Only then a mental model of a part of the reality can develop, both from the domain involved and from the field in which the student works or will work.

As is shown in this chapter the information and communication technology made it possible to simulate complex industrial environments. This type of instructional environments allows the participants to study the effects of their manipulations on several organizational variables. The simulation provides the participants a rich environment to develop a mental model of the organization and knowledge about the results of an intervention, without risk for damage. For the development of a mental model of the organization the simulation has two special advantages. It can be used in and outside the real organization and with or without peers. The first advantage means that students who are still at school for vocational training can develop a mental model and those who are still working can study what will be the effect of an intervention. The results are promising.

Though the developments shown did solve difficult instructional problems a few unsolved problems on cognition and learning still result in criticism of the instructional designs. The first problem concerns the relationship between the information given (the explanation) to the students and the amount of time to be spend to the students' own problem-solving activity. Mostly both are necessary. The second activity is crucial for firmly embedding the knowledge into the students' cognitive structures. The second problem concerns the development of and embedding of abstract concepts, principles and procedures from different contexts in the learner's cognitive structure. Learning from one situation does insufficiently foster the development of abstraction. Those concepts that are firmly rooted in cognitive structure are supposed to be well applicable. Both problems need further study.

References

Achtenhagen, F. (2004). Curriculum development as modeling of complex reality. In N. M. Seel & Dijkstra, S. (Eds.). Curriculum, plans and Processes in Instructional Design. (pp. 193-210). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Brown, J. S., Collins, S., & Duguid, D. (1989). Situated cognition and the culture of learning. Educational Researcher, 18, 32-42.

Bruner, J. S., Goodnow, J. J., & Austin, G. P. (1956). A study of thinking. New York: Wiley.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. Journal of verbal learning and verbal behavior, 8, 240-247.

Collins, A., & Stevens, A. L. (1983). A cognitive theory of inquiry teaching. In C. M. Reigeluth (Ed.), Instructional-design theories and models: an overview of their current status. (pp. 247-278). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Dawes, L., & Dumbleton, T. (2001). Computer Games in Education Project. British Educational Communications and Technology Agency. Electronically available at: http://www.becta.org.uk/page_documents/research/cge/report.pdf

Dick, W., & Carey, L. M. (1978). The systematic design of instruction. Glenview, IL: Scott, Foresman.

Dijkstra, S. (1997). The integration of instructional systems design models and constructivistic design principles. Instructional Science, 25, 1-13.

Dijkstra, S. (2000). Epistemology, Psychology of Learning and Instructional Design. In M. Spector & T. Anderson (Eds). Holistic and Integrative Perspectives on Learning, Instructional Design and Technology. Understanding complexity. (pp. 213-232). Dordrecht: Kluwea Academic Publishers.

Dijkstra, S. (2004a). Theoretical foundations of learning and instruction and innovations of instructional design and technology. In N. M. Seel & S. Dijkstra (Eds*). Curriculum,plans and processes in instructional design.* (pp. 17-24). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Dijkstra, S. (2004b). The integration of curriculum design, instructional design and media choice. In N. M. Seel & Dijkstra, S. (Eds.). *Curriculum, plans and Processes in Instructional Design.* (pp. 131-144). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Gagné, R. M. (1965). *The conditions of learning* (1st ed.). New York: Holt, Rinehart, and Winston.

Gagné, R. M., & Briggs, L. J. (1974). *Principles of instructional design.* (1st ed.). New York: Holt, Rinehart, and Winston.

Gagné, R. M., & Briggs, L. J. (1979). *Principles of instructional design.* (2nd ed.). New York: Holt, Rinehart, and Winston.

Gagné, R. M., Mayor, J. R., Garstens, H. L., & Paradise, N. E. (1962). Factors in acquiring knowledge of a mathematical task. *Psychological Monographs, 76,* 7 (Whole No. 526).

Garris, R., Ahlers R., Driskell J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming: An international Journal, 33,* 441-467.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.

Glaser, R. (1964). Components of the instructional process. In J. P. deCecco (Ed.). *Educational technology*. (pp. 68-76). New York: Holt, Rinehart & Winston.

Glaserfeld, E. von (1996). *Radical constructivism*. London: The Falmer Press.

Gordon, J., & Zemke, R. (2000). The attack on ISD. *Training Magazine, 37*(4), 42-53.

Halttunen, K., & Sormunen, E. (2000). Learning information retrieval through an educational game. *Education for Information 18*, 289-311.

Jacques, D. (1995). Games, simulations and case studies – a review. In: D. Saunders (Ed.), *The simulation and gaming yearbook. Games and simulations for business* (Vol. 3.). London: Kogan Page.

Jansz, J., & Martens, L. (2005). Gaming at a LAN event: the social context of playing games. *New media & society, 7,* 3, 333-355.

Jonassen, D. H. (1992). Objectivism versus Constructivism: Do We Need a New Philosophical Paradigm. *Educational Technology: Research and Development, 39,* 5-14.

Jong, T. de, & Joolingen, W. R van (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68,* 179-202.

Kirriemuir, J., & McFarlane, A. (2004). Literature Review in Games and Learning. *Nesta futurelab series. Report 8.*

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41,* 75-86.

Klawe, M. M., & Phillips, E. (1995). A classroom study: Electronic games engage children as researchers. *Proceedings of Computer Support for Collaborative Learning '95 (CSCL),* Bloomington, Indiana.

Lakoff, G. (1987). *Women, fire and dangerous things.* Chicago: The University of Chicago Press.

Leemkuil, H., de Jong, T., de Hoog, R., & Christoph, N. (2003). KM Quest: a collaborative internet-based simulation game. *Simulation & Gaming: An international Journal, 34,* 1, 89-111.

Leemkuil, H. (2006). *Is it all in the game? Learner support in an educational knowledge management simulation game.* Enschede: University of Twente, Faculty of behavioural sciences.

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

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist, 59*, 14-19.

Melton, A. W. (1964). *Categories of human learning*. New York: Academic Press.

Merrill, M. D. (1983). Component display theory. In Reigeluth, C. M. (Ed.). *Instructional-design theories and models: an overview of their current status.* (pp. 279-334) Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Merriënboer, J.J.G. van. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training.* Englewood Cliffs, NJ: Educational Technology Publications.

Molenda, M. (2003). In search of the elusive ADDIE model. *Performance Improvement, 42*, 5, 34-36.

Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.). *Mental models*. (pp. 7-14). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Perez, R. S., Gray, W., & Reynolds, T. (2006). Virtual reality and simulators: implications for web-based education and training. In H. F. O'Neill and R. S. Perez (Eds.) *Web-based learning. Theory, research and practice.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Peters, V. A. M., & Vissers, G. A. N., (2004). A simple classification model for debriefing simulation games. *Simulation & Gaming: An international Journal, 35*, 1, 70-84.

Piaget, J. (1937). *La construction du réel chez l'enfant*. [The child's construction of the reality.] Neuchâtel: Delachaux et Niestlé.

Reigeluth, C. M. (1983). The elaboration theory of instruction. In Reigeluth, C. M. (Ed.). (1983). *Instructional-design theories and models: an overview of their current status.* (pp. 335-382). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Reigeluth, C. M. (Ed.). (1983). *Instructional-design theories and models: an overview of their current status*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Seel, N. M. (2004). Model-centered learning environments: theory, instructional design and effects. In In N. M. Seel & S. Dijkstra (Eds). *Curriculum, plans and processes in instructional design.* (pp. 49-74). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Seel, N. M., & Dijkstra, S. (2004a). Introduction: instructional design and curriculum development. In N. M. Seel & S. Dijkstra (Eds). *Curriculum, plans and processes in instructional design.* (pp. 1-17). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Seel, N. M., & Dijkstra, S. (2004b). Delivery systems, information and communication technology, and media in education. In N. M. Seel & S. Dijkstra (Eds). *Curriculum, plans and processes in instructional design.* (pp. 257-269). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Seel, N.M., Ifenthaler, D., & Pirnay-Dummer, P.N. (2007, in press). Mental models and problem solving: Technological solutions for measurement and assessment of the development of expertise. In P. Blumschein, D. Jonassen, D. Hung, D., & J. Strobel (Eds.), *Learning by modelling systems.* Rotterdam, The Netherlands: Sense Publishers.

Seel, N. M., & Winn, W. D. (1997). Research on media and learning: distributed cognition and semiotics. In Tennyson, R. D., Schott, F., Seel, N., & Dijkstra, S. (Eds), *Instructional design: international perspectives* (Vol. 1). 293-326. Mahwah, NJ: Lawrence Erlbaum.

Skinner, B. F. (1954). *The science of learning and the art of teaching.* Harvard Educational Review, 24 (2), 86-97.

Skinner, B. F. (1958). Teaching machines. *Science, 128,* 969-977.

Stark, R., Graf, M., Renkl, A., Gruber, H., & Mandl, H. (1995). Förderung von Handlungskompetenz durch geleitetes Problemlösen und multiple Lernkontexte (Fostering action competence through guided problem-solving and multiple learning contexts). *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 27*, 289-312.

Tripp, S., & Bichelmeyer, B. (1990). Rapid prototyping: An alternative instructional design strategy. *Educational Technology Research & Development, 38*(1), 31-44.

11. Bridging a Gap between Psychology and Instructional Practice

Toward applied Model-Based Theory Construction

Andrei I. Podolskij

Moscow State University

Abstract: One of the most important and sharply discussed aspects of scientific knowledge significance is the problem of possibilities for practical applications and results. The application of psychological knowledge in different types and cases of schooling, training, and instruction is a bright illustration of that problem's current state. The aims of this paper are: 1) to consider possibilities and difficulties of such application, 2) to analyze the reasons for both success and failure, and 3) to try to work out a main path toward the construction of an applied theory to bridge a gap between psychological theory (in particular, learning and developmental psychology) and instructional practice. The paper considers the problems of practical applications of the fundamental psychological theory of Planned Stage-By-Stage Formation of Mental Actions, or the PSFMA Theory, by P. Galperin as the target case.

Keywords: Psychological knowledge; practical application; formation of mental actions; mental models; internalization; applied model-based theory.

Bridging a Gap between Psychological Theory and Instructional Practice: What Type of Psychology is needed?

It is perhaps considered trivial to declare the links between psychology on the one hand and educational and instructional sciences on the other hand. However, it is hardly trivial to discuss specific functional interrelations between them. Considering instructional practice as a "consumer" of psychological knowledge, we might ask the following question: What type of psychology does that practice require? This problem is really very important for the theory and practice of contemporary Instructional Design (ID). Any instructional designer or practitioner is not in need of general speculations but rather requires concrete information on the psychology of learning, development, and instruction which can become the core his or her practical activity. Contemporary instructional practice (and modern Instructional Design research as well) has been creating a real challenge for psychology. This challenge consists of providing knowledge which is sensitive to the heterogeneity and complexity of the social context in which learning processes take place, while at the same time offering the practitioners sufficiently concrete and clear psychological descriptions both of students and of the learning/teaching processes and content.

One might say that both descriptions of human learning and prescriptions concerning the planning and conducting of teaching or instruction have recently begun to come closer to real-life requirements than did the previous behaviorist descriptions. At the same time, despite high public expectations concerning real, practical applications of the theories and models identified by instructional designers, the actual results seem to be far from satisfactory. Ely (1990) has described this situation specifically with a focus on instructional design for schooling (IDS). Ely notes that researchers wonder why teachers and school administrators do not want to use the results of their remarkable and sometimes outstanding ideas, theories, and models. However, teachers and administrators (together with public opinion) wonder why the researchers are not capable of providing them with practical, useful knowledge, expressed in an acceptable and understandable form allowing it to be applied in everyday schooling activities. Twenty years ago, Snelbecker (1987) published a "menu" of teachers' arguments justifying why they did not use ID prescriptions in their everyday professional activity. In evaluating the possibilities of the Instructional Design "blueprints", several teachers claimed that they were already practicing what was recommended by IDS on their own. Other teachers, while acknowledging the innovative nature of the ID recommendations, still doubted the practical possibility of applying the recommendations in their own classrooms. They made statements such as: "I don't need any help in teaching/training", "I am already doing what you advise", "If I use that theory I'll have to change my teaching methods completely", "I already know those theories". These and similar statements were found to be the most common among the teachers (Snelbecker, 1987).

Modern cognitively based instructional design models and theories utilize concepts such as "self-directed learning", "student self-regulation", "meaningful content", and "cognitive and metacognitive strategies". The concept "mental model" seems to be one of the central ones in this list, attracting the attention of both "producers" of scientific knowledge and its "customers" to "the conceptual and operational representations that humans develop while interacting with complex systems" (Jonassen, 1995, p. 186), or "internal scale-model representation of an external reality" (Markham, 1999, p. 24). It is emphasized that individuals form mental representations by organizing symbols of experience or thought in such a way that they effect a systematic representation of this experience or thought as a means of understanding it or of explaining it to others (Seel, 1991) Learning occurs when people actively construct meaningful mental representations from presented information, such as coherent mental models that represent and communicate subjective experiences, ideas, thoughts, and feelings (cf. Mayer et al., 1999). There is no doubt that the concept "mental model" plays a very valuable and heuristic role not only for the construction of a general theory of human learning but also for working out an applied psycho-educational theory.

It is hardly difficult to guess that it is not very simple to get practitioners to accept a system of conditions for any scientific knowledge. Creating such conditions would mean giving teachers, trainers, instructors, et al. a possibility to do more explanation and practice with that knowledge than on the basis of common sense or their own practical experience alone. It means finding out a general intellectual procedure that would not only enable the users to analyze pluralities of concrete instructional situations in terms of modern psychology (in particular, learning and developmental psychology) but would also attract them to do it. In other words, it is necessary to offer a sort of "intellectual tool" which practitioners could use to enrich their competence in instructional technologies. This instrument has to be multifunctional and universal. It should direct the attention of the users to the changes and development in the constructive activity of a learner and focus the user's attention on the mental, internal components of any learning activity. On the procedural (technological) level, such an approach must operationalize sufficiently so as not to be simply a set of speculative declarations or "good intentions". Thus, applied psycho-educational theory requires a strict and simultaneously more explicit form of psychological knowledge. That is, it requires non-metaphoric descriptions of the variables (structural, functional and developmental) that are most essential and that determine the effectiveness and efficiency of learning/teaching processes as well as the description of the interrelations of those variables. In addition, a detailed and, again, non-metaphoric and unambiguous description of the psychologically grounded conditions that should be present within schooling (training, instruction, etc.) environments should be offered. Such descriptions must encompass the whole of the schooling situation and the complexity of the processes and phenomena involved.

It is important to emphasize that these descriptions must also be developmentally sensitive. Two different mechanisms may underlie a lack or even an absence of an ability to act in a mental plan: 1) macro-genetically, a learner's mental plan may be underdeveloped (Galperin, 1992; Piaget, 1970), hence preventing him/her from acting mentally within specific spheres of reality; 2) micro-genetically, the mental actions that are the prerequisites for learning specific content may not have been formed at all (or may have been formed with inappropriate and insufficient properties) in the course of a student's past experience (Galperin, 1969). The developmental dimensions of instructional content are equally clear. For example, it is generally not possible to assimilate certain subject areas before a certain, identified age or developmental point (Piaget, 1970). However, it is possible to overcome such age-related barriers (Galperin, 1992) when a teacher promotes the special formation of a student's mental activity on the basis of functional development regularities.

In discussing the "developmental sensitivity" of modern ID descriptions, one has to distinguish two different aspects: First, the developmental dimension must, as an essential and necessary component of the ID-knowledge base, be taken into account in developing plans for instruction. This requires: (a) planning, designing, organizing the learning/teaching processes in accordance with macro- and micro-developmental regularities, and (b) determining the short- and long-term developmental consequences of these processes and the extent to which learning/teaching processes influence the student's cognitive, personal, moral, social, and emotional development. Second, developmental changes can also be viewed as a direct and immediate aim of the learning/teaching processes. This principle has been formulated in a very general, philosophical manner by Vygotsky as: "instruction is good only when it proceeds ahead of development" (Vygotsky, 1978, p. 132).

"Planned Stage-by-Stage Formation of Mental Actions" Approach

One example of an approach in which the above requirements for a general intellectual tool for ID are met, and met in a sufficiently complete, sophisticated, and operationalized manner, is the "Planned Stage-by-Stage Formation of Mental Actions" approach introduced by Piotr Galperin (1969, 1989, 1992). Throwing a glance superficially, one may say that Galperin's approach resembles the main topic of the recent book entitled Understanding Models for Learning and Instruction. However, this is only the case if one does not consider the approach very profoundly. Indeed, one of the fundamental concepts of the approach is the concept of a mental image or mental representation whose content mostly overlaps with the content of the concept "mental model" as the central theme of the book and the essential part of Norbert Seel's creative work. It's important to add that Galperin's approach, especially in its modernized form (which might be called the "neo-Galperin approach"), represents very important specifications and concretizations concerning both the psychological typology of mental images, models, and representations and the process of the subject's acquisition of one or another new mental item.

Galperin's approach is the continuation of a trend in developmental and learning psychology which was started by Vygotsky (1978). However, Galperin's approach introduces the following new elements: (a) the approach considers the nature of human mental life, its coming into existence, and its further development in the context of philogenetic, anthropogenetic, and ontogenetic processes; and,

(b) it considers the system of psychological conditions which enables the formation of mental actions, images, and representations with the desired and prescribed outcomes. According to Galperin, mental action is a functional structure that is continually being formed throughout an individual's lifetime. Using mental actions and mental images and representations, a human being plans, regulates, and controls his/her performances by means of socially established patterns, standards, and evaluations. Mental action can and should be considered as the result of a complex, multimodal transformation of initially external processes performed by means of certain tools. In other words, from a nomothetical point of view concrete mental actions, images, and representations are the results of the internalization of external processes (Galperin, 1989).

Mental actions and images reflect and are the product of both human needs and the demands and conditions of the objective situation. They can therefore be characterized by a set of primary and secondary properties. The following properties are considered to be primary: (a) the composition of the action's objective content; (b) the extent of differentiation of essential elements of a problem situation from its non-essential elements; (c) the degree of internalization of the action; and (d) "energetic" (speed and enforcement) parameters. Secondary properties are: (a) reasonability; (b) generalization; (c) consciousness; and (d) criticism. The secondary properties are the result of specific combinations of primary properties. Both primary and secondary properties represent socially estimated and evaluated qualities of human activities and refer to any sort of activity, whether individual or collective, material or mental.

The final values of these properties determine the specific action and/or image that is formed. Galperin considered the values of the properties to be the direct outcomes of action formation conditions. He therefore defined a system of conditions that ensure and guarantee the achievement of prescribed, desired properties of action and image; this system is termed the "System of planned, stage-by-stage formation of mental actions" or the PSFMA system.

The system of "Planned Stage-by-Stage Formation of Mental Actions:" Subsystems Included

The PSFMA system includes four subsystems: (1) the conditions that ensure adequate motivation for the subject's mastering of the action; (2) the conditions that provide the formation of the necessary orientation base of action; (3) the conditions that support the consecutive transformations of the intermediate forms of action (materialized, verbal) and the final, end-transformation into the mental plan; and, (4) the conditions for cultivating, or "refining through practice", the desired properties of an action (Galperin, 1989). Each subsystem contains a detailed description of related psychological conditions, which include the motivation and operational areas of human activity.

The first subsystem (conditions for motivation) makes explicit a number of links and connections between learning motivation and the dynamics of the internalization processes.

The second subsystem (conditions for orientation) contains a description of hierarchically organized components which offers a framework for the formation of a concrete action and provides a learner with the conditions for an adequate ("complete" according to Galperin) orientation within a problem situation. These components are the representations of the subjective and objective characteristics of a problem situation and taken together were termed by Galperin the "complete orientation base of action" (Galperin, 1992). The structure of complete orientation base of mental action contains: 1. Representation of the final product of an action; 2. Representation of intermediate products; 3. Representation of the general plan for achievement of the final product; 4. Representation of plans for achievement of the intermediate products; 5. Representation of tools used to achieve those products (both orientation & execution tools); 6. Representation of the plan and tools for control and correction of actions as they are being executed; 7. Representation of the entire structure of a complete orientation base of action (Galperin, 1992). It is hardly difficult to discover a familiarity of the representations forming an orientation base and functional descriptions of mental models of different types. One may add that three psychologically different but interconnected levels of orientation base may be distinguished: 1) executive orientation base, which is a scheme of human orientation in how to do; 2) goal orientation base, which is a scheme of human orientation in what to do; 3) sense orientation base, which is a scheme of human orientation in why (for what) to do. There are both ascending and descending affections between these levels of orientation base Human understanding in how to do, for instance, reflects the upper level sense and goal representations, and vise versa the former are affected by possibilities and features of execution or sense and goal coming into existence (Podolskij, 2003).

The third subsystem represents the stages of internalization or transformation of the action into a mental plan. Galperin introduced six stages of internalization as the fundamental base of any learning process:

1. Formation of a motivation base of action;
2. Formation of an orientation base of action;
3. Formation of the material (materialized) form of action;
4. Formation of the external socialized verbal form of action (overt speech);
5. Formation of the internal verbal form of action (covert speech);
6. Formation of the mental action; final changes, the action's automatization and simultaneouzation Galperin, 1992).

The last, and fourth, subsystem contains a description of the three base problem situation types and of their combination and presentation during the formation processes. Three basic types are distinguished: (a) the "psychological" type, in which the conceptual and perceptual or visible features of a problem situation are opposed; (b) the "logical" type, in which necessary and essential parameters are contrasted with unnecessary or "noisy" parameters of a problem situation, and (c) "object" type, in which all of the possible forms of a specific action object content are varied. Different problem types are offered in a sequence which is meaningful for learners (Galperin, 1989).

If the four subsystems work together harmoniously, they produce an action with the desired primary and secondary properties.

To describe the frames of this approach generally in terms accepted by contemporary instructional design scholars, one may say that: (a) instructional content is presented as a set or a system of interconnected actions, concepts, and representations planned to be formed; (b) goals of instruction are defined and specified in terms of action parameters; (c) instructional plans are elaborated as didactic projections of stages of formation; (d) learners' characteristics are first considered in terms of the students' motivational and cognitive readiness to acquire projected mental actions and concepts.

The system of "Planned Stage-by-Stage Formation of Mental Actions:" General Procedure, Stages

The procedure of planned stage-by-stage formation of mental actions (PSFMA) (Galperin, 1992) can be presented in the most general form as follows: At the first stage, the subject's initial attitudes toward the goals and objectives of the forthcoming process as well as toward the concrete learning-teaching situation are constituted. These attitudes may be changed during the formation process. At the second stage, the scheme of orienting, or the scheme of orientation base of action, is elaborated. Guided by the scheme, a subject constructs, explores, reflects, and performs the action being formed. The extent of autonomy of the subject to construct such a scheme may vary from full dependence on a teacher to an almost full independence; it is the function of the aims and goals of the concrete learning-teaching process and of the learner's characteristics. For instance, the younger the learners are the more necessary it is to present an orienting scheme in a guided form (as a rule).

The general macrostructure of this scheme is relatively indifferent to the features of the special domain content of the action and to the level of expertise of the actor. Essential differences may be found if one compares concrete specifications of each element of orientation schemes in the actions of beginners and experts; of disabled, ordinary, and gifted children, and so on. The macrostructure is also relatively indifferent to kinds and sorts of actions being formed, for example if one deals with concrete specific domain actions, actions that belong to the cognitive meta-strategies, actions that underlie the heuristic methods, etc. The general function of the scheme is to provide the learner with a powerful orientation means or tool which enables him/her to plan, to direct, and to control the solving of different kinds of problems related to the field involved. It should be emphasized that in general such a scheme is not an "algorithm" for solution (although in some cases and under definite conditions, there are several kinds of "algorithmic prescription"; but this is an exception rather than the rule). This scheme is a learner' tool for his/her orientation in both the objective content of action and in the operations needed to handle this content in accordance with concrete learning aims and goals. The construction of an orientation base is a real creative task for the participants of the learning/teaching interaction. Furthermore, it is necessary to stress that this scheme plays the role of a synchronizator for the development of knowledge and skills (see Dijkstra, 1997) related to the content of the action. The scheme of orientation base contains the necessary and essential information base both for learner's analysis of the objective content of the action and for the application of this content to the definite problem situation. In other words, it has a function very close to the most general function of mental models.

At the third stage, the learner starts to solve different problem tasks, organized and presented in the definite sequence and manner (see the fourth subsystem above), using the scheme of the orientation base of action elaborated at the previous stage. The form of the scheme may vary from detailed descriptions of an order and a content of operations to be executed to very general hints and heuristics. As for the external view of the scheme, all kinds of representations are possible: A scheme of the orientation base may be represented as an arrow-scheme, a flow-diagram, a "solution tree," a text, a picture, a graph, a formula, etc., presented either as a whole, or part by part, or hierarchically. The representation is dependent on the three variables mentioned above: the objective content of the action, the learning goals, and the learner's characteristics. The constancy of the action's essential general macrostructure, enforced by verbally reasoned solving of the sequence of specially designed problem types, leads to the point that it is no longer necessary for the student to use the scheme of the orientation base as a material (materialized) learning aid. At that time its main content (see earlier—the second subsystem) is fully represented in the subject's socialized speech (socialized means understandable for other persons). This socialized speech becomes the base for the new action to be formed.

With this step, the action moves into the fourth stage of formation—the level of overt, socialized speech. Once the set of varying problem situations has been solved, the so called "melting" of the external phonetic form of speech takes place.

The main content of the intermediate fifth stage of action formation is the formation of the action-internal verbal mode (covert speech level).

At the last, sixth stage of formation, the mental action passes through final changes, which are the result of simultaneouzation and automatization. The new mental action begins its own "psychological life." It is able either to be included in other psychological structures enriching them or to subsume other psychological structures to be enriched and developed.

Thus, as a result of a stage-by-stage formation an externally mediated and successive action appears to be transformed into a "pure mental act": after estimating the problem situation a learner makes a decision on the spot. The results of planned stage-by-stage formation closely correspond to the most desirable aims of contemporary Instructional Design: Acquisition of generalized, meaningful, synchronized knowledge and cognitive skills is a result of authentic student learning activity transformations.

The System of "Planned Stage-by-Stage Formation of Mental Actions": Success and Failure of Implementation

Since the late fifties, a significant number of authors (both researchers and practitioners) have tried to use Galperin's theory to improve schooling processes and results. Studies concerned the very different kinds and types of schools (primary, secondary, vocational, special schools). Subjects (learners) were ordinary, disabled, and gifted children of different ages (from 5 to 18). Specific domains were also very different: writing and arithmetic, native and foreign languages, math, scientific and humanitarian disciplines, drawing, music, physical training. At last, psychologically heterogeneous structures were the objects of planned stage by stage formation: separate specific-domain mental actions and, connected with them, concepts and representations; groups and systems of actions and concepts; actions which underlie cognitive as well as metacognitive strategies and heuristics (Galperin, 1992; Podolskij, 1997; Talyzina, 1982).

Looking back at the more than fifty year history of Galperin's approach, one may note that the sixties, seventies, and early eighties were periods of the great optimism concerning the effectiveness and efficiency of its practical application. As has been convincingly demonstrated by hundreds of experimental and applied studies, the whole set of main objectives that any schooling is aimed at could be fulfilled: (a) the guaranteed acquisition of the educational course is ensured for practically all of the learners (all, of course, who have the necessary level of preliminary knowledge and skills) without necessitating the prolonging (sometimes even with reducing) of the time allocated to it, and practically without any additional costs; (b) the division into the acquisition of knowledge and its application is minimized or even disappears; (c) the learners acquire abilities to transfer to a new situation: not only knowledge and skills being formed, but also the way of acquiring them; (d) the learners get more and more interested in the very processes of acquiring knowledge and in knowledge itself (P. Galperin; A. Podolskij; N. Talyzina et al.).

However, in comparing the 60s – 80s' and later publications one can easily discover a significant decrease in the wave of optimism concerning the PSFMA's application. Moreover, anyone who is familiar with the current school education situation can hardly discover the PSFMA's extensive real practical applications in contemporary schools or in schools of the nearest past. Of course, there have been and there still are a lot of interesting experiences in different parts of the world which demonstrate the successes, failures, and problems of PSFMA's practical usage; however, its range of use is rather limited.

Besides the obvious socioeconomic and sociopsychological reasons that motivate the implementation of any psycho-educational innovations, there is one more reason of a theoretical and methodological nature for using Galperin's approach. Historically, it has been established that in most psychological research performed along the lines of this approach, the substantial pedagogical results of planned stage by stage formation of mental actions first came to the fore. However the proponents' enthusiasm about really unusual and hopeful results had a reverse side: it led to serious misunderstandings concerning the status of Galperin's approach. Sometimes the approach has been interpreted not as a general description of laws and regularities which try to explain the concrete dynamics and results of human mental activity formation, but rather as a set of technologies and prescriptions on how to teach. Indeed, such an interpretation distorts reality and transforms the approach to "absolute" knowledge like a sort of "philosophers' stone".

One should not forget that any scientific research ("pure" scientific or practically oriented study) is always based on a system of accepted abstractions. A direct use of research methods created on the basis of such abstractions has a number of fundamental restrictions. Positive results achieved in experimental procedures will never be preserved in practical schooling if the circumstances which were abstracted from in the course of an experimental formation procedure come to the fore in the actual situation. This means that a very carefully done procedure of exact definition which follows a "zone of admitted and assumed abstraction" (ZAAA), that is developmental or/and learning variables or determinants which the researcher sees as being important, or at least as existing, but leaves out of consideration due to certain reasons) is needed (Podolskij, 2003). Another important prerequisite for applied psycho-educational theory construction is an objective analysis of truly successful and unsuccessful attempts at applying the PSFMA theory in different types and kinds of schooling and instruction. It is also important to realize a necessity to have at one's disposal an integrative knowledge to synthesize appropriate information from modern learning and developmental psychology, and also a systemic description of practical requirements presented to psychological theory by schooling practice.

The heterogeneous structure of a learner's orientation in the problem task, or in the instructional situation in general, in learner/teacher and learner/learner interrelations, and the non-linear character of an action's orientation formation preclude speaking about an application of any constant, or, so to speak, "absolute" PSFMA procedure (see Podolskij, 1993, 2003). The sequence of stages, the general structure of the orienting base of action, and other cornerstone elements of the planned stage-by-stage formation system mentioned above should be considered as the most complete, normative and, according to Galperin, the nomothetic description of human mental action formation process (Galperin, 1992; Podolskij, 1993).

Providing Successful Application of the PSFMA Approach in the Instructional Situation: Three-model Framework

In emphasizing the nomothetically orienting role of the general PSFMA system, the successful application of the PSFMA statements does not imply a literal reproduction of some abstract, extremely general procedure. Rather, it refers to the creative design of a system of necessary and sufficient psychological conditions for instruction. The elaboration of such a procedure occupies an intermediate position between fundamental psychological knowledge and the real process of schooling, instructing, or training (Podolskij, 1993, 1997). This intermediate position is operationalized in the consecutive elaboration of three models of the instructional situation. These are the psychological, the psychological-pedagogical, and the methodical, or technological model (Podolskij, 1993, 2003).

The psychological model includes: (1) a description of the knowledge and skills to be acquired in terms of the learner's mental actions, images and concepts; (2) a description of the macro- and microstructure of the multi-level learner's orientation as the basis for a new mental action, concept, or image to be formed; (3) a description of age-related and individual characteristics of students that are relevant to instruction and schooling; and, (4) a description of the specific system of psychological conditions needed for the formation of the planned action. It is clear that in different applications of the PSFMA system, application emphasis should be placed on different constituents of the psychological model.

The main function of the psychological-pedagogical model is that the psychological model demands projection onto the specific objective and subjective conditions of schooling and teaching. Such conditions include: instructional activities and organization and distribution of different organizational forms during a lesson or a sequence of lessons; a quantity of in-class and homework activities; and amount of individual, small group and whole class learning activities; use of available technical aids for teaching (CAL, for example). Rephrasing one famous saying one might declare that the psychological-pedagogical model represents the "art of the possible", that is to say that it reaches an optimal compromise between the strict requirements of the psychological model and the restrictions constructed by objective and subjective components of the reality. Sometimes it is necessary to reduce such strict requirements in favor of requirements implementation (at least part of them), and sometimes – by contrary – it appears to be real to overcome traditional learning environment resistance in favor of innovation implementation.

The last, procedural, or technological model of instructional situations includes a detailed description of the teaching process, distributed between units of definite form and time, with a precise description of the goal of each unit and the means to achieve it. It also includes a complete list of teaching documentation: schemes, different types of learning and assessment tasks, a description of the order in which technical aids should be applied, and a number of other materials specified for different types and kinds of schooling/instructional situations. The methodical model looks like the traditional well-done "teacher's lesson plan"; however, one has to remember that this model is based upon the consideration outlined by psychological and psychological-pedagogical models (Podolskij, 1993, 1997).

It is also necessary to consider the three-model framework as an intellectual tool, not just as an algorithm which prescribes a teacher "how to act". This framework, while used in an appropriate and sophisticated way, gives a teacher the ability to orient, plan, control him/herself completely, and correctly design, arrange, and carry out different instructional activities. In other words, this framework may provide us with an applied psycho-educational theory that occupies an intermediate position between fundamental psychological knowledge and educational/ instructional practice.

To conclude the recent chapter I might declare that to bridge a gap between psychological science and schooling (instructional) practice one needs to deal with two categories of mental models. First, one must take into account a hierarchical system of the student's mental models, which forms schemes of action orientation on different levels; those models come into existence and reach required features by means of application of the special procedure of mental action formation. Secondly, one must form a system of teacher's mental models, the contents of which are to be constituted by the three-model scheme of the instructional situation. Such a scheme may become a basis for applied model-based psycho-educational theory construction.

References

Dijkstra, S. (1997). Educational technology and media. In S. Dijkstra, N.M. Seel, F. Schott, & R.D. Tennyson (Eds.), *Instructional design: International perspectives. Volume II: Solving of instructional design problems* (pp. 137-144). Mahwah, NJ: Lawrence Erlbaum Associates.

Ely, D.P. (1990). Trends and issues in educational technology. *Educational Media and Technology Yearbook, 16*, 5-30.

Galperin, P.Ia. (1969). Stages in the development of mental acts. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary soviet psychology* (pp. 249-273). New York: Basic Books.

Galperin, P.Ia. (1989). Organization of mental activity and effectiveness of learning. *Soviet Psychology, 27*(3), 65-82.

Galperin, P.Ia. (1992). Stage by stage formation as a method of psychological investigation. *Journal of Russian and East European Psychology, 30*(4), 60-80.

Jonassen, D.H. (1995). Operationalizing mental models: Strategies for assessing mental models to support meaningful learning and design-supportive learning environments. Proceedings of the Computer Support for Collaborative Learning Conference (CSCL'95), 17-20 October. Bloomington, IN: Indiana University.

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

Mayer, R.E., Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communication by minimizing cognitive load. *Journal of Educational Psychology, 91*(4), 638-643.

Piaget, J. (1970). Piaget's theory. In: P.H. Mussen (Ed.), *Carmichael's manual of child psychology* (Vol. 1). New York: Wiley.

Podolskij, A.I. (1993). Psychological theory as base of instructional design and as part of curriculum in post graduate education. In C.Terlouw (Ed.) *Instructional development in higher education: Theory and practice*, (pp. 66-78). Amsterdam: Thesis Publishers.

Podolskij, A.I. (1997). Instructional design for schooling: developmental issues. In: S. Dijkstra et al. (Eds.), Instructional Design: International Perspectives, Vol. 2, (pp. 289-314). HillsLawrence Erlbaum.

Podolskij, A.I. (2003). On scientific status of the Gal'perin's psychological approach. *Hungarian Journal of Psychology, 4*, 78-96.

Seel, N.M. (1991). *Weltwissen und mentale Modelle*. Göttingen: Hogrefe.

Seel, N.M., & Schenk, K. (2003). Multimedia environments as cognitive tools for enhancing model-based learning and problem solving. An evaluation report. *Evaluation and Program Planning, 26*(2), 215-224.

Snelbecker, G.E. (1987). Instructional design skills for classroom teachers. *Journal of Instructional Development, 10*(4), 33-40.

Talyzina, N.F. (1982). *The psychology of learning*. Moscow: Progress Publishers.

Vygotsky, L.S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

12. Model-Based Learning Environments

From Sophisticated Intelligent Tutoring Systems to Simple Adaptive Learning Environments

Gerhard Weber

University of Education Freiburg, Germany

Abstract: The development of intelligent tutoring systems (ITS) was in the focus of computer-based learning during the last twenty years of the last century. But most ITS did not find the way from the laboratory to use in standard learning settings. During the last ten years, adaptive learning environments emerged, a promising alternative to ITS. Both, ITS and adaptive learning environments use some more or less sophisticated type of user modeling. At hand of the web-based learning environment ELM-ART that combines both an ITS for problem solving support as well as adaptive features for navigating through the learning materials it is shown how user modeling can be used to facilitate learning. Finally, NetCoach, a Web-server for presenting and authoring adaptive on-line learning environments will be introduced.

Keywords: Intelligent tutoring systems; adaptive learning environments; user modeling.

Introduction

Learning is the perhaps most important core competence in the modern knowledge society. That is one of the reasons why computer based learning is one of the fast growing areas, especially in on-line learning scenarios in the world-wide-web. Computer based learning offers the opportunity to learn on demand. That is learners don't have to wait for a traditional course on the topic to be learned and they can concentrate on just the information they need without bothering with all the other topics offered with courses in further education. Learning on the job is the second opportunity of computer-based leaning. Learners can learn in parallel to their job or at home without the need to leave the job for a day or more and don't have to move to another city to participate in a very time consuming course.

But these opportunities will only convince or show an advantage if learning results of computer-based learning are comparable to or even better than results of traditional courses in further education. The drawback of learning on demand and learning on the job settings may be the absence of a human teacher or tutor. In situations where a human teacher can help individually when a learner gets stuck in problem solving or in understanding an introductory text, a simple computer-based learning environment may not be able to offer the help needed to learn successfully. An answer to this problem was the development of sophisticated intelligent tutoring systems that are able to play the role of an individual human tutor and that can communicate with the learner during the learning process (Wenger, 1987). However, though a lot of successful intelligent tutoring systems have been developed during the last three decades, most of them were only tested and used in laboratory settings and only a few of them are used in schools or in further education, e.g., the mathematics tutors developed by the Carnegie Mellon University (CMU) (Corbett, Koedinger, & Hadley, 2001). A reason may be that the costs developing these sophisticated systems are too high compared to the advantage over other forms of learning support.

During the last ten years another type of advanced computer-based learning has emerged, the adaptive learning environments. Adaptive means that these learning environments are able to adapt in one or more aspects to particular needs of the learner. Both, intelligent tutoring systems and adaptive learning environments are based on some form of a user model, in this context called learner model. Such models consist of assumptions about the learning state of a learner. This information can be assessed in advance so that learners can be assigned to categories (e.g., novice, advanced, expert). Learners belonging to different categories will get different advice and different learning support. Or the system is able to accumulate information about the user during learning with the system and, therefore, will be able to adapt to particular problems during the learning process.

In this paper, we will show how both intelligent tutoring systems and adaptive learning environments are able to support the individual learning process on the basis of information gathered in a learner model. At hand of the sophisticated learning environment ELM-ART (Weber & Brusilovsky, 2001) we will describe a web-based learning environment combining an intelligent tutoring system that supports basic problem solving tasks (described in Section 2) with a simple adaptive learning environment that supports the instruction phase of learning basic information topics in a new domain (described in Section 3). Section 4 introduces the web-server NetCoach, a system that allows authors to easily create and edit adaptive web-based learning courses. ELM-ART was the first complex web-based learning environment implemented in NetCoach. The conclusion in Section 5 will summarize the main ideas of this chapter and give a brief glance of further directions in the development of adaptive learning systems.

Intelligent Tutoring Systems

Intelligent tutoring systems (ITS) played the most prominent role from the mid eighties until the end of the last century. Intelligent tutoring systems usually are complex systems combining different components needed to support learning. These are the domain knowledge, the learner model, the tutoring module, and a user interface. The domain knowledge describes the domain in terms of concepts (declarative knowledge) and rules (procedural knowledge). The learner model stores information about the user's learning state (in terms of a user profile or a detailed description of the concepts and rules the learner possesses of). The tutoring module guides the individual learning process based on information from learning goals, from the learner model and, especially, from the diagnosis of learners' answers to questions, quizzes, or problem solving tasks. The user interface, on the one side, allows the user to input solutions to tasks and, on the other side, allows the system to present tasks und feedback to the user. The "intelligent" aspects of these systems mainly stem from the user-modeling component, the learner model. Information stored with the learner model is used by the tutoring component to suit the learning support (hints, feedback) and the learning path to the individual learner.

A lot of sophisticated ITS have been developed during the last decades, the most important ones in more formal domains like mathematics, computer science and programming. Many successful intelligent tutoring systems stem from the group of J. Anderson. The first one was the tutoring system for the programming language LISP (Anderson, Conrad, & Corbett, 1989), lessons learned with this LISP-Tutor have been described in Anderson, Corbett, Koedinger, & Pelletier (1995). The geometry tutor (Anderson, Boyle, & Yost, 1986) was the first one of several mathematics tutors that have been further developed and are successfully used in schools today (Corbett et al., 2001). Corbett (2001) has shown that these cognitive tutors can be at least as effective as human tutors in single tutoring situations. All these tutoring systems are built upon the ACT* theory (Anderson, 1993), a rule-based cognitive architecture to model human cognitive skills.

The ELM Architecture

In this paper, I will present the episodic learner model (ELM, Weber, 1996a). It supports learning the programming language LISP. ELM is a hybrid learner modeling architecture using both rules (as the programming tutors from the CMU group, e.g., Anderson et al., (1989)) and information from cases. Figure 1 shows an overview of the ELM architecture. In ELM, knowledge about a learner is stored in terms of episodes in a learner model. Episodes from the learner model are used to explain learners' solutions to programming problems. This individual case base grows with the number of programming tasks the learner works at. In the beginning, the interpretation of the learner's code and the explanation of bugs are only based on information from the domain knowledge. Explanations of learners' solutions to programming problems are stored in an individual, dynamic learner model. Episodes from the learner model can be used to trigger the explanation of learners' solutions to further programming problems. Additionally, and perhaps more important, episodic information can be used as reminders to the learners of examples and of previous solutions.

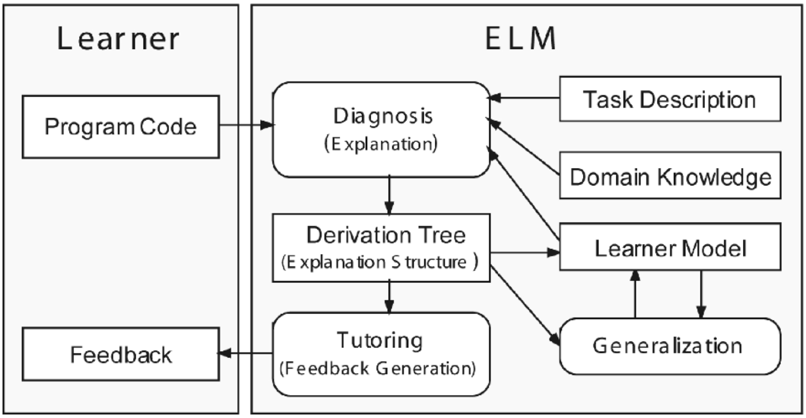

Fig. 1. The ELM architecture.

ELM works like an EBG method in two steps. In the first step, the code produced by a learner is analyzed by an automatic diagnostic component, comparable to the explanation step in EBG (Mitchell, Keller, & Kedar-Cabelli, 1986). This diagnosis results in a derivation tree (an explanation structure in terms of EBG). In the second step, information from the derivation tree is stored in the case base and generalized. The diagnostic step draws information from four sources: (a) a description of the task the learner is working on, (b) the domain knowledge, (c) the individual learner model, and (d) the program code produced by the learner (see Figure 1).

The task description is a high-level description of a programming problem. It consists of a plan addressing concepts and schemata from the knowledge base. Plans are hierarchically organized. They consist of calls to concepts from the knowledge base where arguments may in turn be sub-plans.

The knowledge base is represented in terms of hierarchically organized frames of concepts and rules. Concepts comprise schemata, higher-level programming concepts, and concrete LISP concepts. These frames consist of slots describing features of the concept, especially rules and transformations. The rules slot contains a list of rules that may be applicable to plans addressing this concept. The transformations slot contains a list of plan transformation that can transform a plan addressing this concept into a different, semantically equivalent plan. Rules are sorted by their internal priority. High, medium, and low priorities stand for so-called "good", "suboptimal," and "bug" rules, respectively. The set of bug rules in the knowledge base is comparable to a bug library in other systems (e.g., PROUST, a tutoring system for the programming language PASCAL (Johnson, 1986), or LISP-Tutor, an ITS for learning the programming language LISP (Anderson et al., 1989)). Bug rules reflect known bugs and misconceptions observed from students in previous LISP courses.

The learner model is built from episodic information about a particular learner. Concepts including plans and rules identified during the diagnostic process and captured in the derivation tree are integrated into the learner model as instances of corresponding concept frames of the knowledge base. Episodic instances in the learner model are generalized in the second step of the ELM algorithm.

The system accepts LISP code produced by a learner as a more or less complete solution to the problem description. In case the program is coded in a LISP structure editor, an incomplete solution may be submitted. In the current on-line version ELM-ART, only complete solutions to the programming problem can be submitted.

To construct the learner model, the code produced by a learner is analyzed in terms of the domain knowledge on the one hand and a task description on the other hand. This cognitive diagnosis results in a derivation tree of concepts and rules the learner might have used to solve the problem. These concepts and rules are instantiations of units from the knowledge base. The episodic learner model is made up of these instantiations. In ELM, only examples from the course materials are pre-analyzed and the resulting explanation structures are stored in the individual case-based learner model. Elements from the explanation structures are stored with respect to their corresponding concepts from the domain knowledge base, so cases are distributed in terms of instances of concepts. These individual cases are used for two different adaptation purposes. First, episodic instances can be used during further analyses as shortcuts if the partial code and plan match corresponding patterns in episodic instances (Weber, 1996a). Second, cases are used by the analogical component to show similar examples and problems for reminding purposes (Weber, 1996b).

The Diagnosis of Program Code

Let us explain the diagnosis of program code, the construction of the individual learner model and the impact of the learner model on further diagnoses at hand of an example from programming recursive functions. In list recursion tasks (typical for programming in LISP), the recursion usually stops when the recursion list is empty. That is, in the case decision of the recursive function one has to test whether the list to work at is empty. In the plan describing how to solve the programming task, the sub-plan for this terminating case is to test on whether the recursion list is empty. The sub-plan may look like this: *(NULLTEST (PARAMETER ?LIST))*. This plan addresses the concept NULLTEST, a concept describing how to test on the empty list. The plan has one argument, the coding of the parameter ?LIST that is bound to the corresponding parameter definition in the arguments list of the function definition.

The concept NULLTEST is a special case of a predicate with one argument and with this argument expected to be a list. The default rule for coding such a predicate is the *Unary-Func-Rule* that describes how to program a function with one argument by finding an appropriate function operator and coding the argument of this function. A slot of the concept NULLTEST describes that an appropriate operator would be a NULLTEST-OP. So, the default rule *Unary-Func-Rule* establishes two sub-plans, a plan *(NULLTEST-OP)* that addresses the concept NULLTEST-OP and a sub-plan *(PARAMETER ?LIST)* to code a parameter of the function definition (see Fig. 2a). The NULLTEST-OP can be coded with the LISP functions `NULL` or `ENDP`.

The definition of the concept NULLTEST contains two other, sub-optimal rules. The rule *Not-Instead-of-Null-Rule* describes how to write code for the plan with using the function NOT by calling the sub-plan *(NOT-OP)* (see Fig. 2b). The other sub-optimal rule *NIL-Comparison-Instead-of-Null-Rule* describes how to code the plan by directly comparing the parameter to the atom `NIL` with the function `EQUAL`. Sub-optimal means that coding the program this way will result in a correct solution but coding the program following the default rule (or possibly another correct rule) would be more convenient in LISP programming.

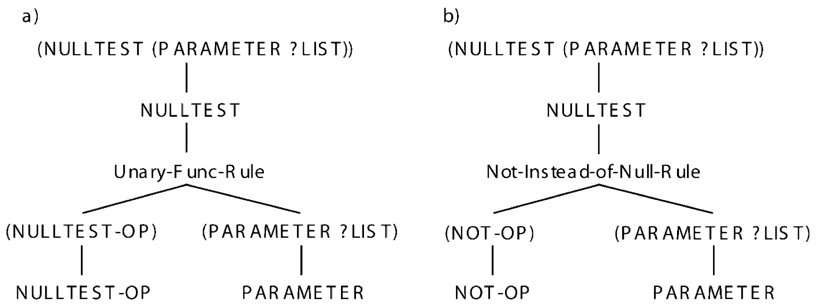

Fig. 2. Hierarchy of the plan *(NULLTEST (PARAMETER ? LIST))* with sub-plans for a) the default rule *Unary-Func-Rule* and b) the sub-optimal rule *Not-Instead-of-Null-Rule* of the concept NULLTEST.

In case the learner coded the function definition using a sub-optimal rule, the code will be accepted as a correct solution but he or she will receive the feedback that the program could be coded in a more convenient way. The learner can decide whether he or she will optimize the function definition. However, in case of an error at another place in the function code, the learner will get feedback only on the wrong part of the code and will get help even in the context of a sub-optimal rule. The main goal is to solve the programming task correctly; a sub-ordinate goal is to code the function more elegant or convenient. Additionally, the concept NULLTEST contains a transformation that describes how the current plan can be transformed into a plan that tries to solve the task by testing on the length zero of the recursion list.

Besides correct rules (e.g., the default rule to code a unary function) and sub-optimal rules (as the rules described above) the systems possesses of bug-rules that describe typical programming bugs. E.g., in an arithmetic task, a typical error may be to confuse operators (like + or *), to intermix the sequence of arguments in non-commutative functions, or to miss coding an argument.

With these rules and transformations a large room of possible solutions (correct or wrong) to the programming task is built up. The learner's solution to the programming task is not compared to a canonical correct solution as in many simple learning systems. Instead, a wide range of correct, sub-optimal, or wrong solutions can be identified by the diagnosis component of ELM. Rules contain text patterns for explaining where and why the code is sub-optimal or wrong. These text patterns are used by the tutorial component to give feedback to the learner. The rules are arranged according to their priority, with correct, sub-optimal, and bug rules assigned to highest, intermediate, and lowest priorities, accordingly.

In principle, the diagnostic component tries to solve the programming task in the same way as the learner. It starts with the plan stated in the task description with applying possible rules and transformations recursively. That is, it tries to produce the same code as the learner with preferring rules with the highest priority in case of competing different explanations leading to the same code. The diagnosis works like the first step of the explanation-based generalization (EBG) method (Mitchell et al., 1986). It results in a derivation tree containing all concepts and rules used to explain the learner's solution. Figures 3a and 3b show the derivation trees for two different program solutions `(null reclist)` and `(not reclist)`, accordingly for the sub-plan to code a NULLTEST with the recursion-list parameter of the function definition (in this case `reclist`).

The derivation tree is the basis for the tutorial component of the system to provide appropriate feedback to the learner. As mentioned above, the text patterns

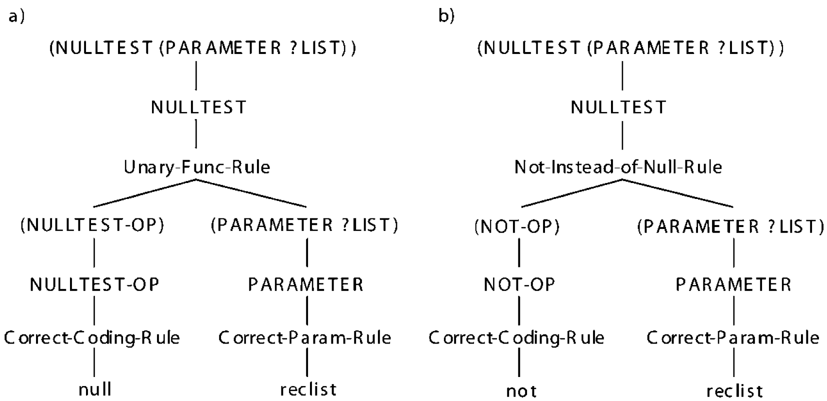

Fig. 3. Derivation trees of the solution codes a) (null reclist) and b) (not reclist) for the sub-plan (NULLTEST (PARAMETER ?LIST).

stored with the rules addressed in the derivation tree are used to generate the feedback text. In the first place, all bugs identified by bug rules are explained to the learner. In case all bugs have been diagnosed, possible sub-optimal solutions are explained to the learner. In case of correct solutions, only a short feedback is given without any additional explanation.

The Episodic Learner Model

The individual learner model is built up with information from the diagnostic process. Information captured in the derivation tree is integrated into the learner model in terms of instances of their respective concepts and rules of the knowledge base. Special slots in these instances refer to the context where these instances occurred (especially the current task). Other slots refer to transformations of concepts, and to argument bindings. The set of all instances (plus further generalizations) constitutes the episodic learner model. In subsequent analyses by the diagnostic component, these episodic instances can be used to trigger the selection of transformations and rules.

This form of episodic learner modeling shows a strong similarity to approaches in *case-based reasoning* (Kolodner, 1993; Schank, 1982) and *explanation-based learning* (Lebowitz, 1986). One single event (or case) is interpreted with respect to the knowledge base and the learner model. The result of this interpretation is integrated into the learner model. So, it is a form of single-case learning as in *explanation-based learning* programs (DeJong, 1988; Mitchell et al., 1986).

The derivation tree (representing an episode) is not stored as a whole but distributed over all concepts and rules mentioned in the derivation tree in terms of episodic frames. The set of episodic frames of a particular episode constitutes a case in the case library, the episodic learner model. An episode frame for each case indexes all episodic frames contributing to the episode. Starting from the episode frame, the complete derivation tree can be reconstructed by scanning all episodic frames and re-establishing the old plan contexts and rules. That is, in ELM, cases are not stored as a whole but are divided into *snippets* (Kolodner, 1993) that are stored with respect to the concepts in the system’s knowledge base.

Individual episodic information is used to accelerate the diagnostic process. Without episodic information, rules are applied sequentially according to their priority. That is, rules with the highest priority are tested first. With episodic information, rules that have been used to solve a plan in previous episodes successfully are tested first. Only if these rules fail, other rules are tested in the order of their priority.

Analogies and Remindings

The second advantage of an individual, episodic learner model stems from its potential to find analogies and remindings to examples from the learning materials and to solutions from previous programming episodes. For this purpose, a fast, explanation-based retrieval algorithm (EBR; Weber, 1996b) has been developed. The EBR method is based on the potential of episodic modeling to predict code that programmers will produce as solutions to new programming tasks (Weber, 1996a). These predictions can be used to probe the case base for examples and remindings that are similar to the expected solution and, therefore, are useful for solving the new task.

Searching for a best analog to a given new programming problem works in four steps. First, rules used by the diagnostic component to analyze program code can be used by a generation component to generate an expected solution to a new programming task (Weber, Bögelsack, & Wender, 1993). This results in an explanation structure of all concepts and rules used to generate a solution. Second, this predicted explanation structure is stored temporarily in the episodic learner model. Third, from computing the values of organizational similarity (Wolstencroft, 1989) to other episodes (examples and remindings) stored with the learner model, a best match can be computed and offered to the learner as an example solution to a similar programming problem. Fourth, the temporarily stored case is removed.

The rationale of using remindings to programming tasks that have been solved or explained previously is to reduce the mental effort (in terms of the cognitive load) of the learner during problem solving. If an example is shown to the learner that he or she has never seen or worked at before, the learner has to switch from the current problem to the new example. He or she has to understand the task description, understand the solution to the task and explain how this works. Than, he or she has to switch back to the current problem, re-understand the current task descriptions and has to explore the similarity of the example to the current task. In case the example is a reminding to a previously solved or explained programming problem, the learner already knows the task description of the example and is reminded of the solution he or she already solved on his or her own or is reminded of the explanation he or she already understood when reading the example in the text book. So, it is easier to switch back to the current task and to draw similarities and explain differences between the reminding and the current task. Burow & Weber (1996) have shown how drawing analogies to example solutions can be done automatically by the system. This automated example explanation explains the similarity of the example to the current task. And it can focus the learner on those aspects of the example that differ from the current task by explaining what the different goals are in both situations.

Generalization of ELM

The episodic learner model has been developed especially for supporting users learning to program in the programming language LISP. So, most programming concepts and rules of the knowledge based are specific for this domain. One may question whether it is possible to generalize this approach to other domain. Generally, the basic ideas of the ELM model can be applied to other problem solving domains with problem solving tasks that can be described by a hierarchy of subgoals and sub-plans. This is true not only for programming languages but also for many topics in mathematics, natural sciences, and technical domains that can be formally described. However, though the general ELM architecture can be used in these domains, concepts and rules have to be developed specifically for each domain.

As it is very difficult to create new concepts and rules, inexperienced authors like teacher will be overstrained by this task. So, trained knowledge engineers can only do this. This is not only true for ELM but also holds for all other rule-based intelligent tutoring systems. This is one reason why more lightweight adaptive systems (as described in the next Section) have evolved during the last years. In Section 4, an authoring system for developing adaptive learning courses will be introduced. This authoring system can be applied to a wide range of topics in different domains and can be used even by inexperienced authors.

Adaptive Learning Environments

The development of intelligent tutoring systems did hit its peak at the end of the last century. The effort to develop more and more sophisticated systems was too high compared to the relative outcome. Many systems resulted from doctoral theses and were mainly tested or used in laboratory settings. Besides the intelligent tutoring systems developed by the ACT-group at CMU (Corbett et al., 2001), only few systems did find their way out of the laboratory into classrooms or into the web. Instead, the development of adaptive systems emerged. Adaptivity means that the system adapts one or more features of the system to a particular user according to information from an individual user model, e.g., the presentation of adapted texts in learning materials, the annotation of links in hypermedia, the individual guidance through the learning space, or the adapted presentation of hints or feedback.

While intelligent tutoring systems have been developed prominently in more formal domains like mathematics, computer science, or physics (especially for the support of problem solving), adaptive techniques can be applied to a much wider range of learning topics and do not require as sophisticated learner models as an ITS. In the following, an overview of adaptive techniques in hypermedia systems will be given together with comments on user models used. Then a special user model suited to adaptation, the multi-layered overlay model (used in ELM-ART), will be described.

Adaptive Hypermedia

Web-based learning systems belong to the large group of hypermedia. Very early, hypermedia has been the subject of adaptation techniques. The well-known and often mentioned 'lost in hyperspace' problem that describes the experience of getting more and more disoriented when following links in hypertexts and hypermedia led to the need of developing techniques to support hypermedia users and learners in hyperspace navigation.

Current adaptive hypermedia use different types of adaptation techniques (Brusilovsky, 1996). The most prominent adaptation techniques are adaptive presentation, adaptive curriculum sequencing, adaptive navigation support, and adaptive hints and feedback. Additionally, adaptive collaboration support, specially designed for the context of Web-based learning environments, is the most recent branch of adaptation techniques.

Adaptive presentation techniques are used to adapt the content of hypermedia pages (texts and multimedia materials) in learning systems to the learners' goals, knowledge, and other information stored in the learner model. In a system with adaptive presentation, the pages are not static but adaptively generated or assembled from different pieces for each learner. For example, with several adaptive presentation techniques, expert users receive more detailed and deep information, while novices receive more additional explanation. However, this simple technique based on a stereotype learner model only is not very well supported by pirical evidence. There is no general model or theory to describe how detailed in-information should be presented even for a simple dichotomy as expert versus novice users. Therefore, this adaptation technique is best coupled with features of adaptability, that is, users can tell the system their preferences of how detailed text or other materials should be presented. An advanced feature offers the possibility to inspect the user's own user model and to modify this model in such a way as to present materials in the preferred form (Kay, 2001). A more flexible system, the adaptive hypermedia system AHA! (De Bra, Smith, & Stash, 2006) uses adaptive text presentation based on sophisticated rules that authors can develop in a rule editor.

Curriculum sequencing (also referred to as instructional planning technology) provides the learner with the most suitable individually planned sequence of knowledge units to learn and with a sequence of learning tasks (examples, questions, problems, etc.) to work with. In other words, it helps the learner to find an "optimal path" through the learning material. Principally, there are two different types of curriculum sequencing. First, a course of pages is selected for a specific user in advance and the user is presented with the sequence of these pages. This type of curriculum sequencing is often based on a stereotype learner model or on a learner model that is *fixed* during the course and will be updated after completing the course. Second, the next best page or task to work at is computed on-line based on a *dynamic* learner model that is updated with each learning step and with each task the learner worked at. This technique is usually used in intelligent tutoring systems and in adaptive systems that utilize advanced, dynamic learner models. An example of such a dynamic learner model, the multi-layered overlay model (Weber, 1999), will be described in the next section.

Adaptive navigation techniques are used to support the learner in hyperspace orientation and navigation by changing the appearance of visible hyperlinks. The system can adaptively sort, annotate, or partly hide the links of the current page to make easier the choice of a next link to proceed. However, sorting and partly hiding links has the disadvantage of providing the learner with an inconsistent user interface. The number and the location of links change over time. So, besides the actual learning goal the learner has to adapt to the user interface. This is the reason why current learning systems prefer adaptive link annotation. Adaptive navigation support can be considered as an extension of the curriculum sequencing technology into a hypermedia context. It shares the same goal - to help learners to find an "optimal path" through the learning material. At the same time, adaptive navigation support is less directive than traditional sequencing: it guides learners implicitly and leaves the choice of the next knowledge item to be learned and next problem to be solved to the learner.

Adaptive hints and feedback depend on more ambitious learner modeling. An example of learner modeling that allows for explanation of errors in problem solving, the episodic learner model, has been described above. Adaptive hints and feedback are typically used in intelligent tutoring systems that implicitly use adaptive techniques.

All these adaptation techniques use some type of learner modeling. While presentation adaptation and simple curriculum sequencing can be based on stereotype learner models, adaptive navigation support and dynamic types of curriculum quencing require at least overlay learner models. The most sophisticated adaptation techniques used for adaptive hints and feedback require more advanced AI-techniques that typically are used in intelligent tutoring systems (e.g., rule-based or case-based reasoning). In the following section, the multi-layered overlay model, the further development of the frequently used overlay modeling technique (Carr & Goldstein, 1977), will be described. This type of learner modeling is used in the complex learning environment ELM-ART (Weber & Brusilovsky, 2001) to support learning new programming concepts by adaptive annotation techniques and dynamic curriculum sequencing.

The Multi-Layered Overlay Model

In order to understand the multi-layered overlay model, the knowledge representation used in ELM-ART has to be described shortly. ELM-ART supports learning the programming language LISP. It is a model-based learning environment that combines an adaptive hypertext with a problem solving support tool. Figure 4 gives an overview of all components of ELM-ART.

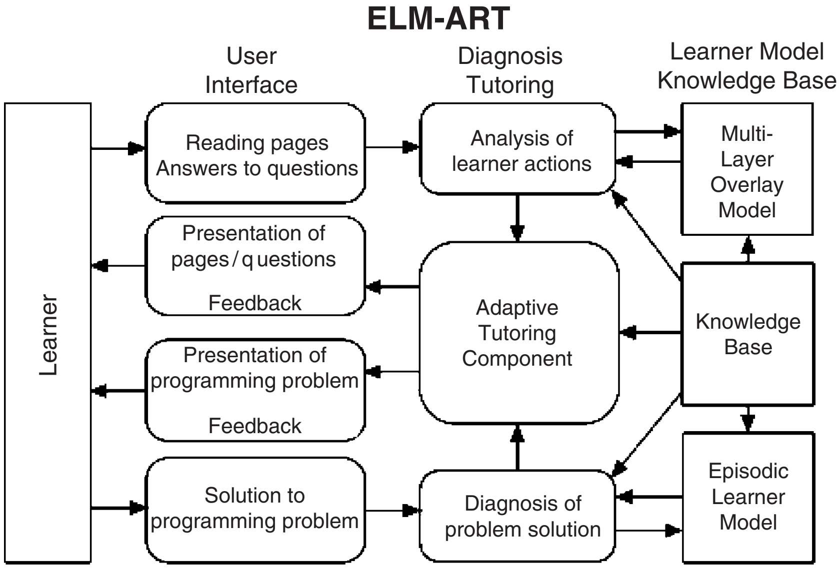

Fig. 4. Components of ELM-ART.

The learner model ELM used in the problem solving support tool has been described above. The adaptive hypertext consists of text pages introducing concepts, the single knowledge units of the knowledge base. Concepts are represented in terms of a conceptual network. Units are organized hierarchically into lessons, sections, subsections, and terminal (unit) pages. Terminal pages can introduce new concepts, present lists of test items to be worked at, or offer problems to be solved. Each unit is an object containing slots for the text unit to be presented with the corresponding page and slots for information that can be used to relate units and concepts to each other. Slots store information on prerequisite concepts (the pages that the learner has to be worked at successfully in advance so the learner is assumed to possess the required knowledge to understand the current page) and related concepts (the concepts that the system assumes to be known if the user worked through that page successfully). Additionally, each unit can have a list of test items, a so-called test-group, or a programming problem to be solved by the learner.

The learner model used for adapting the presentation of the course materials to learner is related to the concepts of the system's knowledge base. Therefore, the model used in ELM-ART is a type of an overlay model (Carr & Goldstein, 1977). An overlay model assumes that the knowledge of a user about a domain can be related directly to the user's state of knowledge about particular concepts. In simple overlay models, this knowledge state can have the values "known" or "not known". In ELM-ART, an extended type of overlay model is introduced (Weber & Brusilovsky, 2001), the multi-layered overlay model. Layers of the overlay model describe different aspects of users' knowledge of a particular concept.

The first layer describes whether the user has already visited a page corresponding to a concept. This is the same information a browser has during a session about the visiting state of hyperlinks. Links to pages already visited are displayed differently. So they can be distinguished from links to pages that have not been visited at all. However, this information gets lost when the learner closes the browser. In the learner model, this information is stored permanently. As an extension to this layer, the number of times a learner visited the page and total amount of time the learner stayed with this page can be recorded and stored with the learner model.

The second layer contains information on which exercises or test items related to this particular concept the user has worked at and whether he or she successfully worked at the test items up to a criterion. The computation of the "solved" value and the comparison with the criterion will be described in the next section.

The third layer describes whether a concept could be inferred as known via inference links from concepts the user already worked at successfully. That is, if the "solved" value of a concept (in the second layer) reaches the criterion, all concepts that are related to this "solved" concept by an inference link will be marked as inferred in this third layer.

The fourth layer describes whether a learner has marked a concept as already known. Learners can inspect their own learner model and mark single pages or whole sections as already known. Therefore, the multi-layered learner model is a scrutable or cooperative user model (Kay, 2001).

Information in the different layers is updated independently. So, information from each different layer does not override information from other layers. For example, learners can mark and unmark pages as already known at any time. This will change only the according value in the fourth layer. Values in the other layers will not be influenced by these changes.

The multi-layered overlay model supports both the adaptive annotation of links and individual curriculum sequencing. Links to pages of the course are shown in an overview on each page or in a separate table of contents. These links are visually annotated according to the learning state of the corresponding concept. Depending on the multi-layered learner model, five different learning states of a concept can be distinguished.

1. A concept is annotated as 'already learned' if enough exercises or test items belonging to that concept or the programming problem have been solved successfully. This means, the "solved" value in the second layer has reached the criterion.
2. The concept is annotated as 'inferred' if the concept is not 'already learned' and if it was inferred as learned from other concepts (third layer).
3. The concept is annotated as 'stated as known by the learner' in case the learner marked this concept as already known and there is no information that the concept is 'already learned' or 'inferred'.

4. A concept is annotated as 'ready and suggested to be visited' if it is not assigned to one of the first three learning states and all prerequisites to this concept are assigned to one of the first three learning states.
5. A concept is annotated as 'not ready to be visited' if none of the other four learning states hold.

Link annotation is used as a hint only. That is, a learner can visit each page even if its learning state is not 'ready and suggested to be visited'.

Individual curriculum sequencing in ELM-ART means that the system's suggestion of the next page to visit is computed dynamically according to the general learning goal and the learning state of the concepts as described above. The next suggested page will belong to the concept that is not assigned to one of the first three learning states and that is the next one ready to be learned.

The next section will describe how the "solved" value of a concept (second layer) is computed based on information from solving test items during working at the course.

Modeling the Learner's Knowledge in ELM-ART

The learner-modeling component of ELM-ART uses a simple mechanism to compute the "solved" value of a concept that is needed to determine the learning state of a particular learner. Each test item has a difficulty value that is pre-set by the author of the learning course. In the beginning, this value has to be estimated by the author or has to be set to a default value. The number of hits and errors of users solving this test item can be recorded in a list. So, with increasing usage of the system, the difficulty value of a particular test item can be computed with respect to the recorded numbers of hits and errors. Each concept has a list of test items, the so-called test-group. Each test item in a test-group is associated with a weight value that may describe the importance of the test item to the particular concept. Usually, this weight is set to a default value. If a user has solved a test item successfully, the difficulty value of the test item is multiplied with the weight value of the test item in this test-group. The resulting value increments the "solved" value of the concept. Each test-group has a critical value (set by the author of the system). If the "solved" value of a concept reaches this critical value, the concept is marked as 'learned' in the second layer of the learner model.

The power of the adaptation techniques used in ELM-ART not only stems from computing the learning state of a page directly from observing the success of solving test items that belong to this page but also from the elaborate inference mechanism. As mentioned in the previous section, concepts have slots describing links to related concepts. If the "solved" value of a concept changes, this change will spread to the related concepts and by a simple inference mechanism the "solved" value of the related concept will be changed, too.

Test items have slots with links to related concepts, too. If a learner visits a page and solves a test item presented with this page successfully, not only the "solved" value of the concept (page) will be changed but also the "solved" values of the concepts addressed by the inference links to related concepts of the test item will be changed accordingly. With this inference mechanism, pre-tests at the beginning of the course or of a section can be used to gather information about the initial learning state of the pages of the course or of the section. If there is enough evidence from solved test items in the pre-test, that is the "solved" value of the inferred concept reaches the criterion, this concept will be in the state 'already learned' and learners will not longer be guided automatically to these pages (though learners are not prohibited from visiting theses pages). Weibelzahl & Weber (2003) could show in an evaluation study that adaptation based on such a pretest can lower the learning effort by reducing the total learning time and the number of pages a learner has to read or to work at.

Authoring Adaptive Learning Environments

Adaptive learning systems will only find their way out of the laboratory into schools and further development settings if the time and effort of developing such a system will show a reasonable relation to its additional outcome. To make the development of adaptive systems easy the authoring system and Web-Server NetCoach (Klein & Weber, 2002) has been created. NetCoach can both present adaptive learning environments in the Internet and support the development of such systems with an easy to use authoring system.

NetCoach (url: www.net-coach.de) is based on CL-HTTP (Mallery, 1994), a LISP-based Web-server developed to integrate artificial intelligence tools directly into a Web-server. As a first step, the adaptive learning environment ELM-ART (Weber & Brusilovsky, 2001) was developed in CL-HTTP. Since more than ten years ELM-ART is running in the Internet (url: art2.ph-freiburg.de/Lisp-Course) and has been used by thousands of users from all over the world to learn the first steps of programming in LISP. In a second step, ELM-ART was the basis for developing NetCoach by adding extended authoring tools that allow authors to develop their own adaptive learning courses and learning environments without any programming knowledge required.

The presentation component of NetCoach offers a lot of tools that support learners while working with the learning environment, e.g., communication tools (chat, discussion panels, document exchange, e-mail communication with tutors), text search, modifying preferences, etc. Weber & Brusilovsky (2001) describe these features in more detail. This paper concentrates on explaining the tools for authoring and editing the adaptive features of the learning environments.

Developing Adaptive Courses

The author of an adaptive NetCoach course can edit and manage pages with the course editor (see figure 5). The editor allows for creating and inserting new pages into the page hierarchy, to move pages or sub-trees to different places in the hierarchy, to delete pages or sub-trees, and to import pages from other courses.

The title and the text of the page can be edited (directly in HTML-code or with support of a wysiwyg-editor). In case of a multi-language course the language of the texts to be edited can be chosen. In the parameter settings of a page as well features for the presentation of the page as information necessary for the computation of the adaptive navigation support can be set. First, the prerequisite pages that have to be worked at before visiting the current page can be set. Second, inference links can be set to pages that the system assumes to be already known by the learner if he or she successfully worked at the exercises of this page. Figure 6 shows an example of a page with a hint that the learner should read or work at a prerequisite page before trying to solve the exercises with this page. The system computes this adaptive hint based on information from the learner model.

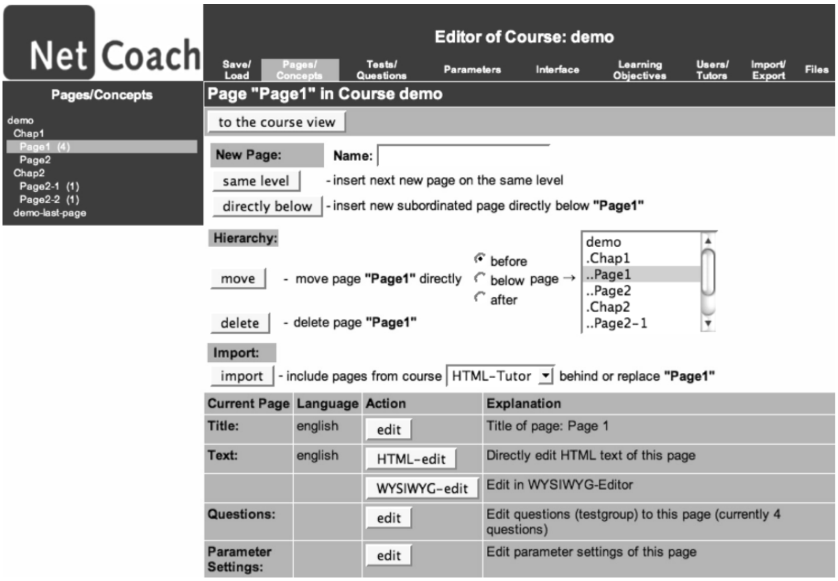

Fig. 5. Editor for Authoring Pages in NetCoach.

It tests whether all pages that are marked as prerequisite pages to the current page are read or worked at successfully. In this case, the learner did not read the prerequisite page "Chapter 1" that the author assumed to be necessary to understand the contents of the current page and to solve the exercises successfully. However, the learner is not required to follow this adaptive link and can continue with working at the current page in order to prove that he or she did already learn enough to solve the exercises.

Authoring Tests and Exercises

Tests and exercises are the main sources for updating the learner model. Single test items from a pool of tests can be assigned to pages. They are collected in a test-group stored in the concept description underlying the page. In a test-group editor, the author can set how the test items will be presented on the page, how many test items are presented at once, and how feedback will be given on wrong and/or correct answers. The author sets the critical value that has to be reached by learners to show that the concept underlying the page has been learned successfully, that is, the "solved" value in the learner model for this concept reaches the critical value. For each test item the weight value, describing a relative weight of the test item in the test-group can be set.

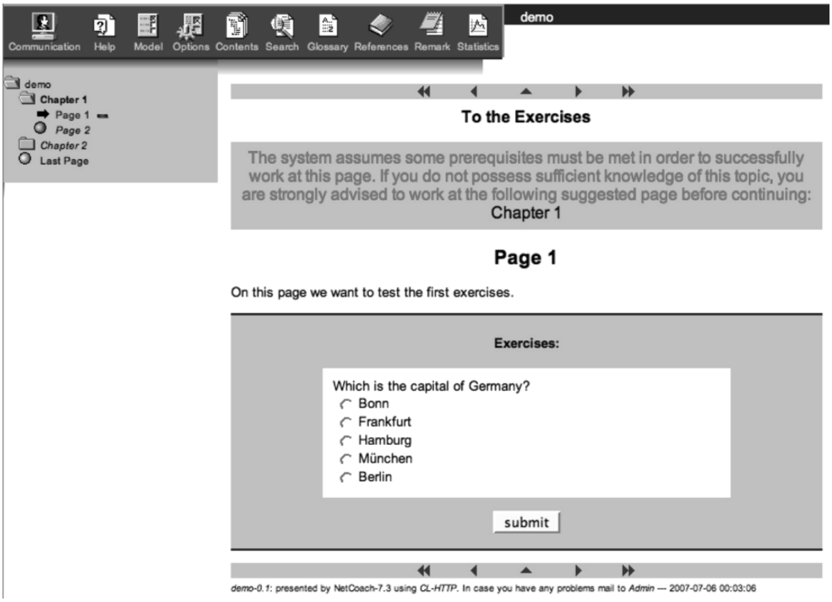

Fig. 6. Presentation of a Page with Adaptive Navigation Support in NetCoach.

Single test items can be created and edited in a separate test item editor. Different types of tests are supported in NetCoach: forced choice and multiple-choice tests, gap-filling tests (with single or multiple gaps), free-input tests (simply testing on the equality of strings), and Flash tests (that present a Flash applet and get results from the Flash applet via a javascript API). The editors for free-input and gap-filling test items allow for describing a list of possible correct answers to the question asked with this test item. So, semantically similar solutions or different spellings of the solution can be identified by the system when evaluating the answers of learners to the test items.

Authoring Learning-Goals

The default learning-goal in NetCoach courses is to read all pages of the course and to work at all pages with tests and exercises successfully. Very often, a final test at the end of a course will be used to decide on whether the learner has worked at the course successfully. In complex courses, different learning goals may be appropriate. For example, in further education settings, only single chapters of a course may be interesting to learners. Or the learner may already know the basics of a domain offered in the course and is only interested in some advanced topics.

The author of the learning environment can describe such cases in a learning-goal editor. He or she can assign a text to the goal that will be presented to learners in a choice list at the beginning of the course. The author can describe the goal by marking all concepts (pages) that have to be worked at to learn the goal successfully. Only the final concepts have to be marked because all prerequisite concepts to these final concepts are computed by the editor recursively. Additionally, concepts that the learner may already know can be marked and, therefore, will not be included with the individual learning path for the particular learner. When a learner selects such a goal with concepts marked as to be known already, these concepts are marked temporarily in the fourth layer of his or her learner model. When the learner switches to another learning goal, these marks will be removed. As changes within one layer of the multi-layered learner model do not interfere with information from other layers, switching marks in the fourth layer on and off can be done at any time during the learning process.

The HTML-Tutor (url: art.ph-freiburg.de/HTML-Tutor) is an example of an adaptive learning system that explicitly uses learning goals. A slightly modified version of this tutor was used in an experiment by Weibelzahl & Weber (2003) to evaluate the inference mechanism of the adaptive techniques used in NetCoach-Tutors and to show that adaptive curriculum sequencing reduces the learning effort to successfully solve the learning goals of the course.

Conclusion

The paper has outlined the changes in the development of model-based learning environments. It started from the development of sophisticated intelligent tutoring systems that are based on expensive rule-based and case-based learner models using advanced AI-techniques to infer the learning state of a single user. In the last ten years, the development of more simple adaptive learning environments has emerged. These learning systems adapt to a particular learner using more or less complicated learner modeling techniques. While only few intelligent tutoring systems have found their way from development and evaluation in a laboratory to regular use in classrooms or in further education, adaptive learning environments are used more widespread these days.

With ELM-ART, an example of a complex model-based learning environment, it could be shown how different learner modeling techniques can be used to support the individual learning process. While the hybrid rule-based and case-based learner model ELM supports the acquisition of programming skills when solving programming problems, the multi-layered overlay model is used to adapt the user's learning path through the learning materials and to adapt the presentation of hyperlinks.

Based on the experiences with thousands of users learning with ELM-ART during more than ten years, the model-based Web-server NetCoach has been created. This system supports authoring adaptive on-line courses that use the adaptation techniques 'individual curriculum sequencing' and 'adaptive link annotation' based on the multi-layered overlay model. Up to now, a large number of adaptive on-line courses have been developed with NetCoach. Weber, Lippitsch, & Weibelzahl (2004) report on several NetCoach courses field-tested during the last years. Experimental studies already mentioned above have shown how adaptive learning environments are successful especially in further development settings.

From the current point of view one can expect that the development of model-based learning environments, especially the implementation of the more simple adaptation techniques, will be used in a growing number of on-line learning environments. Not only proprietary systems like NetCoach (Klein & Weber, 2002) or AHA! (De Bra et al., 2006) stemming from research laboratories will be used in the future. New developments indicate that adaptation techniques will be implemented in future standards for learning environments like SCORM, too (e.g., Rey-López, Fernández-Vilas, Díaz-Redondo, Pazos-Arias, & Bermejo-Muñoz, 2006).

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R., Boyle, C. F., & Yost, G. (1986). The geometry tutor. *The Journal of Mathematical Behavior*, *5*, 5-20.

Anderson, J. R., Conrad, F. G., & Corbett, A. T. (1989). Skill acquisition and the LISP tutor. *Cognitive Science, 13*, 467-505.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences, 4*(2), 167-207.

Brusilovsky, P. (1996). Methods and techniques of adaptive hypermedia. *User Modeling and User-Adapted Interaction*, *6*, 87-129.

Burow, R., & Weber, G. (1996). Example explanation in learning environments. In C. Frasson, G. Gauthier & A. Lesgold (Eds.), *Intelligent Tutoring Systems – Proceedings of the Third International Conference, ITS'96* (pp. 457-465). Berlin: Springer.

Carr, B., & Goldstein, I. (1977). *Overlays: A theory of modelling for computer aided instruction.* Unpublished AI Memo 406, Cambridge, MA: Massachusetts Institute of Technology, AI Laboratory.

Corbett, A. T., Koedinger, K. R., & Hadley, W. H. (2001). Cognitive tutors: From the research classroom to all classrooms. In P. S. Goodman (Ed.), *Technology enhanced learning: Opportunities for change* (pp. 235-263). Mahwah, NJ: Lawrence Erlbaum Associates.

Corbett, A.T. (2001). Cognitive computer tutors: Solving the two-sigma problem. *User Modeling: Proceedings of the Eighth International Conference*, UM 2001,137-147.

De Bra, P., Smits, D., Stash, N. (2006). Creating and Delivering Adaptive Courses with AHA!, *Proceedings of the first European Conference on Technology Enhanced Learning, EC-TEL 2006, Springer LNCS 4227* (pp. 21-33), Crete, October 1-4, 2006.

DeJong, G. (1988). An introduction to explanation-based learning. In H. E. Shrobe & AAAI (Eds.), *Exploring artificial intelligence* (pp. 45-81). San Mateo, CA: Morgan Kaufmann Publishers.

Johnson, W. L. (1986). *Intention-based diagnosis of novice programming errors.* London: Pitman.

Kay, J. (2001). Learner control. *User Modeling and User-Adapted Interaction, Tenth Anniversary Special Issue*. *11*(1-2), 111-127.

Klein, B., & Weber, G. (2002). Die Realisierung von Adaptivität mit dem Internet-Autorensystem NetCoach. In U. Rinn & J. Wedekind (Eds.), *Referenzmodelle netzbasierten Lehrens und Lernens* (pp. 101-118). Münster: Waxmann.

Kolodner, J. L. (1993). *Case-based reasoning*. San Mateo, CA: Morgan Kaufmann.

Lebowitz, M. (1986). Concept learning in a rich input domain: Generalization-based memory. In R. S. Michalski, J. G. Carbonell & T. M. Mitchell (Eds.), *Machine learning: an artificial intelligence approach* (Vol. II, pp. 193-214). Los Altos, CA: Morgan Kaufmann.

Rey-López, M., Fernández-Vilas, A., Díaz-Redondo, R. P., Pazos-Arias, J. J., & Bermejo-Muñoz, J. (2006). Extending SCORM to create adaptive courses. *Innovative Approaches for Learning and Knowledge Sharing*, *4227*, 679-684.

Mallery, J. C. (1994). A Common LISP Hypermedia Server. *Proceedings of The First International Conference on The World-Wide Web*, Geneva: CERN, May 25.

Mitchell, T. M., Keller, R. M., & Kedar-Cabelli, S. T. (1986). Explanation-based generalization: a unifying view. *Machine Learning, 1*(1), 47-80.

Schank, R. C. (1982). *Dynamic memory*. Cambridge, MA: Cambridge University Press.

Weber, G. (1996a). Episodic learner modeling. *Cognitive Science, 20*(2), 195-236.

Weber, G. (1996b). Individual selection of examples in an intelligent programming environment. *Journal of Artificial Intelligence in Education, 7*(1), 3-31.

Weber, G. (1999). Adaptive learning systems in the World Wide Web. In J. Kay (Ed.), *User modeling: Proceedings of the Seventh International Conference, UM99* (pp. 371-378). Wien: Springer-Verlag.

Weber, G., Bögelsack, A., & Wender, K. F. (1993). When can individual student models be useful? In G. Strube, & K. F. Wender (Eds.), *The cognitive psychology of knowledge. The German Wissenspsychologie project* (pp. 263-284). Amsterdam: Elsevier.

Weber, G., & Brusilovsky, P. (2001). ELM-ART: An adaptive versatile system for Web-based instruction. *International Journal of Artificial Intelligence in Education, 12*(4), 351-384.

Weber, G., Lippitsch, S., & Weibelzahl, S. (2004). Virtuelle Lernkurse zur Einführung in die Psychologie für Pädagogik- und Lehramtsstudierende. In U. Rinn & D. M. Meister (Hrsg.) *Didaktik und Neue Medien - Konzepte und Anwendungen in der Hochschule* (pp. 149-166). Münster: Waxmann.

Weibelzahl, S., & Weber, G. (2003). Evaluating the Inference Mechanism of Adaptive Learning Systems. In Brusilovsky, P., Corbett, A., & de Rosis, F. (Eds.), *User Modeling: Proceedings of the Ninth International Conference, UM2003* (pp. 154-162) (Lecture Notes in Computer Science). Berlin: Springer.

Wenger, E. (1987). *Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge.* Los Altos, CA: Morgan Kaufmann Publishers.

Wolstencroft, J. (1989). Restructuring, reminding, repair: What's missing from models of analogy. *AI Communications, 2*, 58-71.

13. Model-Centered Learning and Instructional Design

Patrick Blumschein

Albert-Ludwigs-University, Freiburg, Germany

Abstract: This chapter discusses how the approach of mental models can be applied to Instructional Design (ID) and eventually leads to a theoretical concept of model-centered Instructional Design. The various critiques of Instructional Design which describe ID as too fixed, slow and which doubt that there is room for ID in the Information Society, will serve as the starting point for the discussion. This contribution takes up these critical views and opens up a look at ID from a model-centered perspective. The argumentation starts out with a reference to the underlying epistemological foundations and is based on the fundamental understanding of human cognition as the construction of mental models when confronted with more or less complex challenges. Further-on the application of the concept of mental models to learning and Instruction is the "building block" to ID. Thus ID is understood as a higher-order process of model construction. Thereby the approach of mental models can lead to a new perspective on Instructional Design, which can successfully defend against the critiques of Information Technologists. However, Model-Centered Instructional Design may help to see existing excellent layouts of ISD products from another perspective, which helps to understand the mystery of the human learning processes more precisely than other can do. The question if this leads to a paradigm shift in Educational Science hasn't been answered yet.

Keywords: Instructional design; mental model; ADDIE; learning, model-centered instruction.

Challenges for Learning and Instruction

Instructional Design in the Information Age

In the course of the technological revolution of the Internet and Information Processing Systems the death of Instructional Design (ID) was predicted. Some critics for instance argue that ID will be subsumed under the category of Information Design (see Duchastel, 1999; Myers, 1999). Other critics postulate that ID will have no future for its incapability of meeting the requirements of the rapidly changing Technology and Information Age (Zemke & Rossett, 2002; Gordon & Zemke, 2000; Thiagarajan, 1999; Gayeski, 1998). Another argument is that ID will not be able to cope with the heterogeneous groups of learners and the challenges of problem-based learning. ID is considered too slow, too imprecise, and too much focused on reference points. Also it has been said that ID is based on an out-dated understanding of truth (see Clark, 2004a, b).

A Delphi Study on the future of Instructional Design by Ritchie and Earnest (1999) identified 6 important future trends of ID. The authors conclude that Instructional Designers should direct their focus more towards intercultural and global learner. The developmental processes should be faster and more adapted, less linear and fixed and evaluation processes should be usefully integrated. Furthermore, computer-based, worldwide directly assessable learning environments as well as a still further consideration of the users needs are important future trends. Also Instructional Designers should continuously extend their competences. On the following issues the experts opinions were ambiguous: cost-benefit orientation, automation through knowledge systems, standardization, and performance orientation instead of focusing on domain knowledge. In the following some of the arguments will be addressed. For more details of definitions of ID, ISD as well as a historical overview see Seel's Introduction to Instructional Design (2003c; cf. Andrews & Goodson, 1980; cf. Merrill, 2002b).

Linearity in the ID Process

A point that might be criticized about ID results from instructions how to act which are linearly deduced from steps of analysis. This contradicts with the underlying constructivist epistemological theory and therefore cannot be accepted per se. The following example illustrates this idea. The simple fact that calculation and geometry are important basics needed to determine the distance from one place to another, especially considering that various means of transportation must be taken into account (see e.g. Jasper-adventure by Bransford et al., CTGV, 1997), does not necessarily demand a hierarchical and structured knowledge transfer in a learning environment as suggested by Gagné (Gagné & Briggs, 1979 or Reigeluth, 1987); however, it is neither stated that a sequenced procedure is wrong. Detaching diagnoses procedures of lacks of competences from direct instructional prescriptions (e.g. learning hierarchy) is clearly pointed out by the design layouts by Bransford et al. (2001) and Schank et al. (2001). In both cases students are confronted with complex learning environments that demand a high degree of self regulation. This does not, however, imply a devaluation of the traditional approaches to ID which would be a post-hoc fallacy (cf. Lowyck & Elen, 1991). These models and procedures should not be understood as linear instructions how to act but as attempts to give a thorough explanation of human thinking, memory, and action with regard to the prevailing scientific era. Schank's (1993) approach of Goal-Based-Scenarios (GBS) is a good example of breaking with linear structures. GBS initially demand an intensive and highly structured analysis and formulation of goals but they then provide a rich problem-based learning environment, which offers the learner a very high degree of self regulation and individual responsibility. Also, the well-known ADDIE model (Analysis, Design, Development, Implementation, and Evaluation) in Instructional Design may give at first sight the impression to demand a linear procedure to develop a successful educational concept (Gayeski, 1998; Gagné, 2005). However, it does not force anyone to do so. Also, a counter model has not been developed.

Taking into account several critics of ID, this chapter presents model-centered approaches to thinking, learning, instruction, and instructional design as a way to overcome such reproaches. The following paragraphs are therefore dealing with epistemological issues as a foundation for doing instructional design, model-centered learning theory, and model-centered instruction, before focusing on a model-centered instructional design approach. This will show the advantage of Instructional Design over bare Information Design. But to begin, a central starting point is the question of how to overcome the gap between the actual and the desired state and the meaning of being an expert as a learner and also being an expert as an Instructional Designer.

The Gap between the Actual and the Desired State – a Gap between Knowledge and Action

A world that is becoming more and more complex is characterized by two main developments: its processes are influenced by a greater number of systems and those systems are becoming more and more complex themselves. Complex systems usually do not have unambiguous cause-and-effect systems; they consider time lags and often many intervening variables (for more details see e.g. Levin; Wiener; Bateson; Parsons). Also Instructional Design processes face these conditions. In further education programs most of the mentioned problems become apparent very clearly. A typical situation is that conditions often change during the developmental phase so substantially that one actually has to start all over again and again. Furthermore working and learning processes more and more merge together. Learning is integrated into the daily work routine to be more effective; this is done for instance with help of software tutorials. On the other hand, face-to-face trainings for software applications are often not effective because they overload students with new termini, routines, and a lack of understanding and practical orientation. So in many cases Instructional Designers should say good bye to the idea that “training” is the universal solution to competence deficits – if at all the problem lies in competence deficits. Other training plans can often just partly solve the given problems since staff qualification does not solve organizational problems. Therefore a strategic-systemic concept is necessary (see Rothwell & Kazanas, 2004; Kaufman, 2003). Kaufman (2003) therefore points out the central questions: What Is? And What Should Be? Where Are The Gaps? For “Mastering The Instructional Design Process” Rothwell and Kazanas (2004) are also discussing non-instructional solutions to some problems. Even if ID can solve only a part of the problems that must be faced, its use is justified for the phase of analysis which can bring to light the conditions that are necessary for solving the problems. The insight that processes of change must accompany educational trainings is a result of this analytical process as well. To find out about the gap between what is and what should be, a comparison of novice and expert behavior can be useful.

Experts as Models to Optimize Processes of Learning and Instruction

Today expertise research has a remarkable tradition within the field of Educational Science (Chi, Glaser & Farr, 1988; Sternberg, 1998; Bransford, 2004; Seel, 2003). In Instructional Design expert behavior serves as a competence model for the qualification of novices. Expert behavior is analyzed and compared to novice behavior in order to learn more about the gap between the actual and the desired state of affairs. Making this comparison, some very interesting phenomena occur. Experts do not necessarily know more, and they are not always faster than novices. Differences can be observed, however, in the way they work. The experts’ knowledge seems to be structured more effectively and is therefore easier to recall. Surprisingly, experts are not always better at describing what they do and how they do it. The observation of expert behavior and the corresponding instruction of novices is an important idea within Instructional Design (Collins, Brown, & Newman, 1989). This strategy is very well demonstrated in “How people learn” by Bransford (2004). A crucial question in this context is what expertise the Instructional Designers have (Hardré, Ge, & Thomas, 2005). The authors identified three dimensions of expertise in Instructional Design that are related to three layers of knowledge. The first layer is the “domain specific factual and structural knowledge about instructional design principles. A second layer involves the skilful and adaptive use of various instructional design tools in different design contexts. A third layer of expertise involves metacognition, the self-awareness and self-regulation of one's problem solving activity." (Hardré et al., 2005, 55). Accordingly, Model-Centered-Instruction (MCI), thus supporting closing the "Gap", has to reflect the learner's strategy how to solve the problem in comparison with the expert solution in order to develop a model for a sufficient learning environment. The most important question to be answered in this context deals with a strong reference to human cognition. Therefore we need to take a look at the epistemological basis of a model of learning.

Model-Centered Learning

The Epistemological Basis for a Model-Centered Learning Theory

In the 90's the constructivist school of thought, which was already known in ancient times, started to become more popular in the field of didactics (cf. Duffy, Lowyck, & Jonassen, 1993; Gerstenmaier & Mandl, 1995). In retrospect, the discussion on the philosophical foundations of constructivism and the demands for "constructivist" instructions in didactics was confusing and inadequate (see Winn, 1993; Strauss, 1996; Chadwick, 2004). For a useful discussion on the role of constructivism within educational science, it must be regarded as a theory of cognition. In a next step, a theory of learning and instruction must be deduced, and after that a prescriptive theory of ID with principles and rules can be unfolded. This point of view is of utmost importance for the following discussion on model-centered learning and instruction, since constructivism suggests that each individual person constructs a subjective concept of reality. Therefore reality is not regarded as a given or objective state individuals adapt to. If this was the case, knowledge about the world would consist of a never ending number of assimilation processes (Groundhog Day) and therefore we would have a hard time to think about inspiration, creativity and new constructions. Of course, this can never be a desired state of human existence. For disputants of constructivism a typical counter argument is about a chair that is said to be a part of an objective reality. This may be true at first sight, but what is the point if people do not have cognitive access to it? Individuals' minds have to perceive a chair as such and know how to experience and modify it in order to act in the world. Consequently reality is always constructed individually. And this process is "radical" since one can take over a position for or against it (cf. von Glasersfeld, 1987). A chair only makes a chair, if the individual perceives it as such – half a chair is no chair. The fact that there are many people who share the idea that a given chair can be identified as a chair, indicates that ontology defines humans as social creatures and suggests that different people undergo the same socialization and education processes. This shared knowledge is decisive for successful communication. Also, it explains why differences in culture and education can lead to a failing communication (Schöfthaler et al., 1984; von Glasersfeld, 1987; Bateson, 1995). As demonstrated, the individual constructs his or her subjective world. Consequently, learning is an individual construction process and as well a reflection about this process. Since humans are social beings and as such essentially oriented to communication, the idea of individual knowledge construction must be kept in mind in all respects of human thinking and acting. Reflection about the own thinking and acting processes is therefore a crucial part of any human learning process. This can be considered one of the major fallacies of behaviorism. Since the individual has no direct access to the perception and thinking processes of his or her fellows, he or she can only speculate about their nature. To do so, he or she develops a model of how to comprehend and think and thus develops a mental model. This model must be regarded as incomplete, imperfect, partial, instable, changeable and absolutely individual and therefore idiosyncratic (cf. Seel, 1991, 2003b). Human thinking and acting can thus be considered an extensive process of model construction. In order to describe this process, Piaget (1972) introduced the termini assimilation and accommodation. According to his philosophy, human beings strive for a balance state. However, reaching the state of total balance would ultimately mean the end of life. Therefore, the striving for balance must be understood as a drive but not as a desired aim. This striving is the basis for the development of cognitive structures, which then will be stored as models-of to be at one's disposal on a long-term basis. This must be the foundation of a model-centered learning theory.

Assimilation Learning – the Foundation of Model-Centered Learning

Piaget (1972) calls it assimilation if perceived information can be matched with existing cognitive structures (see Case, 1987). Assimilation is one possible process of transforming information into knowledge. In this case, the individual cognitive structure is expanded, however the structure itself remains the same. Cattell (1973) names this crystallized intelligence (see also Lohmann, 1989; Tuijnman & van der Kamp, 1992). Human beings thus develop a rich and highly differentiated knowledge, they optimize their knowledge structure through repeated application and practice, which can also described as proceduralize and decontextualize knowledge (Anderson, 1983; Oerter, 1985). This optimization of representation can be regarded as a process of testing, differentiating, and expanding active mental models. (Piaget, 1972; Gagné, 1971) Even if mental models are usually not associated with assimilation processes, one can describe such optimization processes as “fleshing out”. (Johnson-Laird, 1983, p. 452; Seel, Al-Diban, & Blumschein, 2000, p. 136; Seel, 2003, p. 197). This means that mental models, understood as evolving schemata, are – within the process of interaction with the surrounding world – tested and expanded until they can not be further optimized. The resulting well-functioning models are stored as successful cases, abstracted as schemata or in form of strategies for action. This is well illustrated by the approach of meaningful verbal learning by Ausubel (1968). In this approach learners are exposed to information, which can be integrated into the cognitive structures or they may expand them. These structures are strengthens by practicing. Schank and Abelson (1977) developed a script-oriented approach of memory, which also explains these processes well, particularly episodic knowledge. The conception of Goal-Based-Scenarios (GBS) (Schank, 1993, 2001) which takes up these ideas and provides important elements of model-centered-instruction will be presented later on in this chapter. Now we will take a look at the process complementing assimilation to which Piaget refers to as accommodation.

Accommodation Learning – Learning with Mental Models

Besides Piaget’s assimilation strategy there is the accommodation strategy which in a way contradicts with human laziness. The main idea is that an individual, when confronted with a new situation for which he or she has no matching schema, develops a new schema. However, how can this be done? Dietrich Dörners study on the “Logic of Failure” (Dörner, 2001) demonstrates typical errors human do in such situations. Usually the subjects lack a matching (viable; cf. von Glasersfeld, 1987) model for system relations. In this case, important effect relations cannot be identified, causal factors cannot be identified or are identified incorrectly, delays within the system cannot be recognized, or regulations are quantitatively inappropriate. It is – from a social-psychological perspective – an interesting phenomenon that individuals of one group delegate competences to their fellow group members in order to reduce their own responsibilities. Such behavior can have fatal consequences; according to Dörner this happened in the Chernobyl catastrophe (Dörner, 2001; Janis, 1972). In the same matter as assimilative competences belong to crystallized intelligence, accommodation is a part of fluid, context-free intelligence (Cattell, 1987; Lohmann, 1989; Tuijnman & van der Kamp, 1992). Usually, children show a high amount of such intelligence which is not surprising since they are permanently exposed to new situations and they have only little knowledge new situations can be linked to. The most important accommodation strategies are to find certain analogy models (cf. Gick & Holyoak, 1983) or to imitate action. In both cases the concept of mental models takes effect perfectly (cf. Gibbons, 2001).

Social Learning with Models

Bandura's theory of social learning is one classical example of learning psychology (Bandura, 1971). Bandura describes learning with models as a substantial form of how humans learn. Learning with models implies imitating or manipulating behavior of models (e.g. experts) which is then recognizable in other social conditions. Doing so, new schemata can be constructed and also be strengthened through practice. Learning through imitation illustrates the feedback effect of learning from a model-perspective. An observed behavior is imitated on the basis of one's own knowledge and thus a reaction from the environment is expected. Then the mental model will be adapted accordingly. Thus, learning takes place through generating a structure and strengthening of this structure on the basis of action and reaction. Accommodation processes can occur either through these social learning experiences where humans interact or through the examination of new challenges and situations. In the latter case, the regulative moment takes place through just the one learning person.

Learning with Perceptual and Thought Models

In a problem situation individuals usually stack. They lack essential schemata or proceed with inadequate ways to solve these problems. This leads to incorrect solutions and unsatisfying actions. Therefore the human mind is facing a cognitive crisis. Often this is even a substantial crisis. In this case, it is assumed that the individual generates a mental model of the problem in order to gain back his or her ability to act: Mental models are thus the self-esteem's fire brigade. This process of mental model construction aims at the construction of viable schemata. Craik (1943) introduced the term *internal model* which functions as a working model and mediates between the individual's own knowledge and the outside world (cf. Seel, 2001, p. 407). If an individual constructs mental models, it constructs reality, because whatever he or she perceives is perceived on the basis of his or her own knowledge. At this point the process of constructing mental models has to solve two problems at the same time. First, the human mind does not understand the world which means that it lacks a strategy to solve the problem. Second, the human mind is not able to generate a new strategy to solve the problem. Therefore mental model construction functions as a heuristic-tool-machine. One of the most important strategies of human problem solving and thus a heuristic strategy is finding analogies as well as the testing of new constructions. New structuring, linking up to prior knowledge and testing new knowledge are essential processes of the generation of new structures (cf. "fleshing out"). According to Stachowiak's general model theory (1973) we can distinguish between perceptual and thought models. Perceptual models are regarded as structure models of a fact (Johnson-Laird, 1987; Seel, 1991). They simplify a thinking process by producing an interface of the outside world and the cognitive structure. However, they are supposed to be static and picture-like and therefore a first individual interpretation of the environment. It is important to keep in mind that models shape just parts of reality (e.g. isomorphism, see figure below). As such they are utilitarian models. Thought models, in contrast, are simulations and processes of internal actions or thought experiments.

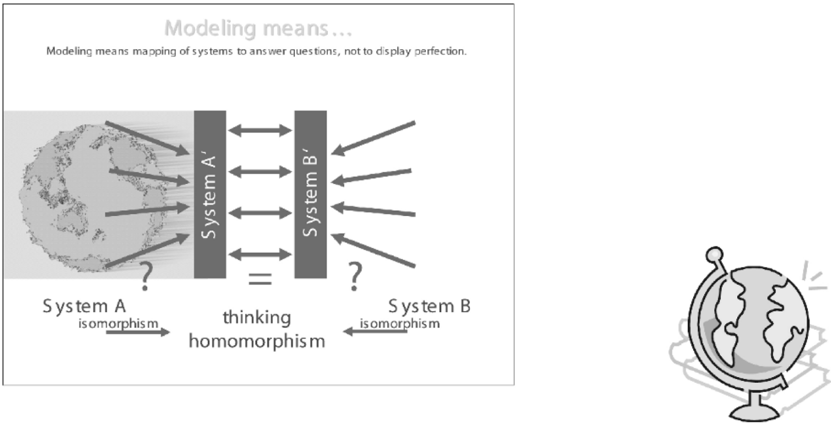

Fig. 1. A mental model approach to perception and cognition (Seel, 1991).

By contextualization of procedural abstract knowledge, for instance a globe to answer the question of what the earth looks like, a perceptual model is constructed. This model generation cannot function without representation systems. What the individual can see depends on what the individual knows. Therefore perception is highly dependent on sign-meaning relations. Seel refers to this with Piaget's term semiotic function (Piaget, 1972; Seel, 2003b, p. 60). Thus perceptual models are based on symbolic representations. Thought models, however, are dynamic. They are idiosyncratic to a high degree since they only consist of individual knowledge and interpretations of the world. A reference to the environment is possible but not necessary. In a sense, thought models built up perceptual models as a homomorphism. Naturally this does not always work out because this process is influenced by the inadequacy of the semiotic constellation of the individual's mind. Also isomorphism means not all of the elements of the origin is represented in the model, which makes it difficult to re-construct the reality. But this process is optimized to solve the given problem. And doing so homomorphism is the process working with analogies, comparing the elements of system a' and b'. From a learning theory perspective the question needs to be answered, how the process of representation can be enhanced through symbols in order to have more efficient model adaptations. In this context, one has to reflect about how interventions can be designed in order to provide useful scaffolds (Bransford, Zech, Schwartz, Baron, & Vye, 2000). Assimilation and accommodation processes are interacting processes which in the course of the model construction from the first mental model to the tightened schema may oscillate and finally merge.

"Mental models play a central and unifying role in representing objects, states of affaires, sequences of events, the way the world is, and the social and psychological actions of daily life. They enable individuals to make inferences and predictions, to understand phenomena, to decide what action to take and to control its execution, and, above all, to experience events by proxy." (Johnson-Laird, 1983, p. 397).

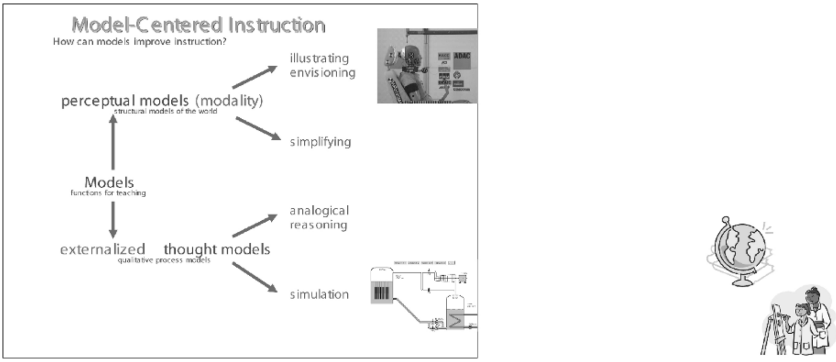

Fig. 2. Learning and Instruction from a model-centered perspective

Models, MOPs, and TOPs

First, one has to distinguish between short-term mental models and permanent models. The latter are for instance an individual's concepts of how the world works, such as the idea that school prepares for life. Permanent models are usually schematized concepts which might change over time. As such they are not mental models but schemata. These processes will be discussed in the next paragraph in more detail. Further-on, externalized, communicated models must be distinguished from (internal) mental models (short-term and permanent models) (see also figure 2). E.g. Concept-maps as presented by Al-Diban (in this book) belong to the group of externalized models. A mental model is a cognitive-psychological construct and cannot be provided from outside as a fact and the same. Mental models are idiosyncratic, accommodative and they mediate between the outside world and subjective knowledge. Consequently they are of short duration, they give meaning to the individual and they are not restricted to a primary form of representation. An instantiated working model might be transformed into a schema if it is strengthened through cognitive or physical practice. In this case, a Best-Practice model is represented which serves as an explanation model. At this point we have a connection to Schanks (1982) script theory. In his terminology he would call this a memory organization paket (MOP) (Schank, 1982; Kolodner, 1993, p. 61). MOPs can be understood as a kind of mega scripts. A visit at a restaurant which consists of various scenes and scripts is a classical example of a MOP. Besides scripts and MOPs he introduced the thematic organization points (TOP), which can be understood as objectives or structural patterns for complex behavior (see Figure 3). TOPs can also be regarded as abstract rules, adages, or wisdom (cf. long-term model). In this sense, they can be considered heuristic tools for mental models. Markman (1999) discusses the advantages of Schank's latest conception of scripts, scenes, MOPs, Meta-MOPs, and TOPs with the example of a train ride. The graphic below illustrates the relations. Over all, the script theory explains very detailed how human knowledge can be stored and recalled (cf. Ramirez, 1997). Schank developed on the basis of this theory his famous Goal-Based-Scenarios approach (GBS) (for examples see: http://www.engines4ed.org/hyperbook/nodes/-NODE-301-pg.html). Also Kolodner (1993, p. 70) took up the hierarchical concept of representation from Schank's "Dynamic Memory" (Schank, 1982) as a basis for the development of case-based learning environments, which are pretty close to GBS and Model-Centered-Instruction. But she also states that novices often are not good at case-based-learning since they lack prior experience to link up to (Kolodner, 1993, p. 105).

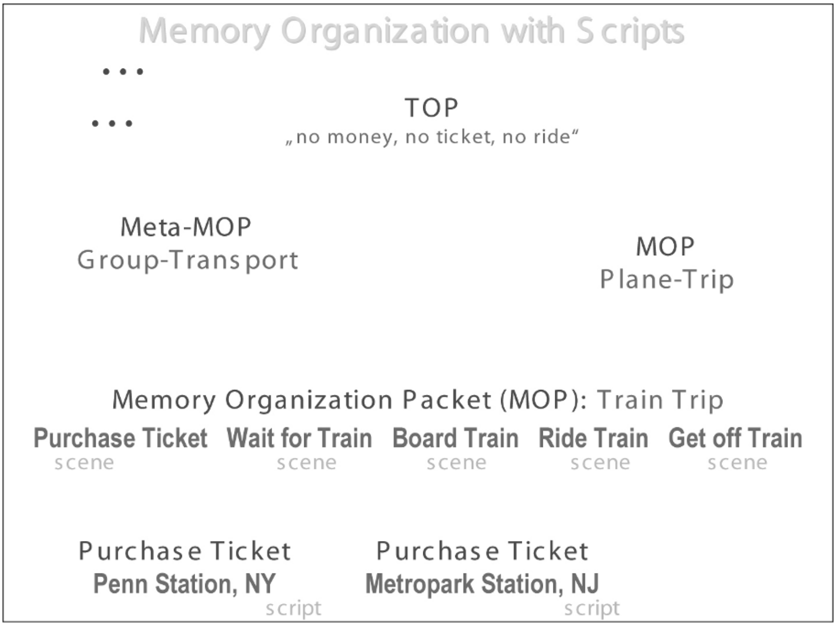

Fig. 3. The relationship of MOPs, TOPs, Scenes and Scripts (Schank, 1993; Markman, 1999; Ramirez ,1997)

However, the script approach is fairly static and hierarchal. One of the thesis of this chapter is that mental models can help to overcome this deficit. Mental models are the creative engine of human thinking when they produce scripts, MOPs, and TOPs and also recall them more rapidly and dynamically as scripts can do. Schank's indexing concept may only be a little help out of the dilemma (cf. Schank, 1993a, b). When applied to complex and dynamic systems with a permanently growing network of knowledge, this tool is too restricted and slow, especially because it is based on the same script concept. Kolodner (1993, p. 82) points out how models can be described and distributed through MOPs and therefore mentions the following advantages of MOP models: MOPs describe elements that functioning models must have, how these elements are interconnected, how models and instances can be integrated (concrete, abstract) and most importantly how models develop at all. In this context Kolodner states: "In particular, my own view of dynamic memory views MOPs, scenes, and TOPs as components of models of the everyday world" (Kolodner, 1993, p. 80).

Schank's script theory is of great importance for model-centered-learning and instruction bacause it provides a plausible construction of human memory as well as many points to link up to when dealing with the construction of models. Furthermore Goal-Based-Scenarios are great learning environments seen through the glases of MCI. Following this vein the question is how models develop over time.

The Progression in Model-Centered Learning

Snow (1990) discusses the learning progression as a path starting with a mental model in form of a pre-concept or misconcept to an causal explanation model that can be regarded as the final state of a learning process and thus as a settled knowledge of how to perform adequately. In doing so, he develops a framing concept that distinguishes motivational and self-regulating functions, learning strategies, declarative structures, and procedural skills. The central concern in this context is how mental models develop from their alleged wrong starting point to a stable explaining model and how these processes can be captured and controlled (cf. Seel et al., 2000, p. 131). Anderson (1983) describes this process as knowledge acquisition, and automation. Also Norman (1982) illustrates a developmental process of the learner that starts with the elimination of old, incorrect concepts and leads to the fine adjustment of the new concepts and its application. "Learning is thus seen as a progression across a range of simpler to more complex mental models of a domain, as well as progression in conceptual understanding." (Snow, 1990, p. 459). Decisive for the learning process to be successful is in Norman's as well as in Anderson's approach the depth of processing and its meta-cognitive control (Flavell, 1979; Seel, 2003, p. 226). Also it is important to keep in mind that learning does not build up on incorrect models; those have to be unlearned at first (Snow, 1990, p. 459). Thus, when learners are confronted with complex problem situations he or she has to be provided with new and well matching models which are plausible and explanatory to the learner. However, Snow (1990, p. 469) indicates that the evaluation methods and procedures must be expanded to generate reliable statements about the improvement of such learning processes. And this is crucial for the conception of interventional measurements and the design of learning environments (Snow, 1990, p. 469). The work of the Freiburg-group, in particular the contribution of Al-Diban (2003), Ifenthaler & Seel (2004), Ifenthaler (2006) and Pirnay-Dummer (2006) show possibilities of how the progression of the model construction can be evaluated and how such models can be identified on the basis of text material without provoking test effects (see the respective chapters in this book). The essential improvement of diagnostic strategies and − instruments is the most substantial answer to the question of how model-centered learning works. Jih and Reeves (1992) accused the mental model research of not generating any descriptive instructions for Instructional Design. Notwithstanding, Seel et al. (2000) demonstrated possibilities of how this can be done. Besides this research related problem, the question remains how learning environments should be designed to facilitate model construction as a central part of problem-based learning. One thing is a certain fact: pure information design and free access to any knowledge of the world will not help. Therefore Kolodner (1993) draws the conclusion that, in order to facilitate case-based learning, supportive systems have to be developed. In particular inexperienced learners need support. "The challenge in building such systems is to be able to represent, segment, and index cases appropriately so that they can be retrieved at appropriate times and presented to users in accessible ways." (Kolodner, 1993, p. 107).

Model-Centered-Instruction

A Model-Centered Approach to Instruction

As the first section of this article dealt with the epistemological foundations of thinking, model construction and modification, the following section takes a look at instruction from a model-centered perspective. With reference to Kluwe and Haider (1990) Seel provides a conceptual frame for the positioning of instructional interventions. This approach covers up Achtenhagen's approach (2001) to a high extend: "According to recent didactic research approaches two steps of didactic modeling should be taken into account (cf. Achtenhagen, 1992): (a) modelling reality, and (b) modelling models of reality under a didactic perspective." Gibbons (2001, and in this volume) deals with this issue too. He distinguishes three fundamental models from an instructional perspective: First, he mentions the environment itself; it can either be real or artificial and in mediated form. Second, he distinguishes systems as cause-effect-systems and knowledge systems, which have a lot in common with conceptual models. Finally, he mentions expert models, which to a high extend resemble the learning with models approach (Bandura's role model behaviour) and the approach of Cognitive Apprenticeship (Collins, Brown, & Newman, 1989). Since in accordance with the epistemological foundation, there is no direct access to the outside world, information is generated through individual processing of knowledge. In this way, also scientific and other models of the world are generated, seen as conceptual models reflecting the character of being abridging and goal-oriented (Stachoviak, 1973; Achtenhagen, 2001; Seel, 2003a). Achtenhagen (2001) refers to them as "original 2" as a referential basis for the learner's individual model construction. Usually this world is also referred to as "shared knowledge of a discipline" (Seel, 2003a, p. 67). From my point of view, this term is confusing for knowledge is communicated in the form of information. Communicated knowledge is not internal anymore but sent as information in representational forms of the outside world (Bateson, 1995). The instructional models have to mediate between the worlds. The whole crux of Instructional Design and curricular development is the genesis of this specific instructional world. In this context, Instructional Design at least faces three enormously complex problems.

Three Core Problems of Instructional Design

In a first step, the learner's conditions have to be assessed correctly. The word "correctly" holds the reflection of the psychological model of the whole process (cf. Seel 2003, p. 66). The Instructional Designer tries to gain a good understanding of the learner's mental model from the respective context of the given problem. How does the learner deal with the given problem? What does the learner already know and how does he or she know it? What strategies is he or she already familiar with, etc… In the context of the ADDIE-model these questions must be dealt with in the analytical phase. Another (2) important modelizing world is the ascertainment of the domain-specific knowledge or scientific concepts of the world (cf. "original 2" Achtenhagen, 2001, p. 365). This is another core competence of the Instructional Designer. Finally (3) the construction of mental models is a reflexive process, particularly in the sense of Kluwe and Haider (1990) who regard psychological models as models that are part of model construction related activities (Seel, 2003a). For the Instructional Designer these model construction activities are the most important reflection instrument. The following example can help to illustrate the point (see also figure below).

A German radio station switches its production technology from analogue to digital. The staff members are technically trained, of middle age and have only little computer skills. A training center develops a training concept for a 2-day training for the staff members. After the first training procedure the trainers realize that

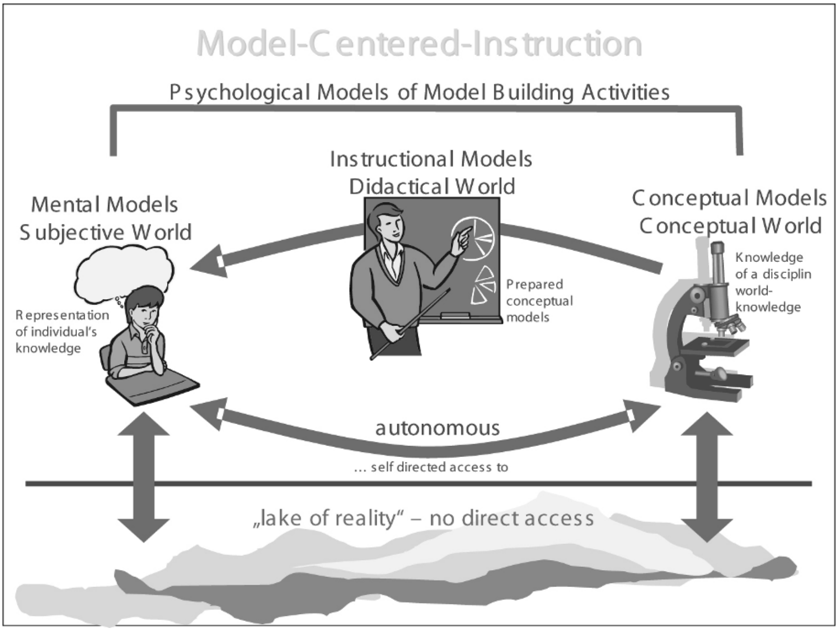

Fig. 4. A model perspective to learning, instruction and reflection.

it is too much for the participants. Also they realize that 2 days of training is not enough to cover the complexity of the topics to be learned. However, there are more complex circumstances, this short sketch shows some typical problems. First, the participants' relevant prior knowledge was assessed only in an indirect and static way. Not at least for financial reasons, a performance and task analysis hasn't been conducted. Also the job description of the typical work processes was only done with respect to expert knowledge. As a matter of fact a skill training of the staff within two days is almost impossible anyway. This should have been kept in mind when chosen a strategy. Perhaps it would have been a good idea to evaluate the staff members' problem solving strategies at first and then design a 2-day training. If an IDler has a model of how learners solve problems – and mental models are the most essential strategies of solving problems – a trainer can impart problem solving strategies in the new domain within two days. In this context, a stable (mental) model of the expert domain is fundamental in order to determine the most important requirements for self-directed learning, which is essential in this case. A psychological model of the relation of the learner's model, the ID model, and the domain-specific model implies the problem that the learners have to have computer skills to be able to act successfully in their new domain. For the learners it's a crucial point to learn related problem solving strategies and generate analogy models for the new defiances. This implies to control their own models of learning and instructional procedures. The instructional designer has to take into account such processes when planning a instructional intervention (see figure above).

The Macro Design of Learning Environments

With reference to Johnson-Laird (1983, 1989) we have distinguished three different forms of learning; accordingly we differentiate three macro-designs for learning environments (cf. Seel et al., 2000). (1) Learning takes place autodidactically e.g. by means of analogical reasoning. This happens in a free learning environment which can hardly be called an instructional environment, as there is only self-directed learning. For this case the information technologist's argument holds true, we must acknowledge that there is almost no use of ID (cf. Duchastel, 1999). However, there is slight hope for the profession of ID in the fact that all information has to be designed and developed, what Instructional Designers would be good for. (2) A second version states that humans learn by means of observation and then develop mental models. This goes along with our idea of guided discovery learning, thus we design learning environments with means of scaffolding or an integrated companion (cf. Gibbons, 2001). Gibbons' learning environments include scaffolding as a general support system. This can either be in the form of a tutor, a teacher or an intelligent software system. However, it always aims at supporting the learner just-in-time. Also goal-based-scenarios as presented by Schank et al. (2001) and Anchored Instruction (Bransford et al., 2000; cf. Blumschein, 2003) are excellent examples of guided discovery learning. Both approaches are presented in much detail and up-to-date in the volume of Forbus and Feltovich "Smart Machines in Education" 2001. (3) A third version of learning with mental models is presented by Johnson-Laird (1989) with explanations by other people. This goes along with conventional teacher-centered-instruction. A classical example for instruction is the cognitive-psychological-based concept of Ausubel known as meaningful verbal learning (Ausubel, 1968). Also the well-known approach of Cognitive Apprenticeship (CA) developed by Collins, Brown and Newman (1989) can be subsumed under this form of learning. Somehow, CA is a reinterpretation of the approaches of Palincsar and Brown (1984) reciprocal teaching of reading, Scardamalia and Bereiter (1984) procedural facilitation of writing and Schoenfelds (1983) mathematical problem solving. All of them focus on expert behavior and –skills to start with in the learning process. In the 1990th the research group of Seel (cf. Al-Diban in this edition; cf. Seel et al., 2000; cf. Seel et al., 1998) developed a learning environment for an economic context (4M: Multimedia and mental models) for applying the Cognitive Apprenticeship approach for its first time in its total range. The complex economic relations of the European Monetary Union were included into a multimedia learning system. One major research question with that was, if the CA is a useful and applicable system for model-centered learning with multimedia. Surprisingly this was the first realization in Educational Science testing this approach at all. With different evaluation measures the learners' mental models were assessed and compared. Within the multi-media environment the learners' path was facilitated by various didactic means. Among other supporting instruments the system provided a conceptual model giving orientation for the individual problem solving process (cf. Snow, 1990; Seel et al., 2000). In the beginning of 2000 Seel's group (Seel, 2005) researched another learning environment according to model-centered discovery learning called "The Forest as an Ecological System" which had included as a major feature a model-building kit (MoBuKi). The students' task was to construct an "explanatory model" of the phenomenon in question. Therefore the learning environment provided several tools to help students to construct analogical models, which were tested (Seel, 2005, p. 83). Lately we developed the learning environment BIRD for Far-East-Asian scholarship holders in the field of biotechnology in this vein.

BIRD – Problem-Based Learning

BIRD is a learning environment we designed that refers with some extend to the concept of goal-based-scenarios (Schank, 2001; Blumschein, Liebhardt, & Kessler, 2002) and MCI. The DVD was developed to prepare Far-East-Asian students of biotechnology receiving a scholarship for staying in Germany. If the scholarship holders are accepted by INWENT (society for international capacity building of Germany) and the Helmholtz Center for Infection research, they complete an extensive training program in Germany (http://www.helmholtz-hzi.de/en/; http://www.inwent.org/ueber_inwent/kurzprofil/wer/index.en.shtml). Scholarship holders complete periods of practical work in biotechnological companies and they are further educated in their discipline by the Helmholtz Center for infection research in Braunschweig. The major goal of these trainings is to qualify the students for independent scientific and economic autonomy in their home country. Therefore the learning software serves the preparation of the candidates in advance in their home countries. The students' prior knowledge differs enormously. Therefore the institutions involved focused on a competence-based didactic approach (Fischer & Blumschein, 2007). This goes along with the overall concept of INWENT, which intends the reinforcement of qualified as well as executive employees' competence to act in political, organizational and operational modification processes (www.inwent.de).

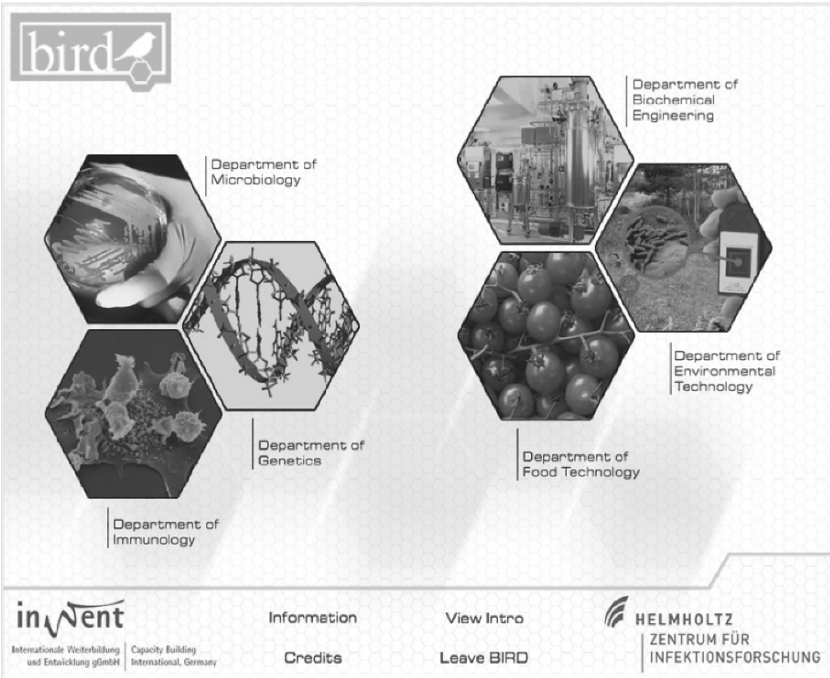

Fig. 5. Introduction page of mission 1, production of lactose free milk.

Figure 5 depicts a screen showing buttons that offer the learner several types of support, including lab tests, ask an expert, ask a colleague, visit the library, make notes, or visit the collaborative Web space. The center of the screen in Figure 5 shows a text version of the video, which provides the missions to be solved. A first usability test has shown that the users can manage the learning surface quite well. The problems are authentic and comprehensible. Also the facilitating means seem to be appropriate. However, from our point of view the major challenges for further work lie in the rather demanding implementation of the (intelligent) Counselor. So far "it" becomes active only if contacted directly by the user or if the user has not shown any activity in a certain time. For this track Gibbons (2001), and Biswas and Schwartz (Blair, Schwartz, Biswas & Leelawong, 2006) present interesting solutions to us; however, these require a lot of programming.

Betty and the Companion

Gibbons (2001) introduced the Companion as a mediator into learning environments. The Companion is seen as a supporter of learning; he or she mediates between the learners' mental models in consideration of their progression and the conceptual models of the problem to be solved. In software-based learning systems for instance he or she is an intelligent tutor; also he or she can take up the role of a human teacher if the intention is to support learning and to mediate – not if the intention is to teach.

"The companion's functionality has many embodiments: peers, advanced peers, teachers, living experts, books, expert computer programs, and other sources capable of arranging, commenting on, supporting, interpreting, or otherwise augmenting experience." (Gibbons, 2001, p. 519).

Biswas and Schwartz (Blair et al., 2006) developed with regard to the research on Anchored Instruction a concept called Teachable Agents (cf. Blumschein, 2003) that function in accordance to the same principles. They support the learner in testing his or her new knowledge in the problem solving situation by demanding to give the agent a detailed explanation of the given task. "Betty" is the latest product of the group of Biswas and Schwartz (Crownover, 2007). "Betty's" knowledge is illustrated in form of a concept-map. The aim is that "Betty" manages to solve the task. Thus "Betty's" behavior provides the students with feedback to his or her own learning process. "By analyzing what they have taught her in a graphic form and observing how Betty uses this information to answer questions, children are assisted in monitoring their own learning process." (Crownover, 2007, p. 4). As a theoretical foundation Biswas and Schwartz (Blair et al., 2006) developed 4 design principles for the Teachable Agents.

- explicit and well-structured visual representations
- enable the agent to take independent actions
- model productive learner behaviors
- include environments that support teaching interactions

Also Seel and colleagues incorporated a "lucky mascot" into their economic systems to support the learner to grasp his or her concept (Seel et al., 2000, p. 142; Al-Diban in this edition). Schank's storytelling concept essentially goes along with this idea of mediating (cf. Burke, 1998). "The person in the story who is playing a role similar to the student's is an important figure for determining the kind of analogy the telling of a story will create. This person is called the student analogue." (Burke, 1998, p. 226). The mediating interface is termed SPIEL: Story producer for interactive Learning. SPIEL was for instance used in the popular GBS YELLO, where communicative competences are trained (Burke, 1998; Schank, 1993a, b). Also Hanke (in this edition) presents a supporting system following the model-centered approach. She assumes that the teacher initially guides the student to the problem, provokes him or her, supports him or her to activate the necessary prior knowledge, provides additional new information and facilitates through appropriate didactic support the consolidation of new knowledge. This phase marks the transfer from a short-term mental model to a stable schema. Even if the mentioned elements of interactive learning environments do not have much in common, they all show how important the mediation between the given state of affairs of the learner (cognitive, emotional, motivational, and meta-cognitive) and the problem situation given in the learning environment is. It almost appears as if the preparation of the material to be learned in its degrees of illustration (cf. Aebli, 1963; Bruner et al., 1971) had fallen behind the matter of designing the interaction between the learner and the given problem. Taking the epistemological postulations seriously, this is not surprising at all (cf. Spector, 2001). Learning is not about creating a "brave new world" but about the ability to act successfully in a demanding and complex world. These mediators thus do not model the content, but support the learner in constructing a mental model of the problem to be solved. That means they aid learners' self-directed learning competencies. This is the important track we have to keep up with even stronger. And so we should even enforce the discussion whether or not there is a future for Instructional Design in the Information Age (cf. Myers, 1999). So ID has to bear the challenge of designing environments to elicit demanding learning-activities according to well defined goals considering concise learner's preconditions. It's hard to imagine how "information design" will charge that goal especially with respect to the unfolded aspects in this chapter. Nonetheless, so far ID didn't bore the palm. Far from it! ID will never make the race when continuing to focus solely on the classic matter of designing contents. But already Collins, Brown and Newman (1989) have shown that the design of the interface is the critical point, and so did blaze the trail for ID of the near future. From a technical point of view we still have to deal with the demanding efforts of software-based realization of intelligent agents and facilitating systems. Even Leutner (2004) states about that: "[…] all those adaptive learning aids that cannot be defined and pre-programmed based on a clear rule system have to be supplied by real-human instead of virtual computer learning assistants (Astleitner & Leutner, 1994)." Unfortunately this is grist for the mill of the antagonists: complex software-based implementations still slow down ISD (cf. Clark, 2004).

Model-Centered Instructional Design

A Model of Instructional Design

From a model-centered perspective it is recommended to dissociate one self from a linear ISD model. Most likely the founders have never had the idea of linear ISD anyway (cf. Dick, 1997). However it is obviously the more convenient way to deal with linear models than choosing the challenge of a high degree of complexity. On the other hand, linear models can be seen as pragmatic solutions that run the business and aim at success, both shouldn't be disclaimed at all. In Merrill's Pebble-in-the-pond approach (Merrill, 2002a) for instance the problem task is in the focus of starting the ISD process. After the problems have been pointed out the progression of such problems of increasing difficulty or complexity which both determines the structure of learning is analyzed. In a next step the necessary knowledge is analyzed. A following step includes reflection about instructional strategies and the implementation (Merrill, 2002a). The essential point in Merrill's modification of his ISD strategy is the much greater emphasis of the problem. Nevertheless the approach is very knowledge-centered, linear, and prescriptive. Gibbons (2001) broadens the way of looking at this significantly by focusing on the analysis and reflection of the central challenge as the initial part of the ISD process and by dealing with modeling the problem at the learner's side as well as in the learning environment. Consequently, he sees Instructional Design from a model-centered perspective as an analysis of different design layers. Accordingly, he shows in his contribution in this edition how a content-centered perspective determines the design of the various layers of action.

Tennyson (1997) presents a systemic approach to Instructional Design. He points out that in the recent decades many ISD approaches have been developed that were specific to certain problems and therefore made the decision for the one right approach more and more difficult. "Proposed in the system dynamics approach is that the richness of ISD can be applied when viewing instructional design from a nonlinear, dynamic systems approach." (Tennyson, 1997. p. 415). In this vein Tennyson names six components of the dynamic ISD that control the quality of the design process:

- situational evaluation (strategy planning and analysis)
- Design (process of analysis, design)
- Production (development and evaluation, distribution)
- Maintenance (care, update, distribution)
- Implementation (implementation, distribution, evaluation)
- Foundation (theory reference)

These overlapping fields dynamically cover up the famous components of the ADDIE approach. So he states: "The view I represent is that instructional design is a nonlinear system that dynamically adapts to the problem conditions of a given situation […]." (Tennyson, 1997. p. 414). With this systemic approach Tennyson (1997) manages to more clearly take up the overlaps within the ADDIE concept as there are for instance in the field of design and development. His model of ISD is relatively unsusceptible to the instructional designer's degree of expertise and the common learning theory or the nature of the problem. Herewith Tennyson provides a model for a dynamic ISD that integrates the necessary and founded components of the ADDIE approach but that is nevertheless non-linear (see figure 6).

The Rapid Prototyping approach (cf. Tripp & Bichelmeyer, 1990) shows at least partly a pragmatic implementation of that. Particularly when dealing with enormous projects one proceeds correspondingly to industrial large-scale production: first a prototype is developed and tested, and then the remaining teaching

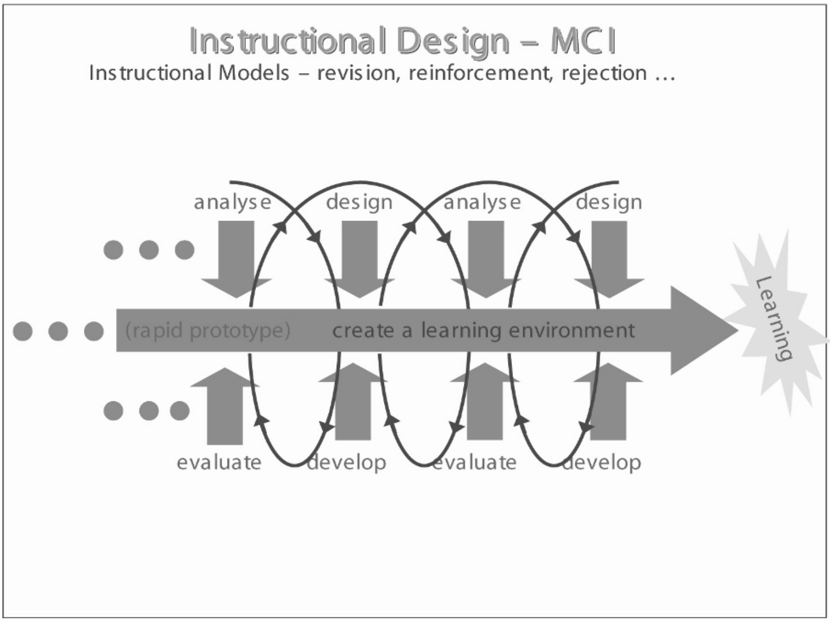

Fig. 6. A systemic perspective of model-centered-instructional design.

objectives are integrated and implemented. This approach is pragmatic since it integrates at an early stage a first "trial and error" into the process. The prototype then serves as a concrete ISD model. MCI-design however, has to reach even further. When adapting the concepts of model-centered learning and thinking and the concept of further development of a "mental model" in terms of "Fleshing-out" (Seel, 2003; Snow, 1990) this might come true. MCI does not only focus on the matters of how external models can be developed and instructionally optimized (cf. Gibbons, 2001) to support the above mentioned sophisticated learning process through mental model construction. MCI rather applies this theory including its epistemological foundations on all its layers. That for this approach is very complex; however, it offers expanded fields of decision making and flexibility in its implementation. Herewith virtually all criticism of ID and ISD can be refuted. And as another benefit the use of automation can absorb this complexity to a high extend (e.g. MITOCAR by Pirnay-Dummer 2006; DEEP by Spector; SMD by Ifenthaler). Taking all this into account one can dare to tackle the following thread of MCI-Design: The ISD-process starts out with the Instructional Designer choosing an ISD model. Then a first situational evaluation starts (Tennyson, 1997; Kaufman, 2003). In this process, the ISD-expert reinforces, modifies or rejects his or her working model (cf. Penner, 2001; Buckley & Boulter, 1999; Snow, 1990).

This activity is highly complex since he or she has to deal with different subsystems: his or her mental models of the world (world 1), conceptual models of the domain specific scientific discipline (world 2), instructional models (world 4), the learner's mental models (world 3) and psychological models of the ID process (world 5). A detailed description of a systemic-strategic perspective of ISD that can be joined with the aforementioned approach has been presented e.g. by Rothwell and Kazanas (2004) and by Morrison, Ross, and Kemp (2004) and with a slightly different focus also by Kaufman (2003). Also, it must be mentioned that the individual's subsystems can change over the course of the process, not only as a result of the ISD-process but also by impacts of external factors since every single model can change over time. Also the acceleration of the ISD-process that has been demanded is a great challenge which in most cases does not allow to linearly working off single steps in terms of ADDIE. Consequently, there is agreement that the analytical phase is the most important part of the ISD-process. Staying on this track we can uncouthly conclude that the whole ISD takes place in the analysis phase. Therefore the development of instruments to guarantee an efficient (fast and precise) analysis of the various conditions of the learners, expert knowledge (expertise) and Instructional Design expertise will be the greatest challenges of the near future in order to win the race for time and dynamics in the era of the information society.

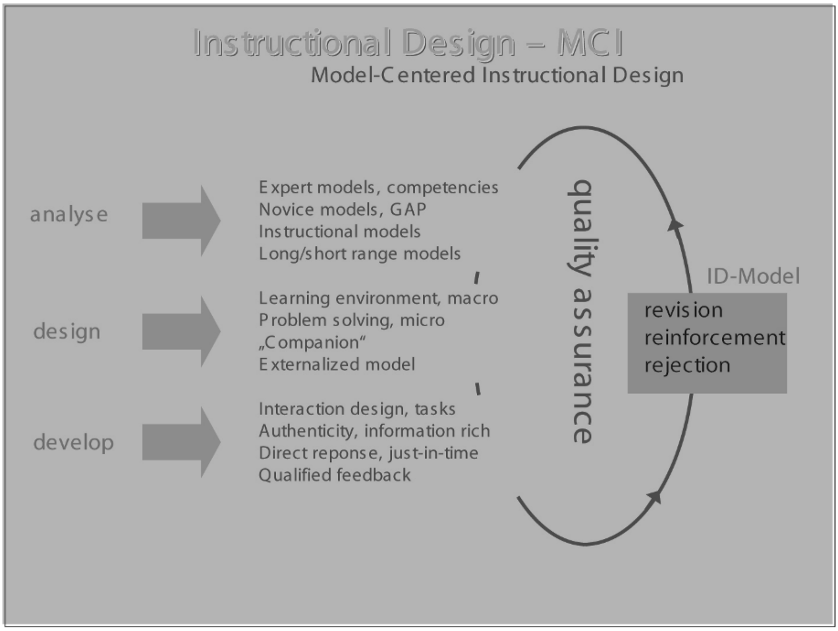

Fig. 7. A model-centered instructional design perspective under recursion of the ADDIE approach.

Model-Centered Instructional Design – a Perspective

Model-Centered learning and instruction is a much promising construct for further research and developments within the scope of Instructional Design (see figure above). MCI offers Instructional Design in many respects new ways, not least of all because it rearranges learning and instruction in all its elements. Possibly, MCI heralds a paradigm shift within educational science which may not be seen at first sight. But, in order to meet its demands, this shift must affect all parts of educational science. The following is the thesis: MCI offers more appropriate patterns of explanation and more effective and efficient rules of design. Critics will accuse this approach of throwing away the results of 100 years of learning and instruction research in favour of a great venture. In 1980 Donald Norman marked with his article *Twelve Issues for Cognitive Science* a central point of reflection within psychologically-oriented educational science. Back then as well as today central questions have not been answered. For MCI these questions must be discussed with regard to its four aforementioned core fields: epistemology, expertise, learning and instruction. From these elements of MCI result the following challenges for the future:

- A central starting point of all research on MCI is the learning human being. We have to gain more insight into its complex nature in order to be able to develop adequate learning environments. Classical categories of knowledge are just one aspect of this. Motivation, emotion, the human social nature and the strong dynamics in the human interpretation of factual and social relations are still an extensive and enormously high-demanding field of research.
- The classical understanding of expertise is currently changing. Expertise is changing more and more rapidly. Therefore expertise must be constantly redefined. Today's expert is tomorrow's novice. Objective orientations towards "the" expertise will be more likely to fail in the future.
- Typical learning biographies change. Life-long learning is becoming more and more important. Consequently, attitudes towards learning and instruction change. Learning and instruction, work and off-time, research and development all will disintegrate.
- Social interaction is expanded through new ways of communication and work. The social human being is a new character.
- Locations of learning and media disintegrate in space and time. Thus it becomes more and more difficult to grasp their design. Learning becomes a hologram. Creating locations for learning means a new challenge.
- Interaction is the key word for the future. How does it happen, where, between whom and at what time? Who controls it? Who designs it?
- The design of software-based systems can not yet meet the demands or its development means yet too much effort.
- ADDIE is not a linear but a parallel dynamic process. Also learning, testing and giving feed-back are no independent events.

MCI has to take up and productively implement these challenges. This can be done by means of a reflexive model perspective of the learning system which obviously we must continue to work on. A model theory is, for its overall openness (terms of representation, dimensions, and systems) the first choice. Indeed, the insights in terms of mental models and model-centered learning are still in their infancy, albeit Norbert M. Seel's work on mental models in 1991 set a cornerstone in the discipline. However, analytical procedures and tools are constantly further developed so that we gain more exact information about prerequisites and learning procedures, The reflexivity of the systems of the learner, the instructor, the learning environment, the environment for application, the living space, Instructional Design as well as science and research can be made manageable through a model theory. Also in this respect, there is still much left that need to be developed. Nevertheless, outstanding quality in design and development of exciting learning environments and research approaches with excellent results have been already created. Most of all the groups of Schank, Bransford, Pellegrino, Penner, Schauble, Seel, and many others must be mentioned in this context (for an overview see: Forbus et al., 2001; Cobb et al., 2000; Lajoie, 2000; Seel et al., 2004). A process of change has thus already been heralded.

References

Achtenhagen, F. (2001). Criteria for the development of complex teaching-learning environments. *Instructional Science, 29*, 361-380.

Aebli, H. (1963). *Über die geistige Entwicklung des Kindes*. Stuttgart: Klett.

Al-Diban, S. (2002). *Diagnose mentaler Modelle*. Hamburg: Kovac.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, Massachusetts, London, England: Harvard University Press.

Andrews, D. H. & Goodson, L. A. (1980). A comparative analysis of models of instructional design. *Journal of Instructional Development, 3*(4), 2-16.

Astleitner, H. & Leutner, D. (1994). Computer im Unterricht und Ausbildung: Neue Anforderungen an Lehrer und Trainer? *Zeitschrift für Pädagogik, 30,* 647-664.

Ausubel, D.P. (1968). *Education psychology: A cognitive view*. New York: Holt, Rinehardt & Winston.

Bandura, A. (1971). *Social learning theory*. Morristown, NJ: General Learning Pr.

Bateson, G. (1995). *Geist und Natur. Eine notwendige Einheit*. Frankfurt: Suhrkamp Wissenschaft.

Blair, K., Schwartz, D., Biswas, G. & Leelawong, K. (2006). Pedagogical agents for learning by teaching: teachable agents. *Educational Technology & Society, Special Issue on Pedagogical Agents*.

Blumschein, P., Liebhardt, H. & Kessler, B. (2002). Docs'n Drugs – die virtuelle Poliklinik: Instruktionsdesign fallbasierter Lernszenarien. In J. Bernauer, M. R. Fischer, F. J. Leven, F. Puppe & M. Weber (Ed.), *Rechnergestützte Lehr- und Lernsysteme in der Medizin.* Aachen: Shaker Verlag.

Blumschein, P. (2003). *Eine Metaanalyse zur Effektivität multimedialen Lernens am Beispiel der Anchored Instruction.* Freiburg i. Br., Univ., Diss., 2004.

Blumschein, P. & Fischer, M. (2007). *E-Learning in der betrieblichen Bildung: Didaktisches Design von E-Learningmaßnahmen*. Mannheim: INWENT.

Bransford, J. D., Zech, L., Schwartz, D., Barron, B. & Vye, N. (2000). Designs for environments that invite and sustain mathematical thinking. In P. Cobb, E. Yackel & K. McClain (Ed.), *Symbolizing and communication in mathematics classrooms. Perspectives on discourse, tools, and instructional design* (pp. 275-324). Mahwah, NJ, USA: Lawrence Erlbaum Ass.
Bransford, J. D. (2004). *How people learn: brain, mind, experience, and school*. Washington, DC: National Academy Press.
Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review, 31*, 21-32.
Bruner, J. S. (1966). *Towards a theory of instruction*. Cambridge, Massachusetts: Harvard University Press.
Bruner, J. S., Olver, R. R. & Greenfield, P. M. (1971). *Studien zur kognitiven Entwicklung: eine kooperative Untersuchung am "Center for Cognitive Studies" der Harvard-Universität*. Stuttgart: Klett.
Buckley, B. & Boulter, C. (1999). Analysis of representation in model-based teaching and learning in science. In R. Paton & I. Neilsen (Ed.), *Visual representations and interpretations* (pp. 289-294). London: Springer.
Burke, P. A., Etnier, J. L. & Sullivan Howard, J. (1998). *Navigational aids and learner control in hypermedia instructional programs 2*(18), 182-196.
Case, R. (1987). *Intellectual development: birth to adulthood*. Orlando [et al]: Academic Pr.
Cattell, R. B. (1987). *Intelligence: its structure, growth and action*. Amsterdam [et al]: North-Holland.
Chi, M. T., Glaser, R. & Farr, M. J. (1988). *The nature of expertise*. Hillsdale, NJ [et al]: Lawrence Erlbaum Assoc.
Clark, D. (2004a). *The attack on ISD – 2000. Comments*. (Retrieved July 15, 2007) http://www.nwlink.com/~donclark/history_isd/attack.html.
Clark, D. (2004b). *A hard look at ISD – 2002. Comments*. (Retrieved July 15, 2007) http://www.nwlink.com/~donclark/history_isd/look.html.
Cobb, P. (2000). *Symbolizing and communicating in mathematics classrooms: perspectives on discourse, tools, and instructional design*. London: Lawrence Erlbaum Associates.
Cognition and Technology Group at Vanderbilt (1997). *The Jasper project. Lessons in curriculum, instruction, assessment, and professional development*. Mahwah, New Jersey, USA: Lawrence Erlbaum.
Collins, A., Brown, J. S. & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing and mathematics. In L. B. Resnick (Ed.), *Knowing, learning and instruction* (pp. 453-494). Hillsdale, New Jersey: Lawrence Erlbaum Ass.
Craik, K. J. W. (1943). *The Nature of explanation*. Cambridge: Univ. Pr.
Crews, T., Biswas, G., Goldman, S. & Bransford, J. (1997). Anchored interactive learning environments. *International Journal of AI in Education, 8* (142-178).
Crownover, A. (2007). What's in Betty's brain? *Eye on the LSI, 4*(2), 4-5.
de Koning, K. & Bredeweg, B. (2001). Exploiting model-based reasoning in educational systems: Illuminating the learner modeling problem. In K. D. Forbus & P. J. Feltovich (Ed.), *Smart machines in education* (pp. 299-330). Menlo Park, CA; Cambridge, Mass., USA; London, UK.: AAAI Press/ MIT Press.
Dick, W. & Carey, L. (1990). *The systematic design of instruction*. New York: Harper Collins.
Dick, W. (1997). A model for the systematic design of instruction. In R. D. Tennyson, F. Schott, N. M. Seel & S. Dijkstra (Ed.), *Instructional design. International perspective. Volume 1*: *Theory, research, and models* (pp. 361-370). Mahwah, New Jersey: Lawrence Erlbaum - Associates.
Dörner, D. (2001). *The logic of failure: recognizing and avoiding error in complex situations*. Cambridge, Mass.: Perseus Books.
Duchastel, P. (1999). *Instructional design in the information age. IFETS Forum. Paper presented as the prediscussion paper for the May 17-28th forum*. (Retrieved July 10, 2007): http://ifets.gmd.de/discussions/discuss_may99.html.
Duffy, T. M., Lowyck, J. & Jonassen, D. H. (1993). *Designing environments for constructive learning*. Berlin, Heidelberg, New York et al: Springer-Verlag.

Flavell, J. H. (1979). Metacognition and cognitive monitoring. *American Psychologist, 34*, 906-911.

Forbus, K. D. & Feltovich, P. J. (2001). *Smart machines in education*. Menlo Park, CA; Cambridge, Mass., USA; London, UK.: AAAI Press/MIT Press.

Gagné, R. (1971). *The conditions of learning*. London: Holt, Rinehart & Winston.

Gagné, R. M. & Briggs, L. J. (1979). *Principles of instructional design 2nd ed.* New York, Chicago, San Francisco et al: Holt, Rinehart and Winston.

Gagné, R. M. (2005). *Principles of instructional design*. Belmont, Calif. [et al]: Thomson Wadsworth.

Gayeski, D. M. (1998). Out-of-the-box instructional design. *Training and Development, 52*(4), 36-40. (Retrieved July 10, 2007) http://www.dgayeski.com/t&disd.html.

Gerstenmaier, J. & Mandl, H. (1995). Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik, 41*, 867-888.

Gibbons, A. S. (2001). Model-centered instruction. *Journal of Structural Learning and Intelligent Systems, 14*(4), 511-540.

Gick, M. & Holyoak, K. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.

Glasersfeld, E. V. (1987). *Wissen, Sprache und Wirklichkeit. Arbeiten zum radikalen Konstruktivismus*. Braunschweig: Vieweg.

Gordon, J. & Zemke, R. (2000). The attack on ISD. *Training Magazine, 37*(4), 42-53.

Hardré, P. L., Ge, X. & Thomas, M. K. (2005). Toward a model of development for instructional design expertise. *Educational Technology, Jan/Feb*, 53-57.

Ifenthaler, D. (2006). *Diagnose lernabhängiger Veränderung mentaler Modelle: Entwicklung der SMD-Technologie als methodologisches Verfahren zur relationalen, strukturellen und semantischen Analyse individueller Modellkonstruktionen.* Freiburg i. Br., Univ., Diss.

Janis, I. L. (1972). *Victims of groupthink: a psychological study of foreign policy decisions and fiascos*. Boston [et al]: Houghton Mifflin.

Jih, H. & Reeves, T. (1992). Mental models: A research focus for interactive learning systems. *Educational Technology Research and Development, 40*(3), 39-53.

Johnson-Laird, P. N. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness*. Cambridge, Mass. [et al]: Harvard Univ. Pr.

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

Kaufman, R. (2003). *Strategic planning for success: aligning people, performance, and payoffs*. San Francisco: Jossey-Bass/Pfeiffer.

Kluwe, R. & Haider, H. (1990). Modelle zur internen Repräsentation komplexer technischer Systeme. *Sprache und Kognition, 9*(4), 173-192.

Kolodner, J. L. (1993). From natural language understanding to case-based reasoning and beyond. In R. C. Schank & E. Langer (Ed.), *Beliefs, reasoning, and decision making* (pp. 55-110). Hillsdale NJ: Lawrence Erlbaum Ass.

Lajoie, S. P. (2000). *Computers as cognitive tools. No more walls*. Mahwah, NJ, USA: Lawrence Erlbaum.

Leutner, D. (2004). Instructional design principles for adaptivity in open learning environments. In N. M. Seel & S. Dijkstra (Ed.), *Curriculum, plans, and processes in instructional design. International perspectives* (pp. 289-308). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Lohmann, D. F. (1989). Human intelligence: An introduction to advances in theory and resarch. *Review of Educational Research, 59*(4), 333-373.

Lowyck, J. & Elen, J. (1991). Wandel in der theoretischen Fundierung des Instruktionsdesigns. *Unterrichtswissenschaft, 19*(3), 218-237.

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

Merrill, D. M. (2001a). *A knowledge object and mental model approach to a physics lesson*.

Merrill, D. M. (2001b). Components of instruction toward a theoretical tool for instructional design. *Instructional Science, 29*, 291-310.

Merrill, D. M. (2002a). A pebble-in-the-pond model for instructional design. *Performance Improvement, 41*(7), 39-44.

Merrill, D. M. (2002b). First principles of instructional. *Educational Technology Research and Development, 50*(3), 43-59.

Morrison, G. R., Ross, S. M. & Kemp, J. E. (2004). *Designing effective instruction*. Hoboken, NJ: Wiley.

Myers, K. L. (1999). Is there a place for instructional design in the information age? *Educational Technology, 1999*(6), 50-53.

Norman, D. A. (1980). Twelve issues for cognitive science. *Cognitive Science, 4*(1), 1-32.

Norman, D. A. (1982). *Learning and memory*. San Francisco: Freeman.

Oerter, R. (1985). Die Formung von Kognition und Motivation durch Schule: Wie Schule auf das Leben vorbereitet. *Unterrichtswissenschaft* (12), 203-219.

Palincsar, A. & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. *Cognition and Instruction, 1*, 117-175.

Penner, D. E. (2001). Cognition, computers, and synthetic science: Building knowledge and meaning through modeling. *Review of Research in Education, 25*, 1-35.

Piaget, J. & Inhelder, B. (1972). *Die Psychologie des Kindes*. Freiburg i. Br.: Walter.

Pirnay-Dummer, P. N. (2006). *Expertise und Modellbildung – MITOCAR. (Doctoral dissertation, University of* Freiburg i. Br., 2006).

Ramirez, C. (1997). *Schemata, frames, and dynamic memory structures*. (Retrieved July 20, 2007) http://www.cs.kent.ac.uk/pubs/1997/221/index.html.

Reigeluth, C. M. (1987). Lesson blueprints based on the elaboration theory of instruction. In C. M. Reigeluth(Ed.), *Instructional theories in action* (pp. 245-288). Hillsdale, New Jersey: Lawrence Erlbaum Ass.

Ritchie, D. & Earnest, J. (1999). The future of instructional design: Results of a Delphi study. *Educational Technology*, 35-42.

Rothwell, W. J. & Kazanas, H. C. (2004). *Mastering the instructional design process: a systematic approach*. San Francisco, CA, USA: Pfeiffer.

Scardamalia, M., Bereiter, C. & Steinbach, R. (1984). Teachability of reflective processes on written composition. *Cognitive Science, 8*, 173-190.

Schank, R. C. & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: an inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.

Schank, R. C. (1982). *Dynamic memory: a theory of reminding and learning in computers and people*. Cambridge [et al]: Cambridge Univ. Pr.

Schank, R., Fano, A., Bell, B. & Jona, M. (1993/1994). The design of goal-based-scenarios. *The Journal of the Learning Sciences. 3*(4), 305-345.

Schank, R. C. (1993a). Goal-based scenarios: A radical look at education. *The Journal of the Learning Sciences, 3*(4), 429-453.

Schank, R. C. (1993b). Goal-based-scenarios. In R. C. Schank & E. Langer (Ed.), *Beliefs, reasoning, and decision making: psycho-logic in honour of Bob Abelson* (pp. 1-32). Hillsdale NJ: Lawrence Erlbaum Ass.

Schank, R. & Neaman, A. (2001). Motivation and failure in educational simulation design. In K. D. Forbus & P. J. Feltovich (Ed.), *Smart Machines in Education. The coming revolution in educational technology* (pp. 37-70). Menlo Park CA: AAAI Press/MIT Press.

Schoenfeld, A. (1983). Problem solving in the mathematics curriculum: A report, recommendations and an annotated bibliography. *The Mathematical Associations of America, MAA Notes, 1*.

Schöfthaler, T. (1984). *Soziale Struktur und Vernunft: Jean Piagets Modell entwickelten Denkens in der Diskussion kulturvergleichender Forschung* (365). Frankfurt am Main: Suhrkamp.

Seel, N. M. (1991). *Weltwissen und mentale Modelle*. Göttingen [et al]: Verl. für Psychologie Hogrefe.

Seel, N. M. (1997). Pädagogische Diagnose mentaler Modelle. In H. Gruber (Ed.), *Wege zum Können: Determinanten des Kompetenzerwerbs* (pp. 116-137). Seattle: Huber.

Seel, N. M., Al-Diban, S. & Held, S. (1998). Didaktisches Design multimedialer Lernumgebungen. *Lernen mit Medien:* (pp. 87-120). Weinheim, München: Juventa.
Seel, N. M. & Al-Diban, S. (1999). Evaluation als Forschungsaufgabe von Instruktionsdesign. *Unterrichtswissenschaft, 27*(1), 29-60.
Seel, N. M., Al-Diban, S. & Blumschein, P. (2000). Mental models & instructional planning. In M. Spector & T. M. Anderson (Ed.), *Integrated and holistic perspectives on learning, instruction and technology: understanding complexity* (pp. 129-158). Dordrecht, NL: Kluwer Acad. Publ.
Seel, N. M. (2001). Epistemology, situated cognition, and mental models: 'Like a bridge over troubled water'. *Instructional Science, 29*, 403-427.
Seel, N. M. (2003a). Model centered learning and instruction. *Technology, Instruction, Cognition and Learning, 1*(1), 59-85.
Seel, N. M. (2003b). *Psychologie des Lernens: Lehrbuch für Pädagogen und Psychologen; mit. 12 Tabellen und zahlreichen Übungsaufgaben*. Basel: Reinhardt.
Seel, N. M. (2003c). Studienbrief Instructional Design (Modul 3). *Enter: Lernen und Lehren mit neuen Bildungsmedien.* Universität Kaiserslautern. Zentrum für Fernstudien und Universitäre Weiterbildung.
Seel, N. M. & Dijkstra, S. (2004). *Curriculum, plans, and processes in instructional design*. Mahwah, New Jersey: Lawrence Erlbaum Ass.
Seel, N. M. (2005). Designing model-centered learning environments: Hocus-pocus or the focus must strictly be on locus. In J. M. Spector (Ed.), *Innovations in instructional technology: essays in honour of M. David Merrill* (pp. 65-89). Mahwah, N.J. [et al]: Erlbaum.
Snow, R. E. (1990). New approaches to cognitive and conative assessment in education. *International Journal of Educational Research, 14*(5), 455-473.
Spector, M. (2001). Philosophical implications for the design of instruction. *Instructional Science, 29*, 381-402.
Spector, J. M. (2005). *Innovations in instructional technology: essays in honour of M. David Merrill*. Mahwah, N.J. [et al]: Erlbaum.
Stachowiak, H. (1973). *Allgemeine Modelltheorie*. Wien, New York: Springer.
Sternberg, R. J. (1998). Metacognition, abilities, and developing expertise: What makes an expert student? *Instructional Science, 26*, 127-140.
Strauss, S. (1996). Confessions of a born-again constructivist. *Educational Psychologist, 31*(1), 15-21.
Tennyson, R. D. (1997). A system dynamics approach to instructional systems development. In R. D. Tennyson, F. Schott, N. M. Seel & S. Dijkstra (Ed.), *Instructional Design. International perspective. Volume 1: Theory, Research, and Models* (pp. 413-426). Mahwah, New Jersey: Lawrence Erlbaum Associates.
Tennyson, R. D., Schott, F. & Seel, N. M. (1997). *Instructional design: International perspective: Vol.1: Theory, research, and models*. Mahwah, New Jersey: Lawrence Erlbaum Ass.
Thiagarajan, S. (1999). *Rapid instructional design*. (Retrieved July 13, 2007) http://www.thiagi.com/article-rid.html.
Thiagarajan, S. (2002). Unconventional instructional design: Faster, cheaper, better. *Educational Technology, 42*(3), 60-62.
Tripp, S. & Bichelmeyer, B. (1990). Rapid prototyping: An alternative instructional design strategy. *Educational Technology, Research & Development, 38*(1), 31-44.
Tuijnman, A. C. (1992). *Learning across the lifespan: Theories, research, policies*. Oxford [et al]: Pergamon Pr.
Zemke, R. & Rossett, A. (2002). A hard look at ISD. *Training Magazine, 39*(2), 27-35.

14. From Cognitive Theory to Instructional Practice: Technology and the Evolution of Anchored Instruction

James W. Pellegrino[1] **and Sean Brophy**[2]

[1] University of Illinois at Chicago

[2] Purdue University

Abstract: This chapter discusses evolution of the thinking of the Cognition and Technology Group at Vanderbilt about teaching, learning, assessment, technology, and instructional design under the broad heading of their extended work on the principles and practices of *Anchored Instruction*. It begins by stating a general set of ideas about the characteristics of powerful learning environments and the instructional design principles that are coupled to them. Then in subsequent sections it illustrates how features of the CTGV's evolving work on *Anchored Instruction* helped lead to and embody those principles. The subsequent sections begin by describing the earliest work on *Anchored Instruction*, the development of a set of multimedia instructional materials known as the *Adventures of Jasper Woodbury*. Later sections then describe work that pushed the ideas of *Anchored Instruction* in new and important directions that led to development of the *SMART* model and development of a general inquiry-learning model known as STAR.Legacy. An example of extending the Legacy instructional design model to higher education instructional settings is provided in the context of work in the VaNTH Center on bioengineering and biomedical engineering education. The chapter concludes with thoughts regarding what has been learned over time, challenges that remain in the areas of theory, research and practice, and the role of technology in the larger enterprise of connecting theory, research, and instructional design.

Keywords: Anchored instruction; instructional design; assessment; cognitive theory; technology.

Introduction and Overview

It is not often that individuals have the opportunity to reflect on a body of work that has evolved over a period of almost 20 years with the goal of trying to explicate some of what was learned in the process, as well as discuss what it might mean for theory and research and educational practice. We very much appreciate the opportunity to do so in this volume honoring the many contributions of Norbert Seel to these arenas of intellectual endeavor. The work that we have tried to overview and discuss in this chapter represents the efforts of many individuals who, over a series of years spanning the late 1980s to the early 2000s, were part of the Cognition and Technology Group at Vanderbilt (CTGV)[9]. We were privileged to be members of the CTGV for extended periods of time and work on many of the projects described subsequently in this chapter. However, we make no claim that what we have to say here is representative of the collective voice of CTGV[10].

Perhaps the one thing that the CTGV became recognized for was its *Anchored Instruction* approach to the design of technology-based instructional materials. As explicated in numerous papers and chapters by the CTGV over a more than 15-year period, the collective body of work on *Anchored Instruction* actually reflects an ongoing dialectic among issues of theory, instructional design, research on learning and assessment, technology, teacher knowledge and professional development, and the realities of diverse learners in diverse instructional settings. We cannot do justice to all of what was learned along the way or how it was learned. Rather, in this chapter we attempt to provide a glimpse of part of the evolution of the CTGV's thinking about teaching, learning, assessment, technology, and instructional design under the broad heading of *Anchored Instruction*.

The remainder of the chapter is divided into the following sections. In section two we have tried to state, in a very concise form, a general set of ideas about the characteristics of powerful learning environments and the instructional design principles that are coupled to them. In essence, this knowledge is the product of many years of research and development pursuing the general logic of *Anchored Instruction* and it is derived from related work by many individuals in the fields of cognition and instruction who were not part of the CTGV. We begin at the end so to speak and present some of the broader conclusions so that in subsequent sections it is easier to see how the features of our evolving work helped lead to and embody the current synthetic view. In section three we then describe the earliest and most well known work on *Anchored Instruction*, the development of a set of multimedia instructional materials known as the *Adventures of Jasper Woodbury.*

---

[9] There were many participants in the CTGV and a complete list would be prohibitively lengthy. In addition to the present authors, the principals included: Linda Barron, John Bransford, Elizabeth Goldman, Susan Goldman, Ted Hasselbring, Dan Hickey, Cindy Hmelo-Silver, Chuck Kinzer, Xiaodong Lin, Allison Moore, Joyce Moore, Anthony Petrosino, Vicki Risko, Dan Schwartz, Teresa Secules, Diana Sharp, Bob Sherwood, Nancy Vye, Susan Williams, and Linda Zech.

[10] Virtually all the members of the CTGV have left Vanderbilt since 2000. We sometimes refer to the dispersed collective group and the ongoing intellectual community as DCTGV – the Distributed Cognition and Technology Group from Vanderbilt.

Section four then describes work that went beyond the design ideas of the *Jasper* series and pushed the ideas of *Anchored Instruction* in new and important directions that led to development of the *SMART* model. One of the most important directions of the work with the *SMART* model was to generalize the *Jasper* challenge-based and problem–based approach and development of a general inquiry-learning model known as STAR.Legacy. The latter model and its instantiation via technology are described in section five. Then, in section six we describe briefly the process of applying the Legacy instructional design model to higher education instructional settings in the context of the VaNTH Center and its focus on bioengineering and biomedical engineering. In each of the settings where we have pursued elaboration and refinement of ideas about *Anchored Instruction*, technology has played a key role in helping to instantiate our ideas and applications. In section seven we provide some final thoughts regarding what we have learned along the way, what challenges remain in the areas of theory, research and practice, and how we view technology in this larger enterprise.

Characteristics of Powerful Learning Environments and Related Instructional Design Principles

Generalizations and principles derived from a large body of contemporary research and theory on learning, instruction, and assessment (e.g., Bereiter, 1990; Brown & Campione, 1994; Bruer, 1993; Collins, 1996; Scardamalia & Bereiter, 1994), including the collective work of the CTGV on *Anchored Instruction* (e.g., CTGV, 1992a; 1997; 2000), have been presented in a series of major reports from the U.S. National Research Council. These reports include *How People Learn: Brain, Mind, Experience and School* (Bransford et al., 2000), *Knowing What Students Know: The Science and Design of Educational Assessment* (Pellegrino et al., 2001), and *How Students Learn History, Mathematics, and Science in the Classroom* (Donovan & Bransford, 2005). The collective body of scholarly work discussed in these reports has led to the generation of four important dimensions of powerful and effective learning environments. These dimensions can and should influence the principles for designing instructional materials and practices and the use of technologies (see Bransford et al., 2000, CTGV, 2002, Goldman et al., 2005/06, and Pellegrino, 2003, 2004 for more detailed descriptions and elaborations of the ideas that immediately follow, their rationale, and ways in which technology enables their realization). The four dimensions, often referred to as the How People Learn (HPL) framework, include:

Effective learning environments are knowledge-centered. This means that explicit attention is given to what is taught – the central subject matter concepts; why it is taught – to support "learning with understanding" rather than merely remembering; and what competence or mastery looks like.

Effective learning environments are learner-centered. This means that educators must pay close attention to the knowledge, skills, and attitudes that learners bring into the classroom. This incorporates preconceptions regarding subject matter and occupational domains and it also includes a broader understanding of the learner. Teachers in learner-centered environments pay careful attention to what students know as well as what they don't know, and they continually work to build on students' strengths and prior knowledge.

Effective learning environments are assessment-centered. This means that it is especially important to make students' thinking visible through the use of frequent formative assessment. This permits the teacher to grasp the students' preconceptions, understand where students are on the "developmental corridor" from informal to formal thinking in a domain, and design instruction accordingly. They help both teachers and students monitor progress.

Effective learning environments are community-centered. This includes the development of norms for the classroom and workplace, as well as connections to the outside world, that support core learning values. These communities can build a sense of comfort with questioning rather than knowing the answers and can develop a model of creating new ideas that builds on the contributions of individual members.

Four principles for the design of instruction are consistent with the ideas just mentioned. These four principles are critically important for achieving learning with understanding.

- To establish knowledge-centered elements in a learning environment, instruction is organized around meaningful problems with appropriate goals.
- To support a learner-centered focus, instruction must provide scaffolds for solving meaningful problems and supporting learning with understanding.
- To support assessment-centered activities, instruction provides opportunities for practice with feedback, revision, and reflection.
- To create community in a learning environment, the social arrangements of instruction must promote collaboration and distributed expertise, as well as independent learning.

In the sections that follow we attempt to concretize what is meant by the preceding principles by grounding the discussion in *Anchored Instruction* design cases. The presentation includes discussion of which of the above characteristics were included in the work, why, and how. As we noted earlier, the preceding more general ideas about learning environments and instructional design principles were in large part the product of the genesis of research and development elaborating and refining the concepts and practices of *Anchored Instruction*. Our current thinking regarding the design of technology-supported learning environments that reflect the broader design principles derives from over 18 years of cumulative research with students and teachers on ways to motivate and assess exceptional learning (e.g., Barron et al., 1995; CTGV, 1994a,b, 1997, 2000).

The *Adventures of Jasper Woodbury Series*: Genesis of Anchored Instruction

Our initial work focused on the problem of learning with understanding in middle school mathematics, a problem identified in both psychological research and educational practice (e.g., NCTM, 1989). The focus of our efforts was on ways in which cognitive theory and research on problem solving might be connected with mathematics instruction. The result was development of the *Anchored Instruction* approach within which teaching and learning are focused around the solution of complex problems or *anchors*. The initial anchors were video-based stories (originally presented on videodisc) that each ended with a challenge to solve. All of the data needed to solve the challenges are contained in the stories. The problems: (a) are complex (more than 14 steps to solve) and require extended effort to solve (at a minimum, in the range of 3-5 hours for most middle school students); (b) are relatively ill-defined and require significant formulation prior to solving; and (c) have multiple viable solutions. The anchors are designed to engage students in authentic problem solving activities that highlight the relevance of mathematics or science to the world outside the classroom. The design of anchored instruction problem solving environments was a very explicit way to focus on using technology in instructional design to emphasize the knowledge centered and learner centered components of powerful learning environments. In the material that follows we briefly describe one such set of anchored instruction materials and the types of learning they afforded relative to the characteristics of powerful learning environments and principles of instruction discussed earlier.

The cumulative work on *The Adventures of Jasper Woodbury Problem Solving Series* (CTGV 1994a, 1997, 2000) is the single best example of our group's attempt to engage in the process of instructional design based on cognitive theory. Through the process of implementing those designs in multiple classrooms we came to understand the complexities of designing and managing powerful learning environments. For more detailed descriptions of this body of work and data, see CTGV (1992a,b,c, 1993a,b, 1994a; 1997, 2000), Goldman et al. (1997), Lamon et al., (1996), Pellegrino et al. (1991), Secules et al. (1997), Sharp et al. (1995), Vye, et al. (1997, 1998), Williams et al. (1998), and Zech et al. (1994, 1998).

The Jasper series consists of 12 interactive video environments that invite students to solve authentic challenges, each of which requires them to understand and use important concepts in mathematics. The challenges are carefully planned around the knowledge to be learned and how it will be assessed. The learning activities related to the challenge are designed to scaffold learners' knowledge construction by fostering a community of learning and inquiry. For example, in the adventure known as *Rescue at Boone's Meadow* (*RBM*), which focuses on distance-rate-time relationships, a video story begins with Larry teaching Emily to fly an ultralight airplane. During the lessons, he helps Emily learn about the basic principles of flight and the specific details of the ultralight she is flying, such as its speed, fuel consumption, fuel capacity, and how much weight it can carry. Not long after Emily's first solo flight, her friend Jasper goes fishing in a remote area called Boone's Meadow. Hearing a gunshot, he discovers a wounded bald eagle and radios Emily for help in getting the eagle to a veterinarian. Emily consults a map to determine the closest roads to Boone's Meadow, then calls Larry to find out about the weather and see if his ultralight is available. Students are challenged to use all the information in the video to determine the fastest way to rescue the eagle. We developed technology tools to help students navigate the video non linearly as a scaffold to their access of important data embedded in the video story.

After viewing the video, students review the story and discuss the setting, characters, and any unfamiliar concepts and vocabulary introduced in the video. After they have a clear understanding of the problem situation, small groups of students work together to break the problem into sub-goals, scan the video for information, and set up the calculations necessary to solve each part of the problem. Once they have a solution, they compare it with those that other groups generate and try to choose the optimum plan. Like most real-world problems, Jasper problems involve multiple correct solutions. Determining the optimum solution involves weighing factors such as safety and reliability, as well as making the necessary calculations.

The Jasper series focuses on providing opportunities for problem solving and problem finding. It was not intended to replace the entire mathematics curriculum. Frequently, while attempting to solve these complex problems, students discover that they do not have the necessary basic skills. Teachers used these occasions as opportunities to conduct benchmark lessons in which they reviewed the necessary concepts and procedures. Solutions to *RBM* clearly require mathematical knowledge. For example, students have to solve distance-rate-time problems such as how long it would take to fly from point A to point B given the cruising speed of the ultralight. But there are big ideas about mathematics, (e.g., concepts such as rate) that are not necessarily revealed by simply solving problems such as *RBM*. We devised three strategies for helping students abstract important mathematical ideas from their experiences with Jasper adventures. The first was to encourage teachers to use sets of similar Jasper adventures rather than only one, and to help students compare the similarities in solution strategies required to solve them[11]. Gick and Holyoak's (1983) work on abstraction and transfer illustrates advantages of this approach. Often, however, teachers wanted to use dissimilar Jasper adventures (e.g., one involving distance-rate-time, one involving statistics); this reduced the chances of abstraction. We also discovered that additional strategies were needed to help students conceptualize powerful mathematical ideas.

A second strategy for making the use of Jasper Adventures more knowledge centered while also scaffolding student learning was to develop analog and extension problems for each adventure that invited students to solve "what if" problems

[11] The 12 adventures encompassed four major mathematical content areas, with three adventures for each content area. The three adventures within a content area represented different situations and were progressively more complex.

that changed parameters of the original problems. For example, given the *RBM* adventure discussed above: "What if the ultralight had travelled at a speed of 20 rather than 30 miles per hour? How would that affect the solution?" Or "What if Emily flew to Boone's Meadow with the help of a 10 mph tailwind and flew back with a 10 mph headwind? Would these two cancel each other out?" (The answer is "no" and it is instructive for students to explore the reason).

Analog problems for *The Big Splash* (*TBS*), a Jasper statistics adventure, further illustrate the importance of promoting what-if thinking. When solving the adventure, students see a particular method used to obtain a random sample and extrapolate the findings to a larger population. Analog problems help students think about alternative randomization procedures that might have been used. Data indicate that the use of analog problems increases the flexibility of students' transfer. Students benefited greatly from opportunities to develop a more general understanding of sampling and statistical inferencing after solving *TBS* (Schwartz, Goldman, Vye, Barron, Bransford & CTGV, 1998). Similarly, students who solved analog problems after solving *RBM* showed more flexibility in their thinking than students who solved *RBM* without analogs. For example, they were more likely to modify their original solution strategies when presented with transfer problems that could be solved more elegantly with a change in strategy (e.g., CTGV, 1997; Williams, 1994).

A third strategy for helping students understand big ideas in mathematics and scaffolding their learning takes the idea of analog and extension problems one step further. Instead of presenting students with environments that involve only *one-shot* problem solving, we help them conceptualize environments where problems tend to reoccur and it becomes useful to invent ways to deal with these re-occurrences. Theorists such as Lave (1988), Norman (1993), Pea (1993), Rogoff (1990) and others argue that people become smart, in part, by learning to eliminate the need to laboriously compute the answer to important problems that tend to reoccur in their environments. One way to do this is through the invention of "smart" tools. Examples of smart tools that can eliminate cumbersome computations include charts, graphs, computer programs, and gadgets such as watches, speedometers and proportion wheels. We did not want to simply give students tools because these can often be applied without understanding, causing people to fail to adapt when situations change (e.g., see Resnick, 1987). We wanted to help students invent, test, and refine their own smart tools.

We pursued development of smart tools in the context of Jasper Adventures such as *Rescue at Boone's Meadow* (*RBM*). Instead of receiving a challenge where they are asked to solve a single problem (rescuing the eagle as quickly as possible), students are invited to imagine that Emily becomes an entrepreneur who sets up a pickup and delivery service for people who go camping in her area. She has several different planes to choose from depending on the needs (e.g., some can carry more payload, fly faster, etc.). When customers call to ask for assistance, Emily asks where they are (or where they want to go) and what they want to carry;

she then needs to tell them the trip time and fuel costs as quickly as possible. Different requests involve problems that vary in terms of the location of the destination, windspeed conditions, payload limits that determine which plane must be used, costs due to fuel consumption, and so forth. To calculate the answer to each individual problem is cumbersome, and Emily needs to be as efficient as possible. The challenge for students in the classrooms is to invent smart tools that allow people to solve such problems with efficiency.

In summary, the *Jasper Adventures* provide one example of using cognitive theory and technology to assist in the design of learning environments that demonstrate knowledge-centered and learner-centered features. In developing the problems and then designing instructional strategies and tools to support learning from these materials we tried to adhere to the principles that instruction should be organized around the solution of meaningful problems and that the environment should provide scaffolds for support of meaningful learning. Technology was a significant component in attempts to achieve these objectives.

The *SMART* Model – Extending the Principles of Anchored Instruction

The development, implementation and evaluation of the Jasper materials was coincident with and impacted other related curricular design projects that extended the ideas of anchored instruction. For example, the research and development team subsequently participated in the construction of a second set of curriculum designs, designated by the title *Scientists in Action* (Goldman et al., 1996; Sherwood, Petrosino, Lin, Lamon & CTGV, 1995). In response to student, teacher, and researcher feedback, these curriculum materials underwent development along distinct but somewhat parallel lines to those in the Jasper Adventure series. Once again, designs evolved to incorporate sophisticated forms of scaffolding to enhance effective student learning. Increasingly, the focus was on the two dimensions of a powerful learning environment that many of the Jasper projects had not focused on deeply – formative assessment and community building.

Various methods were explored in the context of working with both the *Scientists In Action* series and the *Jasper Adventures* to provide frequent and appropriate opportunities for formative assessment (Barron et al., 1995, 1998; CTGV, 1994a, 1997, 2000). These included assessment of student-generated products at various points along the way to problem solution such as blueprints or business plans, and assessment facilitated by comparing intermediate solutions with those generated by others around the country who were working on similar problem-based and project-based curricula. In these examples, assessment is both teacher and student generated and it is followed by opportunities to revise the product that has been assessed. The revision process is quite important for students and seems to lead to changes in students' perspectives of the nature of adult work as well as conceptual growth.

We also took advantage of the fact that different ways of organizing classrooms can also have strong effects on the degree to which everyone participates, learns from one another, and makes progress in the cycles of work (e.g., Brown & Campione, 1996; Collins, Hawkins, & Carver, 1991). We have found it beneficial to have students work collaboratively in groups, but to also establish norms of individual accountability. One way to do this is to set up a requirement that each person in a group has to reach a threshold of achievement before moving on to collaborate on a more challenging project, for example, to be able to explain how pollution affects dissolved oxygen and hence life in the river; to create a blueprint that a builder could use to build some structure. Under these conditions, the group works together to help everyone succeed. The revision process is designed to insure that all students ultimately attain a level of understanding and mastery that establishes a precondition for moving from the problem-based to project-based activity.

The larger model we developed became known as *SMART* which stands for *Scientific and Mathematical Arenas for Refining Thinking*. The *SMART* Model incorporated a number of design features to support the four features of an effective learning environment mentioned in the first section of this chapter. For example, a variety of scaffolds and other technology-based learning tools were developed to deepen the possibilities for student learning. They included: (a) Smart Lab – a virtual community for students in which they are exposed to contrasting solutions to problems in the context of being able to assess the adequacy of each; (b) Toolbox – various visual representations that could be used as tools for problem solving; and (c) Kids-on-Line – a tool that featured students making presentations. By using actors who make presentations based on real students work, we were able to introduce presentations with typical, representative errors. This design feature allows students to engage in critical analysis of the arguments and see same age peers explaining their work in sophisticated ways.

As part of this process of providing resources we found that breaking down the isolation of the classroom could also be a powerful way to support learning through social mechanisms. It proved useful to provide students and teachers with access to information about how people outside their classroom have thought about the same problem that they are facing (CTGV, 1994a, 1997, 2000). Such access can help students be more objective about their own view and realize that even with the same information other people may come to different conclusions or solutions. In addition, discussion of these differences of opinion can support the development of shared standards for reasoning. This can have a powerful effect on understanding the need for revising one's ideas and as a motivator for engaging in such a process. Internet and Web-based environments now provide excellent mechanisms to incorporate these processes within the overall *SMART* Model. Some of these mechanisms include interactive Websites with database components which are dynamically updated as students from multiple classrooms respond to items or probes on the Website. We describe an example of this use of the internet and Web below in the context of implementing the *SMART* Model with a *Scientists in Action* adventure.

Reviewing the evolution of one *Scientists in Action* curriculum, the *Stones River Mystery*, will demonstrate the increasing attention paid to embodying all four elements, especially the assessment centered and community centered elements, while further refining the learner and knowledge centered dimensions already present in the Jasper anchored instruction designs. It also illustrates how Web-based resources were introduced into the learning environment.

In the *Stones River Mystery* video, students are presented with the challenge of understanding the complexities of river pollution. In its early form, students engaged in this curriculum watched a video story of a team of high-school students who are working with a hydrologist and a biologist to monitor the quality of the water in Stones River. The team travels in a specially equipped van to various testing sites, and can electronically submit test results to students back at school. The video shows the team collecting and sorting samples of macro-invertebrates from the river's bottom. They also measure dissolved oxygen, temperature, and other water quality indicators. The video anchor poses various challenges to students: first, test and evaluate the water quality at a particular river site to determine if pollution is present, then localize the source of the pollution, and finally, determine the best method to clean up the pollution.

Students use a Web site as they solve each challenge. The site has three components. One component gives individualized feedback to students about their problem solving that they can use to revise their work. For example, when working on the challenge of how to clean up the pollution in Stones River, students access our Web-based version of the Better Business Bureau. This site contained proposals submitted by various companies for how to clean up the pollution in Stones River. Students using the site were asked to select the best company to hire based on the information provided in the proposals and to chose a rationale for their selection. They also indicated their rationale for not selecting each of the other companies. Some of the proposals contained deliberately erroneous scientific information and clean up solutions. Students' selections and rationales were submitted online and they received individualized online feedback (that could be printed out for later reference) about their submissions. The feedback contained information on any problems with their selections and rationales and suggestions for offline resources for learning more about key concepts.

In addition to the formative feedback component, the site had two other components that were designed to draw on the community building potential of the Internet as well as serve a formative assessment function. The site contained a backend database that collected the data submitted by all student users. Information from the database was used to dynamically build graphs that displayed aggregate data on the choices and reasons that students selected. These graphs could be viewed and discussed by the class. Since students' opinions often differ on which company is best and why, the graphs were conversation starters that make this evident and invited discussion about the merits of various choices and about key concepts that support or argue against a specific choice and/or rationale. Students were motivated to take part in discussions about the aggregate data because they understand that it represented the input of students in their class and other participating classes. Finally, the site contained presentations (text plus audio) by actors in which they explain their ideas on which company to hire and why. As mentioned earlier, these presentations purposely contain misconceptions that students often have about the underlying science content. The actors ask students to give them feedback on their thinking (they submit this feedback which is then posted online). In this way, we once again tried to seed the classroom discussion and focus it on understanding important content by bringing the input of the broader learning community into the classroom.

SMART learning environments were designed to foster the development of high standards, rich content, and authentic problem solving environments. But as compared to earlier curriculum units, special attention was paid to assessment and community building. The former category included assessments of student-generated products such as blueprints or business plans, and assessment facilitated by comparing solutions to problems with others around the country who were working on similar problem and project based curricula. In the second category were a number of tools that enabled students to get feedback from a variety of external communities, including parents, community leaders, expert practitioners, and academics. As they evolved, *SMART* learning environments embodied all four instructional design principles discussed earlier: (1) A focus on learning goals that emphasize deep understanding of important subject matter content, (2) The use of scaffolds to support both student and teacher learning, (3) Frequent opportunities for formative self-assessment, revision, and reflection, and (4) Social organizations that promote collaboration and a striving for high standards. Each of these four characteristics was enabled and supported through the use of various technologies for the delivery of resources and the exchange of information. The ability of students and teachers to make progress through the various cycles of work and revision was dependent on the various resource materials and tools that assisted the learning and assessment process. Our research indicated that students who used these tools learned significantly more than students who went through the same instructional sequence for the same amount of time without using the tools (Barron et al., 1995, 1998; Vye et al., 1998).

STAR.Legacy – A Generalized Inquiry Model and Technology Support Tool

The anchored instruction designs described above for areas of mathematics and science – *Jasper Adventures* and *Scientists in Action* – were very much focused on carrying out an extended inquiry process within regular classrooms with reasonable technology capability. They could be executed at varying levels of sophistication depending on access to specific technology infrastructure. In the the course of pursuing these designs, it became apparent that the most effective learning and instruction was transacted in situations where all four elements of powerful learning environments were present. Doing so demands a wide range of resources and tools and also the capacity to organize and manage the instructional and learning process in a way that is faithful to learner exploration and support. One of the things discovered along the way, especially as we implemented *SMART* inquiry environments, was the need for externalization of the overall process. As a result, we developed a software shell for helping people visualize and manage inquiry in a manner that is learner, knowledge, assessment and community centered. Called STAR.Legacy (STAR stands for Software Technology for Action and Reflection), the environment provides a framework for anchored inquiry. Chances to solve important problems, assess one's progress and revise when necessary play a prominent role in the Legacy cycle. The environment can also easily be adapted to fit local needs; in part by having teachers and students leave legacies for future students (hence the name Legacy).

The STAR.Legacy design grew out of collaborations with teachers, students, curriculum designers, educational trainers, and researchers in learning and instruction. One of its virtues is that it makes explicit the different components of an instructional event focused around an inquiry process. Furthermore, it connects the events with learning theory and the four components of powerful learning environments. STAR.Legacy formalizes those components and their rationale within a learning cycle that is easy to understand and follow. Figure 1 shows the home page of STAR.Legacy (Schwartz, Brophy, Lin, & Bransford, 1999; Schwartz, Lin, Brophy & Bransford, 1999). The software features learning cycles anchored around successive challenges that are represented as increasingly high mountains. As learners climb each mountain, they progressively deepen their expertise. Within each challenge students generate ideas, hear perspectives from experts, conduct research and receive opportunities to test their mettle and revise before *going public* with a synthesis of their ideas. The structure of STAR.Legacy is designed to help balance the features of learner, knowledge, assessment and community centeredness.

The home page of STAR.Legacy helps students become more active and reflective learners by being able to see where they are in a research cycle. The importance of this feature became apparent during research that was discussed earlier (Barron et al., 1995; 1998; CTGV, 1994a, 1997; Vye et al., 1998). When working on complex units, students often got lost in the process. In earlier research, we created a visual map of a curriculum unit that was placed in the classrooms; this

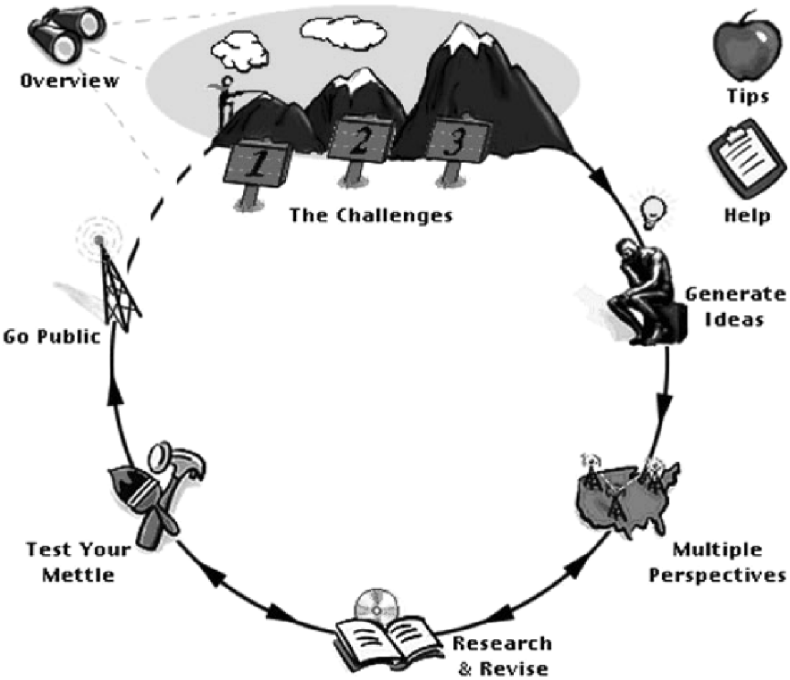

Fig. 1. The STAR.Legacy reference diagram illustrating the organization and sequencing of the components of a multiple challenge inquiry learning cycle.

helped students and teachers understand where they were in the process and where they were going (see Barron et al., 1995 for more information). STAR.Legacy provides a model of inquiry that represents an advance on these earlier designs. It too helps students see where they are and where they are trying to go. White and Fredricksen's (1998) inquiry model for science (which we see as more specific than Legacy but definitely compatible with it) provides an additional illustration of the importance of helping people visualize the processes of inquiry. Teachers can use STAR.Legacy to organize whole class work (by projecting the software on a big screen), or students can work individually or in groups.

The overview section of Legacy (see the top left of Figure 1) allows teachers or students to begin by exploring the purposes of the unit and what they should accomplish by the end. Overviews frequently include pretests that let students assess their initial knowledge about a topic. Teachers can use the pretest to make the students' thinking visible and identify beliefs about the subject matter that may need careful attention (for examples of the value of beginning with students' assumptions, see Hunt & Minstrell, 1994). At the end of the unit students can revisit the pretest and see how much they have learned; the benefits of this are discussed in Schwartz, Brophy, Lin, & Bransford (1999) and Schwartz, Lin, Brophy, & Bransford (1999).

We encourage pretests that are problem-based or design-based rather than simply multiple-choice or true-false tests. The former are more interesting to students initially, and they better reveal students' assumptions about phenomena since their answers are not constrained by pre-determined choices provided on the tests. A pretest for a unit on *rate* may ask students to help a small private flying company design a way to efficiently give people trip time estimates from their airport depending on wind speeds, aircraft speed, and their destination. On a pretest, middle school and even college students usually try to remember formulas and almost never think of inventing smart tools as discussed earlier.

A pretest for one of our units on monitoring rivers for water quality introduces students to Billy Bashinall (a cartoon character) who acts in ways that reveal his assumptions (many are wrong) about water; the students' job is to assess Billy's understanding and prepare to teach him. Students also assess Billy's attitude toward learning (which is quite negative) and discuss their thoughts about how it might be changed. For a more detailed discussion of the idea of using *teachable agents* as part of the instructional design see Brophy, Biswas, Katzlberger, Bransford & Schwartz (1999).

Overall, the *overview* component of Legacy is learner centered in the sense that students are encouraged to state their initial ideas about the overview problem and define learning goals; knowledge centered in the sense that units are designed which focus on important ideas in a discipline; assessment centered in the sense that teachers can gather information about students' beliefs, students can discover that they hold conflicting beliefs as a group, and students can eventually assess their progress when they return to the *overview* problem after completing the unit. The *overview* also supports the development of community centered environments in the sense that teachers can present the pretest as a challenge that all will work on collaboratively (although there can still be individual accountability) in order to communicate their ideas to an outside group (e.g., by *going public* through live presentations or publishing on the Web).

Following the *overview* to a unit, students are introduced to a series of challenges (represented as increasingly high mountains in Figure 1) that are designed to progressively deepen their knowledge. This way of structuring challenges reflects lessons learned about knowledge centered curricula that were discussed earlier. For example, the first challenge for a unit on *rate* in mathematics may ask students to solve the Jasper adventure Rescue at Boone's Meadow, where Emily uses an ultralight to rescue a wounded eagle. It's a calm day, so headwinds and tailwinds do not have to be taken into the account. What is the best way to rescue the eagle and how long will that take? (issues of fuel consumption and capacity come into play in determining this).

The second challenge uses footage of Charles Lindbergh and asks students to decide if he could make it from New York to Paris given particular data such as his fuel capacity and plane speed. Interestingly, Lindbergh couldn't have made the trip without a substantial tail wind. This helps students extend their thinking of rate by considering ground speed versus air speed.

The third challenge invites students to build tools that help them solve entire classes of problems rather than only a single problem. The challenge features Emily's rescue and delivery service. Students are challenged to create a set of "smart" tools that help Emily quickly determine how long it will take to get to various destinations and what the fuel costs will be. Attempts to build tools eventually introduce students to linear and nonlinear functions; for example, changes in wind speeds on different parts of a route introduce students to piecewise linear functions. Examples are provided in Bransford, Zech, Schwartz et al. (1996; 1999).

Within each challenge cycle, Legacy engages students in activities that encourage them to use their existing knowledge, assess their learning, and develop a sense of community with their classmates and the teachers. First, students generate their own ideas about a challenge–either individually, in small groups, or as a whole class. Teachers can listen to students' ideas in whole class sessions and students get a chance to hear the ideas of their classmates. Electronic notebooks in Legacy let teachers capture the thinking of the whole class, or let students who are working in groups capture their thinking and keep it with them as they work through the unit. Teachers can access these notebooks to see how students are approaching the challenges.

After generating their own ideas, students can open *multiple perspectives* and see experts discussing ideas that are relevant to the challenge. An important goal of *multiple perspectives* is to help students (and in some instances teachers) begin to see the relationship between their personal thinking about an issue and the thinking that is characteristic of experts from a scientific community. For example, middle school students may have intuitions that tail winds can help a plane fly faster, but few will have a well-worked out understanding of how to quantify this intuition and how to differentiate air speed and ground speed. In addition, experts can help students understand the value of thinking about the rate of airplane fuel consumption from the perspective of hours of flying time rather than miles per hour.

Opportunities to see *multiple perspectives* have been especially well-received by students. Since they first get to generate their own ideas about a topic (in Generate Ideas), they are able to contrast their own thinking with the thinking of experts. Opportunities to experience contrasting cases are important for helping students develop a deeper understanding of new areas of inquiry (e.g., see Bransford et al., 1989; Bransford & Schwartz, 1999; Schwartz & Bransford, 1998). The fact that Legacy is intended to be an authorable environment also means that teachers can add themselves as one of the multiple perspectives. Having a video of the teacher appear in the media appears to have effects on student attentiveness that go beyond simply having the teacher say the same thing in front of the class.

The "research and revise" component of Legacy provides access to resources for learning, including videos, audio, simulations, and access to the Web. Resources can be different for each of the separate Legacy challenges. For example, a resource for the tools needed for the third Jasper challenge noted above might involve a simple Java-based simulation for building graphs with student-specified scales.

An especially important feature of each Legacy cycle is the *test your mettle* component that focuses attention on assessment, especially self-assessment. For example, students can assess their thinking before *going public* (e.g., by making oral or written presentations or publishing on the Web). For the first Jasper challenge noted above (the Eagle rescue), *test your mettle* might provide students with a checklist to assess whether they have considered important elements such as payload, availability of suitable landing sites, fuel consumption, etc. Alternatively, Legacy might link students to the Jasper Adventureplayer software that allows them to enter their answers to the rescue problem and see a simulation of the results (Crews, Biswas, Goldman & Bransford, 1997).

For the second challenge cycle noted above (the Charles Lindbergh example), *test your mettle* may help students assess their thinking about headwinds and tailwinds (e.g., If Lindbergh takes a test flight where he travels 60 miles to City A with a 10 mph headwind and then flies back with the wind as a tailwind, will the headwind and tailwind cancel each other out?). The Jasper Adventureplayer software (Crews et al., 1997) can also provide feedback on these kinds of variables. The *test your mettle* for the third challenge cycle (the one dealing with smart tools) may provide students with a variety of *call ins* to Emily asking about trip time and expenses to take them to certain locations. Students test their tools to see how efficiently they permit answers to these questions. If the tools need revising (and they usually do), students can go back to *resources*, *multiple perspectives*, or any other part of Legacy. The software is designed to allow flexible back-and-forth navigation–it does not force students into a rigid loop.

At the end of each of the three legacy cycles students have the opportunity to "go public." This includes a variety of options such as making oral presentations in the classroom or posting reports, smart tools, simulations and other artifacts on the Web. The idea of *going public* helps create a sense of community within the classroom because students' and teachers' ideas are being considered by others. For example, in a Legacy used in a college class, students at Vanderbilt went public by publishing essays on the Web that were then reviewed by students from Stanford (Schwartz, Brophy, Lin, & Bransford, 1999). In middle school classrooms, students' essays might be read by other classes or by groups of experts from the community.

At the end of the final challenge in a Legacy unit, students can return to *overview* and revisit the pretest that they took initially. This helps students see how much they have learned in a Legacy unit, and it provides feedback that helps them identify any important concepts that they may have missed.

Most teachers do not have time to create Legacy units from scratch, but we have found that they like to adapt existing units so that they better fit the needs of their classrooms. Teachers may want to choose from several different pretests for a unit, choose from among a number of possible challenge cycles, select a subset of the available resources and *test your mettle*, and so forth. We noted earlier that teachers can also place themselves in the software; for example, as video participants in *multiple perspectives*, *resources*, or *test your mettle*. Over time, Legacy units can be readily adapted to meet teachers' particular goals and needs.

Legacy can also be used to help foster school-wide and community-wide connections by adding videos of local experts who appear in *multiple perspectives*, and resources, and as friendly critics in *test your mettle*. The ability for students within a single class to meet (via video) others from the local school community is powerful. Because students see them visually and learn something about their expertise and interests, our experiences show that they are much more likely to begin conversations and ask questions when they meet these people. We have also observed that students are more likely to go to these people for advice when a project they are working on is relevant to their areas of expertise (Schwartz, Brophy, Lin, & Bransford, 1999; Schwartz, Lin, Brophy & Bransford, 1999).

Legacy is also designed with the idea of having students add their own information to units. The primary mechanism for doing this is to leave a legacy for the next group that explores a unit. Students can add their ideas to *multiple perspectives*, *resources*, or *test your mettle*. For example, students may provide a clip or two in *multiple perspectives* that explains that a challenge initially seemed arbitrary to them and a waste of time but they later realized its value – hence the new students should stick with it. Students might also point toward new resources in the *research and revise* section (e.g., new Web sites) that they found to be particularly valuable. If they wish, students can also add a new *challenge*. Overall, the opportunity to leave a Legacy is very motivating to students and helps them see themselves as part of a community whose goal is to teach others as well as to learn.

Legacy also helps teachers envision a common inquiry cycle that extends across disciplines. Many teachers have found that this helps them communicate with colleagues in other disciplines. A number have used the Legacy cycle to design their curricula even when it is not computerized. The *overview* and *test your mettle* sections of Legacy have been especially important for curriculum development because they focus attention on goals for learning that are more explicit than a mere list of abstract objectives that are often only vaguely defined.

Finally, Legacy also provides a shell for dynamic assessment environments that can be used to assess people's abilities to learn new information. Legacy can capture students' questions, use of resources, tests and revisions. Students who are better prepared for future learning (because of well organized content knowledge as well as attitudes and strategies) should do better in these environments than students who are less prepared (for more discussion see Bransford & Schwartz, 1999).

Anchored Instruction meets Higher Education: Challenge-based Instructional Design in the VaNTH Center

The STAR.Legacy design was intended to be flexible and adaptive to the needs of instructors and learners in a variety of contexts. Evidence of this claim is that it has been employed in both K-12 and adult learning contexts. At the K-12 level it has been used with inquiry-based science learning and with complex mathematics problem solving at the middle school level. With adults it has been used in instructional contexts ranging from electrical engineering (Biswas et al., 1997), educational psychology (Schwartz et al., 1999), and bioengineering education (Harris et al., 2002; Roselli & Brophy, 2006a; Perreault et al., 2006; Martin et al., 2005; 2006), to teacher professional development. For some examples see Schwartz, Brophy, Lin, and Bransford, (1999) and Schwartz, Lin, Brophy, and Bransford, (1999). Across these varied contexts it has proven to be a highly usable design and development environment. It can be as simple or as high tech as one wants and needs. It has the capacity to support complex learning because it links activities to purposes and makes things explicit and manageable. As discussed above, it includes all four components of effective and powerful learning environments. With regard to other technology tools and resources, it can support the use and integration of multiple technology tools and products including multimedia, internet resources, simulations, automated assessments, virtual reality, and communication and dialog systems. Finally, it is highly compatible in its design features with future technology tools and developments in education.

It has proven especially interesting to apply the STAR.Legacy learning cycle in higher education contexts. As we will argue below, it facilitates the enhancement of teaching and learning in higher education contexts that are often very dependent on the traditional lecture and test model of *transferring knowledge*. Such a traditional transmission model of instruction typically supports the development of declarative knowledge and procedures associated with a domain, but does not always provide the experience that learners need to be creative nor does it develop new interest and desire to pursue disciplinary careers. The Legacy cycle provides an alternative approach to designing instruction that supports more advanced learning objectives that develop flexible thinking, sensitivity to multiple perspectives, awareness of professional communities, and greater independence for guiding one's own inquiry. In higher education, learners may often gain these experiences through other activities such as internships, club activities, and other extracurricular professional development experience they have sought out. The Legacy Cycle enhanced with technology provides a mechanism to bring these kinds of experiences into the classroom.

A case in point is the Engineering Research Center (ERC) funded by the U.S. National Science Foundation to enhance bioengineering education by applying advanced theories from the learning sciences and developing accompanying technologies to support learning, instruction and assessment. The Center called, VaNTH, was a collaboration of four institutions including Vanderbilt, Northwestern, Texas (at Austin) and HST (Health, Science and Technology program at Harvard and MIT) (VaNTH 2007). A fundamental conjecture of the VaNTH is that anchored inquiry, specifically challenge-based instruction, is an effective method for achieving the learning outcomes desired for bioengineering students entering a rapidly changing discipline. Further, the lessons learned from the K-12 research presented in the previous sections can apply to undergraduate education of bioengineering. Specifically, the STAR.Legacy Learning Cycle provided a useful framework for organizing learning activities that achieve a balance of the dimensions of the How People Learning (HPL) framework described earlier. Over the past 8 years these frameworks have guided their decisions for what, how and when to teach bioengineering content.

The HPL and Legacy frameworks quickly became a focal point for VaNTH participants to discuss and share ideas about the issues and opportunities for refining their learning environments. Bioengineering domain experts were new to the theories of learning presented by the learning scientists in the Center. However, the explicit description of the Legacy Cycle provided a close analogy to the inquiry cycle engineers use to solve their own problems. Therefore, these frameworks facilitated the communication between experts from the multiple disciplines (domain, learning sciences, assessment and technology) involved in the Center. Now the challenge became bridging these theories into the practice of bioengineering education. Learning Scientists worked with the domain experts to evaluate their current practices and identified opportunities to transform their current practices into enhanced learning opportunities for their students.

For example, biotechnology was a technical elective for seniors in biomedical engineering. One portion of the course revolved around the professor's demonstration of how to design a bioreactor to grow cells in mass quantities. The traditional model of instruction began with the professor's one-minute set up of a context for developing a new process for manufacturing a drug for hemophiliacs with a reduced chance of infection. Subsequent instruction consisted of the instructors modeling of his problem solving process and derivation of mathematical models necessary for sizing a specific type of bioreactor for the application. The context and lectures from this more traditional model of instruction were transformed into a more generative learning environment for the students that engage them in the process of problem solving and modeling experiences. Now, the instruction using the Legacy cycle presented challenges that put students in the role of a bioengineer and presented a grand challenge to transform a market analysis of potential users of a product line into a feasible product line for the company. This set up for their learning experience provides a realistic engineering problem complete with data that needs to be interpreted in order to determine if they have a viable product. In the new instructional model, students began by generating ideas about potential solutions and questions they needed to answer to improve their own comprehension of the grand challenge and the knowledge they needed to better solve the problem. This sets up a series of smaller challenge around factors influencing cell growth, existing methods for growing cells (i.e., the various types of bioreactors) and computational analysis for evaluating the feasibility for a stirred tank bioreactor as one option to the grand challenge (Giorgio & Brophy, 2002; Birol et al., 2002).

The bioreactors challenges organized the learning around the STAR.Legacy cycle beginning with generating ideas around the sub challenges students identified as necessary to solving the grand challenge. They began their inquiry into solving a sub challenge by generating ideas and questions about the challenge and comparing these ideas with perspectives provided by experts familiar with the challenge. Then they conducted their own self-guided research into specific concepts related to these challenge. All of these activities are conducted as a precursor to coming to lecture. Sometimes lectures consisted of the traditional information transfer by the instructor. However, now students were primed to understand the relevance of this knowledge and actively use it to answer their own questions they generated before lecture (Schwartz & Bransford, 1999). For example, in multiple instances students would reference the multiple perspectives that they would want more elaboration on, or clarification of, now that the instructor has provided new knowledge. Therefore, the pre-lecture materials function as a catalyst for formulating questions they want to know more about and the classroom community is established where they are comfortable asking for this additional knowledge. In other situations the instructor would ask students to invent their own mathematical models of the bioreactor systems and the class would discuss the benefits and limits of each of the models. This provides students with valuable feedback on their mathematical competency for representing complex systems. Often the professor would demonstrate his mathematical model of the system. Students would remark that they remember some of the concepts from prior courses, but really did not realize that they applied to these specific situations. However, based on transfer questions after the instructional intervention they demonstrated a better ability to apply these engineering principles. Students would now complete homework assignments with feedback. The final stage in the process was to synthesize everything they learned into a report that articulated a solution to the problem complete with a computational analysis and rational as evidence for the feasibility of their design.

The reformed instructional model of the biotechnology course contained the same content as the traditional instruction and included new activities that required students to process the information in a different way. The learning process was transformed into a knowledge generation activity by students that simultaneously developed their ability to analyze authentic context to identify problems, generate questions to establish personal learning goals, articulate their own mathematical understanding as they apply it to a new problem and synthesize their own ideas to articulate a solution to the problem. This is the *Go Public* phase of the cycle that could end both a written report and an oral report which provides excellent opportunities for students to demonstrate their communication skills and receive feedback on the process.

Multiple examples of bioengineering legacy cycles were critical to the professional development of other faculty in the ERC. The bioengineering instructors quickly comprehended the benefits of the instructional approach of Legacy and the importance of the dimensions of the How People Learn frameworks emphasis on not only the knowledge to be learned, but also on the needs of the learner, methods for providing the feedback and monitor their progress toward course objectives and forming a classroom community. However, the examples from K-12 did not always help them realize practical ways for changing their own practice. Therefore, examples such as the bioreactor challenge became critical to providing a new vision for refining bioengineering classrooms.

VaNTH explored and developed a number of technologies to facilitate instructors' integration of the *How People Learn* instructional design principles and Legacy into their classroom. For example, a classroom communication system involves electronic devices that each student has at their desk that they can use to respond to questions posed during lecture. The results of the students' polls are collected electronically and in seconds a classroom computer can display a histogram of students' responses to the multiple-choice question. This provided the instructor and students with immediate feedback on students' ability to apply their knowledge to conceptual questions (see Roselli & Brophy 2006b). They also developed a courseware authoring and delivery system that managed the presentation of challenges and recorded students' articulation of ideas and questions. The system could also deliver and track students' interactions with simulations and other dynamics and interactive learning during online learning experiences. This system called CAPE (Courseware Authoring and Packaging Environment) provides a highly flexible environment for instructors to present learning activities on the Web to learners outside of class and to record students' ideas and actions during these activities. The instructor can use these records to evaluate students' thinking before students come to class. Now instructors can better anticipate the needs of their students and adjust their lectures and in-class activities appropriately. In addition, the authoring environment contains design patterns of effective learning activities that have been developed over the years. New instructors can use these legacies from prior instructors as building blocks to design their own instruction.

The process of bringing the research from K-12 to higher education leverages similar learning principles. However, the needs of the learners are changing and the infrastructure of the institutions can have an effect on the design. Undergraduates are learning to become more independent learners who possess the potential to flexibly adapt to new situations. They must learn to deal with the ambiguity of dealing with novel situations and develop strategies and attributes for managing the dynamics of these situations. If the classrooms in universities do not provide these opportunities, then breadth of student development as innovative thinkers is delayed until they are on the job. However, the stakes then are higher and the support structures for learning are no longer available. If we are to prepare learners for future learning and increase their potential for noticing innovative opportunities, then our learning environments in higher education must incorporate innovative instructional methods. VaNTH demonstrated the potential for adopting this approach. The Legacy model can increase instructors' comprehension of the guiding learning principles associated with designing effective learning environments.

The process does not need to start from scratch and can build on methods already used in the classroom. It is clear that the *challenge* and *generate ideas* components are key features to drive the learning process, while the other phases can be more flexible depending on the needs of the learners and the available resources. The critical component is that students are engaged in a process of taking action on their current knowledge and reflecting on and refining that knowledge through feedback provided in the learning environment.

Concluding Thoughts and Lessons Learned

It is often noted that technology is really just a tool to support learning and instruction, but it is less often noted or made clear that the nature of the technology tools and their manner of use matters deeply (see for example CTGV, 1996). Our thinking about these issues is the product of over 18 years of work attempting to forge a linkage between learning theory and instructional practice, mediated by advances in technology. As our work evolved over time it became increasingly clear that multiple characteristics of powerful learning environments are needed for maximal payoff. Technology allows for this possibility but also adds levels of complexity. Fortunately, one of technology's affordances is that it is simultaneously capable of supporting the structuring and management of that complexity in ways that also enhance the instructional process. The STAR.Legacy software shell, a technology-based design for inquiry learning, emerged as one example of how to connect learning theory, instructional design and management, and the deployment of multiple technologies.

As the work on *Anchored Instruction* evolved over time, and as technologies became more sophisticated and more powerful, we were able to harness those technologies to instantiate important components of powerful learning environments. We also came to appreciate the fact that there were multiple components that needed to be present and in balance. This was not something we knew at the start and it emerged from multiple attempts to develop materials and practices that were driven by cognitive theory, implement them in real classrooms and instructional settings, study what worked and did not work and why, and then go back and improve the designs. Along the way *Anchored Instruction* got more complicated because what was important was not just the presence of an interesting anchor problem or situation or challenge. Rather, it was the way in which the anchor was used to set up and drive a process of inquiry-based learning that made the difference. So while things got more complicated they also became more straightforward in the sense that the value of a good anchor and what it means to have a good anchor is defined by everything one does with the anchor situation and the affordances it has to support meaningful and deep learning. Not every problem or situation or challenge has these properties and thus it has become an interesting issue to identify with domain experts what are big ideas and important challenges that can anchor a sustained set of instructional activities in a domain.

There are, of course, any number of issues and problems remaining to be solved when it comes to the unpacking of an idea like *Anchored Instruction* and they range from the theoretical to the practical. Also, we would never claim that *Anchored Instruction* is the solution for all instructional and learning situations. That would be a patently absurd claim to make and it would ignore the fact that there are multiple types of knowledge to acquire in a domain and multiple ways in which the learning can and should take place. Nonetheless, if we think in terms of the features of powerful learning environments, we can begin to see where *Anchored Instruction* fits in terms of enabling many critical features of instructional design. What is most interesting to us is that in the process of trying to connect theory, research, practice, and technology, we have gained a deeper sense of the interplay among them and the need to balance academic, aesthetic, and practical needs. Hopefully, in the process the work of the CTGV has provided some food for thought about important principles of cognition and instruction while also providing some useful instructional materials and practices for students and their teachers. This is a collective set of endeavors that we believe Norbert Seel would value highly.

References

Barron, B. J., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., Bransford, J. D., & the Cognition and Technology Group at Vanderbilt. (1998). Doing with understanding: Lessons from research on problem and project-based learning. *Journal of Learning Sciences., 7,* 271-312.

Barron, B., Vye, N. J., Zech, L., Schwartz, D., Bransford, J. D., Goldman, S. R., Pellegrino, J., Morris, J., Garrison, S., & Kantor, R. (1995). Creating contexts for community-based problem solving: The Jasper Challenge Series. In C. N. Hedley, P. Antonacci, & M. Rabinowitz (Eds.), *Thinking and literacy: The mind at work (*pp. 47-71). Hillsdale, NJ: Erlbaum.

Bereiter, C. (1990). Aspects of an educational learning theory. *Review of Educational Research, 60,* 603-624.

Birol, G., McKenna, A., Giorgio, T., & Brophy, S. (2002). *Implementation of educational modules in a biotechnology course: A challenge based education approach.* Montreal, CA: American Society of Engineering Education.

Biswas, G., Schwartz, D., Bhuva, B., Bransford, J., Brophy, S., Balac, T., et al. (1997). *Analysis of student understanding of basic electricity concepts* (Technical Report TR-CS-97-01). Nashville, TN: Vanderbilt University.

Bransford, J.D., Vye, N., Bateman, H., Brophy, S. P., and Roselli, R. (2004). Vanderbilt's AMIGO Project: Knowledge of how people learn enters cyberspace. In T. M. Duffy and J.R. Kirkley, *Learner centered theory and practice in distance dducation*. Mahwah: NJ: Erlbaum.

Bransford, J. D., Brown, A. L., Cocking, R. R., Donovan, S., & Pellegrino, J. W. (2000). *How people learn: Brain, mind, experience, and school (expanded edition).* Washington, DC: National Academy Press.

Bransford, J. D., Franks, J. J., Vye, N. J., & Sherwood, R. D. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 470-497). New York, NY: Cambridge University Press.

Bransford, J. D., & Schwartz, D. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education, 24*, 61-100.

Bransford, J. D., Zech, L., Schwartz, D., Barron, B., Vye, N., & the Cognition and Technology Group at Vanderbilt. (1996). Fostering mathematical thinking in middle school students: Lessons from research. In R. J. Sternberg & T. Ben-Zeev (Eds.), *The nature of mathematical thinking* (pp. 203-250). Mahwah, NJ: Erlbaum.

Bransford, J. D., Zech, L., Schwartz, D., Barron, B., Vye, N., & the Cognition and Technology Group at Vanderbilt. (1999). Designs for environments that invite and sustain mathematical thinking. In P. Cobb (Ed.), *Symbolizing, communicating, and mathematizing: Perspectives on discourse, tools, and instructional design* (pp. 275-324). Mahwah, NJ: Erlbaum.

Brophy, S., Biswas, G., Katzlberger, T., Bransford, J., & Schwartz, D. (1999) Teachable Agents: Combining insights from learning theory and computer science. In S.P. Lajoie & M. Vivet (Eds.), *Artificial intelligence in education, Vol. 50* (pp. 21-28). Amsterdam: IOS Press.

Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229-272). Cambridge, MA: MIT Press/Bradford Books.

Brown, A. L., & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.), *Innovations in learning: New environments for education* (pp. 289-325). Mahwah, NJ: Erlbaum.

Bruer, J. T. (1993). *Schools for thought*. Cambridge, MA: MIT Press.

Cognition and Technology Group at Vanderbilt. (1992a). The Jasper series as an example of anchored instruction: Theory, program description, and assessment data. *Educational Psychologist, 27*, 291-315.

Cognition and Technology Group at Vanderbilt. (1992b). An anchored instruction approach to cognitive skills acquisition and intelligent tutoring. In J. W. Region & V. J. Shute (Eds.), *Cognition approaches to automated instruction* (pp. 135-170). Hillsdale, NJ: Erlbaum.

Cognition and Technology Group at Vanderbilt. (1992c). The Jasper experiment: An exploration of issues in learning and instructional design. *Educational Technology Research and Development, 40*, 65-80.

Cognition and Technology Group at Vanderbilt. (1993a). The Jasper series: Theoretical foundations and data on problem solving and transfer. In L. A. Penner, G. M. Batsche, H. M. Knoff, & D. L. Nelson (Eds.), *The challenge in mathematics and science education: Psychology's response* (pp. 113-152). Washington, DC: American Psychological Association.

Cognition and Technology Group at Vanderbilt. (1993b). Toward integrated curricula: Possibilities from anchored instruction. In M. Rabinowitz (Ed.), *Cognitive science foundations of instruction* (pp. 33-55). Hillsdale, NJ: Erlbaum.

Cognition and Technology Group at Vanderbilt. (1994a). From visual word problems to learning communities: Changing conceptions of cognitive research. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 157-200). Cambridge, MA: MIT Press/Bradford Books.

Cognition and Technology Group at Vanderbilt. (1994b). Multimedia environments for developing literacy in at-risk students. In B. Means (Ed.), *Technology and educational reform: The reality behind the promise* (pp. 23-56). San Francisco: Jossey-Bass.

Cognition and Technology Group at Vanderbilt. (1996). Looking at technology in context: A framework for understanding technology and education research. In D. C. Berliner & R. C. Calfee (Eds.), *The handbook of educational psychology,* (pp. 807-840). New York: Simon & Schuster MacMillan.

Cognition and Technology Group at Vanderbilt. (1997). *The Jasper Project: Lessons in curriculum, instruction, assessment, and professional development*. Mahwah, NJ: Erlbaum.

Cognition and Technology Group at Vanderbilt. (1998). Designing environments to reveal, support, and expand our children's potentials. In S. A. Soraci & W. McIlvane (Eds.), *Perspectives on fundamental processes in intellectual functioning* (pp. 313-350). Norwood, NJ: Ablex.

Cognition and Technology Group at Vanderbilt. (2000). Adventures in anchored instruction: Lessons from beyond the ivory tower. In R. Glaser (Ed.), *Advances in instructional psychology: Vol. 5. Educational design and cognitive science* (pp. 35-99). Mahwah, NJ: Erlbaum.

Cognition and Technology Group at Vanderbilt. (2002). Connecting learning theory and instructional practice: Leveraging some powerful affordances of technology. In H. O'Neill & R. Perez (Eds.), *Technology applications in education: A learning view* (pp. 173-209). Mahwah, NJ: Erlbaum.

Collins, A. (1996). Design issues for learning environments. In S. Vosniadou, E. DeCorte, R. Glaser, & H. Mandl (Eds.), In*ternational perspectives on the psychological foundations of technology-based learning environments* (pp. 347-362). Hillsdale, NJ: Erlbaum.

Collins, A., Hawkins, J., & Carver, S. M. (1991). A cognitive apprenticeship for disadvantaged students. In B. Means, C. Chelemer, & M. S. Knapp (Eds.), *Teaching advanced skills to at-risk students* (pp. 216-243). San Francisco, CA: Jossey-Bass.

Crews, T. R., Biswas, G., Goldman, S. R., & Bransford, J. D. (1997). Anchored interactive learning environments. *International Journal of Artificial Intelligence in Education, 8* (2), 142-178.

Donovan, M. S., & Bransford, J. D. (2005). *How people learn History, Mathematics and Science in the Classroom.* National Academy Press: Washington, DC.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1-38.

Goldman, S. R., Lawless, K., Pellegrino, J. W., & Plants, R. (2005/06). Technology for teaching and learning with understanding, In J. M. Cooper (Ed.), *Classroom teaching skills* (8^{th} edition) (pp. 185-234). Boston, MA: Houghton Mifflin.

Goldman, S. R., Petrosino, A., Sherwood, R. D., Garrison, S., Hickey, D., Bransford, J. D., & Pellegrino. (1996). Anchoring science instruction in multimedia learning environments. In S. Vosniadou, E. De Corte, R. Glaser, & H. Mandl (Eds.), *International perspectives on the psychological foundations of technology-supported learning environments* (pp. 257-284). Hillsdale, NJ: Lawrence Erlbaum Associates.

Goldman, S. R., Zech, L. K., Biswas, G., Noser, T., & the Cognition and Technology Group at Vanderbilt. (1997). Computer technology and complex problem solving: Issues in the study of complex cognitive activity. *Instructional Science*, *27*(3-4), 235-268.

Harris, T. R., Bransford, J. D., & Brophy, S. P. (2002). Roles for learning science and learning technologies in biomedical engineering education: A Review of Recent Advances. *Annual Review of Biomedical Engineering, 4*, 29-48.

Lamon, M., Secules, T., Petrosino, A. J., Hackett, R., Bransford, J. D., & Goldman, S. R. (1996). Schools for thought: Overview of the project and lessons learned from one of the sites. In L. Schauble & R. Glaser (Eds.), *Innovation in learning: New environments for education* (pp. 243-288). Hillsdale, NJ: Erlbaum.

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life.* Cambridge: Cambridge University Press.

Martin, T., Rayne, K., Kemp, N. J., Hart, J., & Diller, K. R. (2005). *Teaching for Adaptive Expertise in Biomedical Engineering Ethics. Science & Engineering Ethics, 11*(2), 257-276.

Martin, T., Petrosino, A. J., Rivale, S., & Diller, K. R. (2006). The development of adaptive expertise in biotransport. *New Directions for Teaching & Learning*(108), 35-47.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Norman, D. A. (1993). *Things that make us smart: Defending human attributes in the age of the machine.* New York: Addison-Wesley.

Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 47-87). New York: Cambridge University Press.

Pellegrino, J. W. (2003). Connecting learning theory and instruction: Principles, practices and possibilities. In F. Achtenhagen & E. John (Eds.), *Milestones of vocational education and training. Vol. 1. The teaching-learning perspective.* (pp. 17-42). Bielefeld: Bertelsmann.

Pellegrino, J. W. (2004). Complex learning environments: Connecting learning theory, instructional design, and technology. In N. J. Seel & S. Dijkstra (Eds.), *Curriculum, plans, and processes in instructional design: International perspectives* (pp. 25-48). Mahwah, NJ: Erlbaum Associates.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Pellegrino, J. W., Hickey, D., Heath, A., Rewey, K., Vye, N. J., & the Cognition and Technology Group at Vanderbilt. (1991). *Assessing the outcomes of an innovative instructional program: The 1990-1991 implementation of the "Adventures of Jasper Woodbury"* (Tech. Rep. No. 911). Nashville, TN: Vanderbilt University, Learning Technology Center.

Perreault, E., Litt, M., & Saterbak, A. (2006). Educational Methods and Best Practices in BME Laboratories1. *Annals of Biomedical Engineering, 34*(2), 209-216.

Resnick, L. (1987). Learning in school and out. *Educational Researcher, 16* (9), 13-20.

Rogoff, B. (1990). *Apprenticeship in thinking*. New York: Oxford University Press.

Roselli, R. J., & Brophy, S. P. (2006a). Effectiveness of challenge-based instruction in biomechanics. *Journal of Engineering Education, 95*(4), 311-324.

Roselli, R. J., & Brophy, S. P. (2006b). Experiences with formative assessment in engineering classrooms. *Journal of Engineering Education, 95*(4), 325-333.

Scardamalia, M., & Bereiter, C. (1994). Computer support for knowledge-building communities. *The Journal of the Learning Sciences,* 265-285.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition & Instruction,16*, 475-522.

Schwartz, D. L., Brophy, S., Lin, X., & Bransford, J. D. (1999). Flexibly adaptive instructional design: A case study from an educational psychology course. *Educational Technology Research and Development, 47,* 39-59.

Schwartz, D. L., Goldman, S. R., Vye, N. J., Barron, B. J., and Bransford, J. D., & the Cognition and Technology Group at Vanderbilt. (1998). Aligning everyday and mathematical reasoning: The case of sampling assumptions. In S. P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in grades K-12* (pp. 233-273). Mahwah, NJ: Erlbaum.

Schwartz, D. L., Lin, X., Brophy, S., & Bransford, J. D. (1999). Toward the development of flexibly adaptive instructional designs. In C. M. Reigeluth (Ed.), *Instructional design theories and models: Vol. II. A new paradigm of instructional theory.* Hillsdale, NJ: Erlbaum.

Secules, T., Cottom, C. D., Bray, M. H., Miller, L. D., & the Cognition and Technology Group at Vanderbilt. (1997). Schools for Thought: Creating learning communities. *Educational Leadership, 54*(6), 56-60.

Sharp, D. L. M., Bransford, J. D., Goldman, S. R., Risko, V. J., Kinzer, C. K., & Vye, N. J. (1995). Dynamic visual support for story comprehension and mental modal building by young, at-risk children. *Educational Technology Research and Development, 43,* 25-42.

Sherwood, R. D., Petrosino, A. J., Lin, X., Lamon, M., & the Cognition and Technology Group at Vanderbilt. (1995). Problem-based macro contexts in science instruction: Theoretical basis, design issues, and the development of applications. In D. Lavoie (Ed.), *Towards a cognitive-science perspective for scientific problem solving.* Manhattan, KS: National Association for Research in Science Teaching.

VaNTH Engineering Research Center. (2007). Retrieved on September 13, 2007, from www.vanth.org

Vye, N. J., Goldman, S. R., Voss, J. F., Hmelo, C., Williams, S., & the Cognition and Technology Group at Vanderbilt. (1997). Complex mathematical problem solving by individuals and dyads. *Cognition and Instruction, 15,* 435-484.

Vye, N. J., Schwartz, D. L., Bransford, J. D., Barron, B. J., Zech, L., & the Cognition and Technology Group at Vanderbilt. (1998). SMART environments that support monitoring, reflection, and revision. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 305-346). Mahwah, NJ: Erlbaum.

White, B. C., & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16* (1), 3-118.

Williams, S. M. (1994). *Anchored simulations: Merging the strengths of formal and informal reasoning in a computer-based learning environment*. Unpublished doctoral dissertation, Vanderbilt University, Nashville.

Williams, S. M., Burgess, K. L., Bray, M. H., Bransford, J. D., Goldman, S. R., & the Cognition and Technology Group at Vanderbilt. (1998). Technology and learning in schools for thought classrooms. In C. Dede (Ed.), *1998 ASCD year book: Learning with technology* (pp. 97-119). Alexandria, VA: ASCD.

Zech, L., Vye, N., Bransford, J., Goldman, S., Barron, B., Schwartz, D., Hackett, R., Mayfield-Stewart, C., & the Cognition and Technology Group at Vanderbilt. (1998). An introduction to geometry through anchored instruction. In R. Lehrer & D. Chazan (Eds.), *New directions in teaching and learning geometry.* Hillsdale, NJ: Erlbaum.

Zech, L., Vye, N. J., Bransford, J. D., Swink, J., Mayfield-Stewart, C., Goldman, S. R., & the Cognition and Technology Group at Vanderbilt. (1994). Bringing geometry into the classroom with videodisc technology. *Mathematics Teaching in the Middle School Journal (MTMS),* 1 (3), 228-233.

15. Complex Skills Development for Today's Workforce

Using Games as a Strategy for Engineering Model-centered Learning Environments

Tristan E. Johnson[1] and Wen-hai D. Huang[2]

[1] Florida State University

[2] University of Illinois at Urbana-Champaign

Abstract: Today's workforce is presented with many challenges including the increasing level of information and technology as well as the loss of expertise. In order to be competitive and successful, organizations are looking for workers that have complex skills. Educational institutions are trying to meet these challenges, but the tradition methods are not enough to prepare learners for current challenges. This chapter presents the benefits of using games as model-centered learning environments. We present how games have the characteristics of MCLE. Even while lacking empirical evidence, games (as a type of MCLE) have the design features to support complex skills development. As such games have the ability to prepare the upcoming workers to meet the challenges found in today's workforce.

Keywords: Mental models; instructional gaming design; complex skills development.

Introduction

With the acceleration of scientific and technological progress, human workforce requirements and demands are changing. Increased international competition and rapid technological advances require today's workforce to have the ability to adapt to changing technologies. With this evolution, skills obtained early in one's career may become obsolete. Many technical fields are specifically dealing with this issue of changing demands and requirements on the workforce (Freeman & Aspray, 1999).

Technological progress has an impact not only on how we do things, but also how we think. What it means to "know" was once being able to remember and repeat information, but now "knowing" is being able to find and use information (Simon, 1996). Current and future work tasks favor strong non-routine cognitive skills such as problem solving, abstract reasoning, communication, and collaboration (Karoly & Panis, 2004). Learning has now become a continuous process throughout ones life that involves training as well as retraining that continues well past initial entry into the workforce. To better meet the needs of the workforce, consideration should be given to how various educational and training systems can adapt to meet such needs.

Educational and training programs can not longer single handedly cover the vast amounts of knowledge and information that are being generated. The field of Education is starting to align its missions in ways that address the need for workforce change (Bransford, Brown, & Cocking, 2000). It is important that educational goals focus towards helping students develop learning strategies and cognitive tools so that students can acquire knowledge in order to think productively. Students need to acquire a basic understanding about key subjects but also they need to learn how to ask meaningful questions in order to better understand the principles of learning so that they can become themselves, lifelong learners.

It has been said that the highest purpose of education is not to just teach facts however important they may be, but to train the mind, to make good citizens, and to develop character. While these are laudable goals, the process that one journeys is not accomplished in a few years nor is it limited to a specific academic or training setting. This process extends beyond the formal structure of an educational institution and it occurs throughout one's lifetime.

Lifelong learning can occur in less formal settings such as the workplace. In fact, the workplace provides tremendous opportunities supporting human development and growth. When experts are asked where most of their learning comes from, they typically respond that most of what they learned came from "on the job" experience and not from the schoolhouse. Workplace experience plays a crucial role in the development of experts, yet this kind of learning is often neglected when it comes to the planning and development of organizational learning policies and strategies.

Experiences do not automatically constitute learning or skill development. There are cases where experiences do not equate to skill development. But experiences coupled with reflection and debriefing can provide the critical learning activities as well as related feedback to influence productive thinking and ultimately complex skills development (Ertmer, Johnson, & Lane, 2001; Kagan, 1992; Seel, 2003a; Stolurow, 1973).

Experiences are linked to the environments where they typically occur. Workplace environments, while offering experiences that are unique to a particular setting, are not available to non-employee populations. In order to obtain a specific experience that may not be available to non-employees, simulation environments can be created to provide experiences to learners in order to facilitate complex skills development prior to entering into the workforce. These experiences may have a meta-level effect by helping the learner better understand the overall process for lifelong learning.

There are various instructional strategies that can be used to prepare students prior to going in to the workforce. One strategy is to use simulations or models to provide a mechanism to help the learner gain experience. Gaining experience facilities workplace performance in areas where skills are complex. Experiential learning are typically better suited for workforce preparation in domains that involve complex skills such as engineering, science, finance, information technology, health care, and military. Complex skills involve tasks like decision-making, strategy development, and also ill structured problem solving. Because these skills are multifaceted, there are some domains, where professional licensure requires the successful completion of an internship or practicum where the student spends months engaged in supervised practice.

One of the key differences between lower performers and higher performers according to Dreyfus and Dreyfus' (1986) five stage skill acquisition process (novice, advanced beginner, competence, proficiency, and expertise), is that someone at a lower stage may have similar cognitive processes as someone at a higher stage, but the lower performers will perform poorly if deficient in actual experience. In order to progress through the states, one needs to acquire experience. Similar to the a problem-based case method, the critical goal for experiential learning is to give learners experience in the sort of decision making required later in their academic and professional careers (Reigeluth & Frick, 1999).

Game-based MCLE can be used prior to entry into the workplace. A game-based MCLE can be used in an academic setting, but also could easily be used in a training/workplace setting. For example, in the military, often soldiers are involved more so in training settings and less involved in the actual performance setting, but they are seeking to gain experience using virtual settings. There are some tasks that are rare and are not often performed in a work setting, or that take a long time to complete a session. Managing a Mutual Fund would be an example of this. Decisions are made, but the implications of the decisions are not materialized for many days and months. In these cases, a game where a player can experience a set of tasks, and engage in decision making in a relatively short time period would merit the use of a gaming strategy.

It has been reported in an exhaustive study (Johnson, Spector, Huang, & Novak, 2007), that games can be used for a wide variety of learning outcomes. The gaming literature suggests that players might experience knowledge gain, skill improvement, and attitudinal change as the result of game playing. Several specific learning outcomes are linked to games, but there was no evidence of an empirical relationship. The gaming literature is not specifically helpful in helping game designers decide when use games as a specific strategy to train to a targeted learning outcome. However, it seems reasonable to use an instruction game strategy for more complex skills development (Johnson et al., 2007). Other instructional strategies would be more appropriate and cost effective for facilitating learning of simpler learning outcomes such as knowledge domain outcomes including comprehension and understanding; factual information; and procedures and processes. No matter what the specific focus for the instruction, the key aspects of a game-base strategy is related to its' ability to model.

In the most basic sense, a simulation is a model or a representation of something. This definition of an operating model contains a dynamic relationship of two or more variables (Reigeluth & Schwartz, 1989). From a learning perspective, an instructional simulation is a program containing a manipulatable model of a real or a theoretical system. Simulations are designed to allow the user to interact with, respond to, and manipulate the underlying model. A simulation allows students to change the model from a given state to a specified goal state by directing it through a number of intermediate steps. Thus, a simulation program may accept commands from the user, alters the state of the model, and when appropriate displays the new state (Thomas & Hooper, 1991). This functionality permits the student to actually carryout the learning activities in a context, constrained by the simulation, yet similar to the real world. The simulation basically simplifies a real-world system in order to help the student solve problems, learn procedures, understand phenomena, and practice skills safely and efficiently (Johnson, 1999).

By focusing more on the experience, instructional simulations can be considered experiential exercises as students can uniquely interact with a knowledge domain by being part of the model (Gredler, 1996). Gredler points out that specific experiential simulations are designed to factor in the student as one of the several dynamic variables. In this sense, simulations are used as heuristic devices for understanding complex relations within a specific domain (Thomas & Hooper, 1991).

In addition to the benefits associated with higher-order learning, simulations can be in certain cases used for teaching facts and knowledge. As compared with interactive tutorials or drills, simulations can be designed to present information and query the level of understanding by question-answer techniques, but also, they can help guide students in acquiring information or skills. Beyond supporting these cognitive levels, tutorials and drills do not offer much. However, simulations go beyond this to provide support for students to engage in practice skills, thereby reinforcing and increasing fluency in a specific knowledge domain.

Unlike a tutorial, a simulation usually does not provide explicit feedback informing the student what action to take, nor does it explain commission or omission errors. It merely responds to the student's actions in accordance with the governing rules of the simulation. On the other hand, a simulation differs from a programming environment in that it has inherent goals that the user is expected to pursue (Thomas & Hooper, 1991).

Simulations have been traditionally used in settings to mitigate the negative effects of environments associated with extreme cost and/or risk. While these are important issue to consider, sustaining motivation for many tasks is problematic. Gaming technology while similar to modeling and simulations, has a unique objective to support significant levels of sustained motivation as well as the shared objectives of providing opportunities for risk taking with minimal negative consequences. While the primary focus of games is on entertainment, games like simulations are a model or a representation of something whether real or hypothetical. Games like simulations are environments that provide experiences for the players in addition to supporting high levels of engagement. While levels of fidelity can differ greatly, gaming environments, whether action games or games of chance, involve the development of mental models as well as system modeling. Due to the potential benefits of games, researchers are seeking to understand and harness the benefits of games in order to impact learning.

In the reminder of this chapter, we will describe how games are a type of model-centered learning environments and then delineate gaming feature as an example of how to engineer a model-centered learning environment with the specific intent to help develop complex skills that apply to a real-world setting.

Games and Model-Centered Learning Environments

While there has been criticism that the field of Instructional Design is faltering in meeting the demands of complex skills development, there has been a stream of arguments that there are positive effects from using model-centered learning environments (MCLE) as a means of dealing with the scientific and societal demands. (Seel, 2003). Based on research investigating the effects of modeled-centered learning ((Johnson-Laird, 1989; Seel, 1991, 1999, 2003a, 2003b, 2003c; Seel, Al-Diban, & Blumschein, 2000; Seel & Schenk, 2003), the presentation, developing, and sharing of mental models allows for the ideal interaction between instructional materials and learner (Reigeluth, 1999).

Games have captured a tremendous amount of attention in recent years. The instructional and training application of games are pervasive in many different types of contexts (K-12 education, higher education, military, and corporate sectors). The main reason given for using instructional games is the belief that they have an ability to engage learners in relevant cognitive and attitudinal aspects of learning. Since playing video games is becoming a primary social activity favored by young learners, educators and employers are investing considerable resources in games as a means to communicate with, motivate, and occasionally train students and employees.

Can a game serve as a meaningful tool to support instruction and facilitate learning? If so what are the key characteristics in games that provide instruction and that facilitates learning? Evidence of the suitability of games for training specific types of skills and knowledge is crucial in order for game-based training solutions to be effective, however such evidence is practically non-existent (Johnson et al., 2007). To give credence that games are a substantive methodology to support learning, we argue that because games are a type of MCLE, they have all of the benefits of simulations, and also support leaner engagement and motivation, games can be effective instructional interventions.

Cognition and learning are a result of a construction of mental models (Seel, 2003b). MCLEs aim to provide support for learners to generate their own of mental models. This support can be in the form of an externally represented model such as games. Models provide four functions that support learning. Consider these functions in the context of games. Models aid in the:

- Simplification of a specific phenomena in a closed domain
- Visualization of a complex system
- Construction of analogies that help identify components in an unknown domain
- Simulation of a system's process.

Games provide students with an experience of the system that they are trying to learn. In this sense the game facilitates learning through experience (Carlson, 1991). Within a game, the game model can be presented as a conceptual model where it can direct the learner's interaction with the learning material.

In previous work where we studied the instructional effects of games, it became apparent that while gaming technology has a strong link to attention getting, satisfaction, and relevance, games are truly linked to specific systems models. Certain games simulate a simple model while others embody a complex systems model. Nevertheless, the underlying challenge in any game is to understand the game model in order to more effectively play the game. There is evidence suggesting that teammates who think in similar ways about difficult problems are likely to work effectively together (Cannon-Bowers & Salas, 1998; Guzzo & Salas, 1995; Hackman, 1990). We believe that a similar effect occurs when a player's mental model converge with the game model. As players better understand the game model, they can better predict the outcomes of their decisions. As mental models become more similar with the game model, the player's decision-making is more effective and efficient during game play.

Several game characteristics are aligned with modeled-centered learning environments. These characteristics include the following.

- Games are a model of a system (real or fantasy).
- Games are a mechanism for learner to experience the system model. Players interact with the systems model to determine what are the components and how are they interrelated.
- Games help players see the system as a whole. This experience provides the players with information they can then develop a system model. As the player's mental models align with the game model, they will be able to compete more successfully in the game.
- Games are designed to be entertaining and thereby promote an interest in and engagement with the game model.
- Games can simulate risk taking (specific experiences that maybe user may never get) without the effects of negative consequences.

Games as an MCLE Strategy

What is a game and how is it related to MCLEs? A game is a context in which individual and teamed players, bounded by rules and principles, compete in attaining the identified game objectives. There is a series of decision-making processes are required by the game players. Avedon and Sutton-Smith (1971) assert that game playing is a voluntary exercise of controlling a system (i.e., the game) intended for a state of disequilibrium. The system is purposefully designed to keep the participants of the system engaging in activities to decrease the degree of disequilibrium. In other words, game players continuously try out new methodologies and strategies during the game playing process based on the system's feedback till they achieve the game objectives or the equilibrium state. The game playing process is very similar to the learning process intended in a Model Centered Learning Environment (MCLE). Learners immersed in the MCLE constantly revise their intrinsic mental models by comparing them with the extrinsic models motivated by the disequilibrium between the outcome of the extrinsic models and learners' intrinsic mental models. The final product of MCLE therefore is learners' mental models that produce compatible performance outcome (e.g., solving an algebra application problem) as does the extrinsic model (i.e., instructions given the instructor). The following section will explore the relationships between games and MCLE with four aspects:

- Games create experiences
- Learning occurred by interactions in games
- Games are complex
- Games are models.

Games are Experiences

Games are known for their capabilities to promote learning-by-doing type of collaborative and active learning (Downes, 2004; Klabbers, 2006; Vygotsky, 1978). Game players learn from their success and mistakes in order to improve their gaming skills and winning strategies. Players learn about the games and how to win the games by experiencing the interactions occurred during the process. The process begins with concrete playing experience. Players observe how the system responds to their actions in the form of scoring, for example. Players then revise their playing strategies and try them out at similar situations. A player can play the game for a long time and never quite get things (experience alone is not sufficient). However, at some point, a player who thinks about what they are trying to learn (reflection and self feedback) can begin to master the game. This happens as the player considers the game experience and makes meaning out of the interaction. Sometimes the player creates an assumption about the game model and then it can get tested out. Other times, there is enough experience to support a conjecture about the game model that there is not need to test it out. In more complex games, the model may not be totally clear and requires the user to explore the game more.

Players Learn about the System by Interacting with it

Games requiring individual participation offer a rich environment for player interaction with the game system (Pivec & Dziabenko, 2004). Systems within games consist of rules, process, and interactions that players must experience in order to attain the desired outcome, typically winning the game. System rules are a crucial aspect of games. Effective game play is governed by rules and principles. To be competitive, players must familiarize themselves with the rules at the earliest stages of game playing. These processes of familiarization of rules in games help players connect the game context with their existing schema. Rules impose limits and guidelines on actions and to ensure that all players take the same paths to achieve the game goals. Rules also represent the evaluation criteria in the form of scoring, and also provide a measure of players' performance during the process. Rules further enable players to analyze the interrelationships between different rules in order to generate feasible and "winning" strategies (Bennett & Warnock, 2007; Björk & Holopainen, 2003; Crookall, Oxford, & Saunders, 1987; Garris, Ahlers, & Driskell, 2002; Hays, 2005; Leemkuil, de Jong, & Ootes, 2000). Interactions in games are considered as structural components allowing players to interface with other players, game context, and the system. It is the interaction component that makes the game rules and the playing process meaningful. (Ang, Avni & Zaphiris, 2007; Asgari, 2005; Crawford, 1982; de Felix & Johnson, 1993; Garris et al., 2002; Kirriemuir & McFarlane, 2006; McGrenery, 1996; Waal, 1995). The interactions within a game allow players to directly acquire first-hand experience to learn about the system. From an MCLE viewpoint, players learn about the model represented by the game by exchanging timely actions and getting immediate feedback with the actual model. Learning is facilitated by the processes of "knowing the rules in the system" and "applying the rules in the system".

Games are Complex and Constructed with Interrelated Pieces

An engaging game is often more complex than a boring one. Players must consider multiple factors before making a winning decision. Sim City™, for example, players are responsible of planning, developing, and sustaining a prospering city. Building a new hospital is usually an effective strategy to attract new residents to move to the "simulated" city. However, it also could backfire if the city infrastructure is not readily available to support the hospital such as the lack of power plants, operational mass transportation systems, safe residential areas, and decent public schools. Therefore the decision-making process for players is everything but simple and straightforward. Another example is the Sudoku game. One falsely placed number would fail the whole logic-based number placement game. One game play or action could impact the overall outcome since all pieces are causally connected by the game rules and intended processes. Games are capable of relating all intended information together and hence to create a complex learning environment, which helps learners in MCLE see the compound nature of given model and develop transferable and foreseeing problem-solving strategies.

Games are Representations of Realistic Models

Games embody abstract concepts and rules in MCLE. The winning game play or winning strategy is the direct translation of problem-solving strategies intended by the model. The game adds contextual information to the model as to how to apply it in different situations. The contextual information is often represented by a storyline that implicitly or explicitly guides the players throughout the process. Simulation games, for example, are powerful in creating authentic situations for players to experience realistic and immediate performance feedback and can be an application-based duplication of the intended model. Learner playing simulation games are directly interfacing with the intended model in a tangible way.

Game Design

The following sections will describe the game-based learning environment in more detail in terms of their characteristics, learning outcomes, and design approaches.

Game's Characteristics

An exhaustive literature review has identified twelve prominent and yet interrelated characteristics found in existing games regardless of their delivery formats (see table 1). While games do not need all of the characteristics to be classified as a game, a game does embody several characteristics. The quality of the game does not depend on the number of characteristics, but rather on the ability to successfully embody the characteristics in the game design.

Table 1. Game Characteristics

1.	Challenge	7.	Story or Representation
2.	Competition	8.	Engagement or Curiosity
3.	Rules	9.	Role Playing
4.	Goals	10.	Control
5.	Fantasy	11.	Tasks
6.	Changed Reality	12.	Multimodal Presentation

Challenge—A challenging activity provides an achievable level of difficulty for game players, which consists of clearly identified task goals, unpredictability, immediate performance feedback, and a sense of accomplishment and conquering after completing the activities (Ang et al., 2007; Baranauskas, Neto, & Borges, 1999; Belanich, Orvis, & Sibley, 2003; Bennett & Warnock, 2007; Csikszentmihalyi, 1990; Garris et al., 2002; Malone, 1981; Malone & Lepper, 1987; McGrenery, 1996; Rieber, 2001).

Competition—Competition stimulates players to take risk-taking actions in a consequence-free environment enriched with social interactions. Players develop their skills during the game playing process by matching and exceeding the opponents' skill levels. The competition can be implemented between individual players, amongst teams, and even between players and the system (Baranauskas et al., 1999; Crawford, 1982; Crookall et al., 1987; Csikszentmihalyi, 1990; Leemkuil et al., 2000; Rieber, 2001; Vockell, 2004).

Rules—Rules of games serve as the guidelines for players' actions. Fair play is also sustained by the enforcement of game rules. Players need to learn about the game rules either by designated training or via the actual playing experience. In the context of games for learning, game rules could be the direct or indirect translations of intended instructional materials such as scientific concepts of economic principles (Bennett & Warnock, 2007; Björk & Holopainen, 2003; Crookall et al., 1987; Garris et al., 2002; Hays, 2005; Leemkuil et al., 2000).

Goals—Goals in games clearly state the final status for players to attain via series of planned tasks and actions by following the rules of game. Sub-goals usually are often presented to present various stages of accomplishment for motivational and evaluation purposes. The presence of goals is also the major difference between games and simulations (i.e., simulations could be goal-less) (Ang et al., 2007; Bennett & Warnock, 2007; Björk & Holopainen, 2003; Crookall et al., 1987; Csikszentmihalyi, 1990; de Felix & Johnson, 1993; Gredler, 1996; Hays, 2005; Hirumi, 2006; Leemkuil et al., 2000; Malone, 1980).

Fantasy—Fantasy creates entirely unreal situations and environments for game players to experience. This characteristic encourages players to take risks in a safe environment. Fantasy also motivates players to follow the storyline to achieve desired game goals (Bennett & Warnock, 2007; Garris et al., 2002; Kasvi, 2000; Kirriemuir & McFarlane, 2006; Malone, 1981; Malone & Lepper, 1987; McGrenery, 1996).

Changed Reality—Changed reality in games allows players to have exaggerated experiences in a specific context, which must reflect to certain degree of reality, but not entirely. Usually changed reality separates itself from reality by altering time, space, role-playing, and the complexity of situations (e.g., simplified reality) (Belanich et al., 2003; Björk & Holopainen, 2003; Crawford, 1982; Csikszentmihalyi, 1990).

Storyline or Representation—Storyline or representation in games provides paths for players to interact, react, and progress. It summarizes the game goals, rules, constraints, role playing, and contexts for players in a seamlessly interconnected and embedded fashion (Ang et al., 2007; Hirumi, 2006; Rieber, 2001). Players usually favor the representation of game rules in stories since it not only informs them the playing guidelines, but also provides a holistic view of the entire game context.

Engagement and Curiosity—These two characteristics are frequently discussed in game design literatures. Engagement created by games allows players to become deeply involved in the game that players lose the sense of realistic self. In other words, players perceive themselves as part of the game and enjoy the intrinsically motivating game playing experiences. Another way to describe an engaging experience is that players lost the track of time when playing the game. Playing the game itself is rewarding enough without extrinsic motivators. Implementing elements of mystery and curiosity is also considered effective in creating engaging game experiences (Asgari, 2005; Bennett & Warnock, 2007; Csikszentmihalyi, 1990; Kasvi, 2000; Leemkuil et al., 2000; Malone, 1980; Malone & Lepper, 1987; McGrenery, 1996).

Control—Control in games enables players to determine and predict the outcome of actions or events. It is essential for players to perceive the sense of "being in control" thus to be intrinsically motivated to complete the game. Providing options or choices to players, for example, is effective approach to allow players to exercise their control over the progression of the game (Belanich et al., 2003; Bennett & Warnock, 2007; Crookall et al., 1987; Csikszentmihalyi, 1990; Garris et al., 2002; Gredler, 1996; Hays, 2005; Kasvi, 2000; Malone, 1981; Malone & Lepper, 1987; McGrenery, 1996; Waal, 1995).

Role-Playing—Role-playing in the game are characters embedded in the storyline of the game. It is a popular format of game playing known as RPG (Role-Playing Games). Usually the player's role is pre-identified with specific position, access to resource and control, dominance over the progression of the game, functionality (if within a team), and behavioral patterns. Role-playing also helps players establish connection with the fantasy world of the game in order to better engage players with the game playing experience (Ang et al., 2007; Björk & Holopainen, 2003; Gredler, 1996).

Task—Task in the game is the building block of a game's goal. Players often are required to take on sequences of tasks along with the progress of the game in order to achieve the game's final goals. Players accumulate experiences upon completions of given tasks. Formative evaluations of accomplished tasks further help players improve their playing strategies (Björk & Holopainen, 2003; Gredler, 1996). From the viewpoint of instructional design, games tasks embody the outcome of varies levels of task analysis and prepare players for the accomplishment of intended learning goals.

Multimodal Presentation—Games usually utilize multimodal presentation to effectively enhance the instructional effect. This is particularly true in video games. Aural, visual, and textual presentations are combined in order to enrich the experience. Animations, for example, are popular as major game components since they seamlessly integrate multimodal presentations and can be easily modified for different game contexts (Belanich et al., 2003; Bennett & Warnock, 2007; Björk & Holopainen, 2003; de Felix & Johnson, 1993; McGrenery, 1996).

The characteristics discussed above are identified at the latent level. We must empirically examine their relationships among them to better understand their design implications in developing effective MCLE.

Design Approaches

Design of game-based MCLE needs to be multidimensional in order to afford an environment for developing complex knowledge base and skill sets. In this section, we propose three approaches in attempting to include the design of interface (player interaction), the design of learning (learning theory), and the process of design (instructional system design), to encapsulate the basic design requirement.

Design From the Player Interaction Viewpoint

This perspective is crucial to achieve the development of an effective game-based MCLE. The design of interaction must include all aspects of learning and game playing. Game players need not only interact with the game (i.e., the system), but in certain cases they also need to maintain meaningful communication with other players. Players also need to have access to the learning environment at large if the game is only part of the MCLE. As discussed earlier, gaining experience is a critical component if players are going to learn from an instructional game. Interacting with all participating elements during the entire playing process is the most effective way to attain desired learning outcome. Game tasks, for example, are often the focal point of interactions in the game. Players must seek out meaningful interactions with the system for game rules and available resources, and peer players (if available) for proven winning strategies.

Design from the Learning Theory Viewpoint

Games are known for their experiential learning approach to develop players' knowledge, skills, and even change their attitudes. Studies have argued that all types of learning theories are applicable in the gaming environment. Behavioral learning can be easily found in simple drill and practice games; the complex game environments are suitable for applying cognitive information processing principles in the context of multimedia learning (Mayer, 2001) or Cognitive Load Theory (Chandler & Sweller, 1991); the same complex game setting can easily afford an ill-structured context for problem-solving activities. The design issues based on learning theories are not only about "which one to apply", but also "how to apply all of them together". An eclectic collection of learning theory, as the foundation of game design in MCLE is necessary for designers, to fully utilize the game's multi-layered features in providing enhanced learning experiences.

Design from the Instructional System Design Viewpoint

In addition to the interaction and learning theories, game design for MCLE also needs to consider the ISD perspective of the design process (Hirumi, 2006). Similar to the design of other instructional methods, game design requires adequate needs assessment, audience analysis, learning and performance task analysis, instructional strategy development, implementation, and formative evaluations, in order to attain effective design outcome. Grounded systematic instructional design approach is well applicable and should be fully applied when designing game-based programs. However, in light of instructional strategies, some are better adopted in games than others. The literatures suggests that active learning, experiential learning, situated cognition, apprenticeship learning, user-centered learning, cooperative learning, and scaffolding are viable instructional strategies in game settings owing to their abilities to induce highly interactive and engaging learning process (Bonk, 2005; Egenfeldt-Nielsen, 2006; Gros, 2006; Hays, 2005; Klabbers, 2006; Squire, 2005; Vygotsky, 1978). However, designers need to be cautious so as to not over motivate learners with embedded game characteristics such as fantasy, challenges, competitions, and multimedia-rich environments.

Design from the Modeling Viewpoint

Craik (1943) described the process of model-based learning with the following steps: learners identify the context and elements of a problematic situation, learners develop their own mental models based on identified problem(s), learners develop tentative solutions for the problem without taking any actions, and finally learners search for counter examples to confirm the feasibility of the tentative solutions before implementing them to solve the problem(s). If a gaming strategy is going to be used to engineer a MCLE, the designer needs to consider various aspects, based on the model-based learning process, such as what is the nature of the model that is driving the game; what is the dynamic feature of the game in support of the model-building process; does the game offer enough opportunities for learners to develop their own mental model during the process; and does the game provide sufficient solution examples for learners to examine their tentative solution model(s)?

The most critical design concern when using games to create a MCLE is how designers can efficiently maintain the integrity of the intended model models and the attainment of learning goals within the game-based learning environments. In other words, game designers should make the intended mental models explicitly observable for learners in the earliest stage of the learning/playing process while cautiously implementing appropriate game characteristics without distracting learners from the core of the intended mental model.

Workforce Implications of Using Modeled-Centered Learning (MCL) Games

While there are many different challenges that need to be addressed in order to strengthen the workforce, there are two trends that have the potential to impact on workforce development and therefore strengthening economic stability. MCL games are a viable solution to address 1) the need to replenish the loss of technical expertise and to 2) promote skills development to cope with the changing workforce requirements and demands. So what are the practical implications for using games as a means for complex skills development for today's workforce? MCL Games have the potential to address the following two trends found in the workforce.

Trend 1: Continual loss of Technical Expertise in the Workforce

The development of a novice to that of an expert has taken place in work settings for thousands of years. As organizations consider developing or revising their policies on workforce learning, it is important to consider the byproducts of providing training that requires "doing" a job task. Knowledge and skills that are developed while "doing" or performing a job are known as tacit knowledge and skills. Because tacit knowledge and skill can only be developed over time from experience, these skills and knowledge can play a critical role in economic stability.

This type of workplace learning is invaluable. While tacit knowledge in it self does not make someone an expert, experts have a tremendous amount of tacit knowledge. This makes sense when considering the process of becoming and expert. Expertise development involves a phenomenal amount of skills development. Some researchers believe that it can take more than 10,000 hours of practice to become an expert with an international reputation (Ericsson, 2004). Realistically given this amount of required practice, perhaps 60% of ones expertise development occurs while they are in the workforce. This development requires a tremendous amount of individual motivation. Nonetheless, the opportunities for individual learning are critical in supporting an individual's progress towards becoming an expert.

Most workers however plateau after 2 or 3 years of practice in a given field. Their knowledge and skills development do not improve by the mere act of "professional practice." In looking at this system of professional development with all of its parts, organizations need to carefully rethink their strategic planning and practice of lifelong learning as they continue to support lifelong learning and human development in the workplace.

A century ago, workforce issues involved manual labor. Today they involve international competition in technology and science. Public, private, business, governmental, and military organizations are grappling with the loss of expertise (e.g., due to retirement, relocation), and they are reaching for some system of organizational knowledge management. This is particularly critical given trends in worker mobility and the aging workforce, both of which are associated with the loss of expertise. This sets a truly noble goal for science: the elicitation and understanding of expertise along with the preservation and sharing of expert knowledge. This is an endeavor of potential benefit to all sectors of society.

Games can provide a solution for mitigating expertise loss. Games can engage workers at an earlier age whereby they will be better prepared to more quickly learn key skills when they enter into the workforce. While we do not necessarily advocate the use of games to specifically develop expertise in the workplace, games can provide a means to pre-workforce development of the complex skills needed for strong preparation prior to entry into the workplace. Entering into the workforce with strong cognitive skills will have a strong impact on workplace performance. While not able to directly gain experience prior to actual workplace entry, games can engage learners in the appropriate activities to help them gain the skills to quickly acquire a system model as well as successfully operate within the system.

Trend 2: Increasing Task Complexity as well as Demands and Requirements on Workforce

The outstanding growth rate of knowledge and information is perhaps more rapid than ever before seen in the history of the world. The nature of work tasks is such that many tasks are too complex and too large to be handled by a single individual (Cooke, Salas, Kiekel, & Bell, 2004; Stout, Cannon-Bowers, Salas, & Milanovich, 1999). Generally, teams are formed because collective resources (e.g., knowledge, skills, and diverse expertise) are required when working on such complex tasks. Individuals need to have advanced skills to successfully complete individual as well as complex team tasks.

Team that think more similarly (the degree to which members of a team share similar conceptualizations of problems and approaches to solutions) are more effective at completing complex tasks (Salas & Cannon-Bowers, 2000).

In situations where teams are important, a MCLE can facilitate individual mental model development that supports developing of shared thinking in teams. Similar mental models can be activated in a team setting to enhance overall performance. Developing instruction and learning activities that support a teams development of shared understanding has the potential to impact an organization's ability to compete and ultimately be successful amiss complex and challenging tasks.

An overarching goal for most MCL games that involve teams is to promote team cognition with the intended outcome of enhancing team performance and outcomes. Within the context of games, the development of shared understanding through effective coordination and communication appears critical to team performance (Eccles & Tenenbaum, 2003). While there are various strategies that are appropriate to improve team process, instructional strategies that use teams for learning focus on overall student benefits such as content understanding, application of knowledge to problem-solving and decision-making, team skills development, and valuing a team approach to solving complex intellectual tasks (Michaelsen, 2004). These are some of the very key objectives that are aligned with MCL games.

Through the use of teams, students and workers have opportunities to develop their skills of knowing how to learn, communicate, problem-solve, negotiate, think creatively, and work as a member of a team. These skills collectively have been recognized as critical elements for the success of business organizations (Carnevale, Gainer, & Meltzer, 1989; Woods, 2000).

Through the use of MCL Games, not only does this strategy deal with the challenge of technological progress, but it can also supports the development of team skills so that workers have the abilities to handle complex tasks.

Conclusion

Games can serve as an example of an instructional intervention that facilitates learners to quickly acquire the cognitive skills required to deal with the tasks that present vast amounts of knowledge and information. Games can also prepare learners to ask meaningful questions to thus help develop the skills associated with individual tasks as well as complex team tasks. Games not only provide a system model that can be complex and extensive in knowledge and information, but specific games can require multiple users to work together to solve problems and meet specific game goals.

Games can be used as a strategy for engineering model-centered learning environments. Even beyond the direct benefits to the learner, games as MCLEs can be used to meet greater societal challenges. Games as MCLE can be used to develop lifelong learning skills in pre-workforce populations. As such MCLEs using games can play a key role in strengthening the workforce. Games can be used to model complex domains as well as initiate long-term practice in these complex domains. While MCL games will mostly not be used over long periods of time, the practice skills learned from them can support the sustained deliberate practice that is needed to develop expertise (Ericsson, 2001a, 2001b; Ericsson & Charness, 1994).

Repeated used of games as MCLE can provide not only an entertainment experience, but can provide benefits such as mental model development, sustained motivation, simulated skill practice, experience acquisition, risk-taking, strategic planning, and complex skill development. As we design MCLE based on game design, we ultimately hope to promote skill development in the workforce to meet the tremendous demands that are place on workers everywhere.

Educational settings can have a direct impact on the workforce. As educational goals focus towards helping students develop learning strategies and cognitive tools, students will be able to acquire knowledge in order to think productively. Students need to acquire a basic understanding about key subjects but also they need to learn how to ask meaningful questions in order to better understand the principles of learning so that they can become themselves, lifelong learners.

As we consider workplace trends, MCLE using gaming strategies offer a unique benefit for complex skills development and ultimately workplace performance. As workplace organizations consider the challenges placed on workers and MCL games, they will be in a better position to make strategic decisions regarding the policies and strategies associated with lifelong learning in the workplace. As the practice of lifelong learning becomes more effective in development of the workforce, the potential for economic stability and strength increases.

Using games as a design for MCLE offer more than just a model for teaching facts. MCL games helps to create lifelong learners, learners that will eventually know how to learn. It has been said that the highest purpose of education is not to just teach facts however important they may be, but to train the mind, to make good citizens, and to develop character. MCL games can provide the critical impetus towards one journey of complex skills development, a journey that is not accomplished in a few years but continues beyond academic settings and into the workplace.

References

Ang, C. S., Avni, E., & Zaphiris, P. (2007). Linking Pedagogical Theory of Computer Games to Their Usability. *International Journal on E-Learning*.

Asgari, M. (2005, June 16-20). *A Three-factor Model of Motivation and Game Design.* Paper presented at the Digital Games Research Conference (DIGRA), Vancouver, British Columbia, Canada.

Avedon, E. M., & Sutton-Smith, B. (1971). *The Study of Games*. New York: John Wiley & Son.

Baranauskas, M. C. C., Neto, N. G. G., & Borges, M. A. F. (1999). *Learning at Work Through a Multi-User Synchronous Simulation Game.* Paper presented at the PEG'99 Conference, University of Exeter, Exeter, UK.

Belanich, J., Orvis, K. L., & Sibley, D. (2003). *Maximizing Training Effectiveness Using PC-Based Games.* Paper presented at the World Conference on E-Learning in Corporations, Government, Health, & Higher Education.

Bennett, J., & Warnock, M. (2007). Instructional game characteristics. Retrieved January 5, 2007, from http://iit.bloomu.edu/Spring2006_eBook_files/index.htm

Björk, S., & Holopainen, J. (2003). *Describing Games: An Interaction-Centric Structural Framework.* Paper presented at the Digital Games Research Conference (DIGRA).

Bonk, C. J., & Dennen, V. P. (2005). *Massive Multiplayer Online Gaming: A Research Framework for Military Training and Education*: Office of the under secretary of defense for personnel and readiness.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington D.C.: National Academy of Science, Commission on Behavioral and Social Sciences and Education.

Cannon-Bowers, J. A., & Salas, E. (Eds.). (1998). *Making decisions under stress: Implications for individual and team training*. Washington, DC: APA.

Carlson, H. L. (1991). Learning Style and Program Design in Interactive Multimedia. *Educational Technology Research and Development, 39*(3), 41-48.

Carnevale, A. P., Gainer, L. J., & Meltzer, A. S. (1989). *Workplace basics: The skills employers' want*. Alexandria, VA: American Society for Training and Development.

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction, 8*, 293-332.

Cooke, N. J., Salas, E., Kiekel, P. A., & Bell, B. (2004). Advances in measuring team cognition. In E. Salas & S. M. Fiore (Eds.), *Team cognition: understanding the factors that drive process and performance* (pp. xi, 268). Washington, DC: American Psychological Association.

Craik, K. (1943). *The nature of explanation*. Cambridge, UK: Cambridge University Press.

Crawford, C. (1982). The Art of Computer Game Design. Retrieved January 5, 2007, from http://www.vancouver.wsu.edu/fac/peabody/game-book/Coverpage.html

Crookall, D., Oxford, R. L., & Saunders, D. (1987). Towards Reconceptualization of Simulation: From Representation to Reality. *Simulation/Games for Learning, 17*, 147-171.

Csikszentmihalyi, M. (1990). *Finding Flow: The psychology of Optical Experience*. New York: Harper Perennial.

de Felix, J. W., & Johnson, R. T. (1993). Learning From Video Games. *Computers in the Schools, 9*(2/3), 119-134.

Downes, S. (2004). Learning by doing: James Paul Gee at RIMA ICEF. Retrieved January 7, 2007, from http://www.downes.ca/cgibin/website/view.cgi?dbs=Article&key=1079385148

Dreyfus, H. L., & Dreyfus, S. E. (1986). Five steps from novice to expert. In H. L. Dreyfus (Ed.), *Mind over machine: The power of human intuition and expertise in the era of the computer* (pp. 16-51). New York: Free Press.

Eccles, D. W., & Tenenbaum, G. (2003). Why an expert team is more than a team of experts: A social-cognitive conceptualization of team coordination and communication in sport. *Journal of Sports and Exercise Psychology*, 9-10.

Egenfeldt-Nielsen, S. (2006). Overview of research on the educational use of video games. *Digital Kompetanse, 1*, 184-213.

Ericsson, K. A. (2001a). Attaining excellence through deliberate practice. In M. Ferrari (Ed.), *The pursuit of excellence in education* (pp. 21-55). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ericsson, K. A. (2001b). Superior Decision Making as an Integral Quality of Expert Performance: Insights into the Mediating Mechanisms and their Acquisition through Deliberate Practice. In B. Brehmer, R. Lipshitz & H. Montgomery (Eds.), *How professionals make decisions*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine, 10*, S1-S12.

Ericsson, K. A., & Charness, N. (1994). Expert Performance Its Structure and Acquisition. *American Psychologist, 49*(8), 725-747.

Ertmer, P., Johnson, T., & Lane, M. (2001). *Vision Quest: Teacher Development Model for Scaffolding Technology Integration*. Paper presented at the 12th International Conference of the Society for Information Technology and Teacher Education.

Freeman, P., & Aspray, W. (1999). *The Supply of Information Technology Workers in the United States*. Washington D.C.: Computing Research Association.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation and Gaming, 33*(4), 441-467.

Gredler, M. E. (1996). Educational Games and Simulations: a Technology in Search of a (Research) Paradigm. In D. H. Jonassen (Ed.), *Handbook of Research on Educational Communications and Technology* (pp. 521-540). New York: Macmillan.

Gros, B. (2006). *Digital Games in Education: The Design of Games-Based Learning Environments*. Paper presented at the IADIS International Conference on Cognition and Exploratory Learning in Digital Age.

Guzzo, R. A., & Salas, E. (1995). *Team effectiveness and decision making in organizations* (1st ed.). San Francisco: Jossey-Bass.

Hackman, R. A. (Ed.). (1990). *Groups that work (and those that don't): Creating conditions for effective team work*. San Francisco: Jossey-Bass.

Hays, R. T. (2005). *The Effectiveness of Instructional Games: a Literature Review and Discussion*. Orlando, Florida: Naval Air Warfare Center Training Systems Division.

Hirumi, A. C. (2006). Serious Games: In Search of Quality. Retrieved Dec. 11, 2006, from http://www.jointadlcolab.org/newsandevents/ifests/2006/presentations/Dr_Atsusi_2C_Hirumi.ppt

Johnson, T. E. (1999). *Using A Computer Simulation As A Cognitive Tool: A Case Study Of The Use And Cognitive Effects Of Identibacter Interactus For The Development Of Microbial Identification Strategies By College Biology Students.* Unpublished Doctoral Dissertation, Purdue University, West Lafayette, IN.

Johnson, T. E., Spector, J. M., Huang, W. D., & Novak, E. (2007). *Instructional Gaming Effects on Learning Outcomes and Instructional Strategy Selection* (Naval Air Warfare Center, Training Systems Division). Tallahassee, Florida: Learning Systems Institute.

Johnson-Laird, P. N. (1989). Mental Models. In M. I. Posner (Ed.), *Foundation in cognitive science* (pp. 469-499). Cambridge, MA: MIT Press.

Kagan, D. M. (1992). Implications of research on teacher belief. *Educational Psychologist, 27*(1), 65-90.

Karoly, L. A., & Panis, C. W. A. (2004). *The 21st Century at Work: Forces Shaping the Future Workforce and Workplace in the United States*. Santa Monica, CA: RAND Corporation.

Kasvi, J. J. J. (2000). Not Just Fun and Games - Internet Games as a Training Medium. In P. Kymäläinen & L. Seppänen (Eds.), *Cosiga - Learning With Computerised Simulation Games* (pp. 23-34): HUT: Espoo.

Kirriemuir, J., & McFarlane, A. (2006). *Literature Review in Games and Learning*: Futurelab.

Klabbers, J. H. G. (2006). *The Magic Circle: Principles of Gaming & Simulations*. Rotterdam: Sense Publisher.

Leemkuil, H., de Jong, T., & Ootes, S. (2000). *Review of Educational Use of Games and Simulations*: EC project KITS.

Malone, T. W. (1980). *What Makes Things Fun to Learn? A Study of Intrinsically Motivating Computer Games*. Palo Alto, CA: Xerox Palo Alto Research Center.

Malone, T. W. (1981). Toward a Theory of Intrinsically Motivating Instruction. *Cognitive Science, 4*, 333-369.

Malone, T. W., & Lepper, M. R. (1987). Making Learning Fun: A taxonomy of Intrinsic Motivations for Learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning, and instruction* (Vol. 3, pp. 223-253). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mayer, R. E. (2001). *Multimedia Learning*. Cambridge: Cambridge University Press.

McGrenery, J. (1996). *Design: Educational Electronic Multi-Player Games – A Literature Review*: University of the British Columbia.

Michaelsen, L. K., Knight, A. B., & Fink, L. D. (2004). *Team-based learning: A transformative use of small groups in college teaching*. Sterling, VA: Stylus Publishing, LLC.

Pivec, M., & Dziabenko, O. (2004). Game-based Learning in Universities and Lifelong Learning. *Journal of Universal Science, 10*(1), 14-26.

Reigeluth, C. M. (1999). *Instructional-design theories and models*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Reigeluth, C. M., & Frick, T. W. (1999). Formative research: A methodology for creating and improving design theories. In C. M. Reigeluth (Ed.), *Instructional-design theories and models volume II: A new paradigm of instructional theory (Vol. 2)*. Mahwah, NJ: Lawrence Erlbaum Associates.

Reigeluth, C. M., & Schwartz, E. (1989). An instructional theory for the design of computer-based simulations. *Journal of Computer-Based Instruction, 16*(1), 1-10.

Rieber, L. P. (2001). *Designing Learning Environments That Excite Serious Play.* Paper presented at the Annual Conference of the Australian Society for Computers in Learning and Tertiary Education (ASCILITE), Melbourne, Australia.

Salas, E., & Cannon-Bowers, J. A. (2000). The anatomy of team training. In L. Tobias & D. Fletcher (Eds.), *Handbook on research in training*. New York: Macmillan.

Seel, N. M. (1991). *Weltwissen und mentale Modelle*. Hogrefe: Gottingen.

Seel, N. M. (1999). Educational diagnosis of mental models: Assessment problems and technology-based solutions. *Journal of Structural Learning and Intelligent Systems, 14*(2), 153-185.

Seel, N. M. (2003a). Designing Model-Centered Learning Environments: Hocus-pocus – or the focus must strictly be on locus. In J. M. Spector, C. Ohrazda, A. van Schaack & D. Wiley (Eds.), *Innovations in Instructional Technology: Essays in Honor of M. David Merrill* (pp. 65-90). Mahwah, NJ: Lawrence Erlbaum Associates.

Seel, N. M. (2003b). Model-Centered Learning and Instruction. *Technology, Instruction, Cognition and Learning, 1*(1), 59-85.

Seel, N. M. (2003c). *Psychologie des Lernens (2. Aufl.)*. Munchen: Reinhardt.

Seel, N. M., Al-Diban, S., & Blumschein, P. (2000). Mental Models and Instructional Planning. In J. M. Spector & T. M. Anderson (Eds.), *Integrated and holistic perspectives on learning, instruction and technology: Understanding complexity* (pp. 129-158). Dordrecht: Kluwer Academic Publishers.

Seel, N. M., & Schenk, K. (2003). An evaluation report of multimedia environments as cognitive learning tools. *Evaluation and Program Planning, 26*(2), 215-224.

Simon, H. A. (1996). *Observations on the Sciences of Science Learning*: Department of Psychology, Carnegie Mellon University.

Squire, K. (2005). *Game-Based Learning: Present and Future State of the Field.* Paper presented at the Masie Center e-Learning Consortium.

Stolurow, L. M. (1973). Lernumwelten oder Gelegenheiten zum Nachdenken. In W. Edelstein & D. Hopf (Eds.), *Bedingungen des Bildungsprozesses. Psychologische und pädagogische Forschungen zum Lehren und Lernen in der Schule* (pp. 351-398). Stuttgart: Klett.

Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors, 41*(1), 61-71.

Thomas, R., & Hooper, E. (1991). Simulations: An Opportunity We Are Missing. *Journal of Research in Computing in Education, 23*(4), 497-513.

Vockell, E. (2004). Educational Psychology: A Practical Approach. Retrieved January 5, from http://education.calumet.purdue.edu/Vockell/EdPsyBook

Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: MIT Press.

Waal, B. D. (1995). *Motivation for Video Game Play: A Study of Social Cultural and Physiological Factors*: Simon Fraser University.

Woods, D. R., Felder, R. M., Rugarcia, A., & Stice, J. E. (2000). The future of engineering education: Development of critical skills. *Chemical Engineering Education, 32*(2), 108-117.

Index

A
ADDIE, 187, 189, 190, 196–199, 207, 247, 249, 260, 267, 269, 270
Analogy, 88, 97, 98, 253, 262, 265, 295
Anchored instruction, 19, 107, 112, 121, 188, 262, 265, 277–281, 284, 286, 287, 294, 298, 299
Application, 48, 50, 65, 76, 90, 100, 107–112, 115–121, 124–126, 130, 133–137, 140, 152, 187, 188, 189, 191–193, 211, 212, 218–222, 247, 250, 252, 258, 271, 279, 295, 309, 311, 313
Aristotle, 4, 5, 12
Assessment, 23, 24–29, 31, 35–38, 44–48, 52–57, 84, 111, 114, 133, 137, 145, 147, 149, 150, 156, 191, 198, 221, 277–280, 284, 286–288, 290, 292–295, 317
  centered, 280, 286, 290
Assumption, logical, 105
Axiom, 106, 123, 128, 136, 137

B
Belief pattern, 28, 31–33, 35, 36
BIRD, 263
Book of Problems, 6, 7

C
Case-based reasoning, 232, 236
Cognition, 7, 10, 14, 18, 23, 25, 45, 47, 50, 51, 63, 65–73, 75, 76, 82, 108–110, 140, 150, 172, 189, 190, 192, 193, 195, 196, 198, 207, 247, 251, 255, 278, 299, 310, 317, 320
Cognition and Technology Group at Vanderbilt (CTGV), 248, 277–281, 283–285, 288, 298, 299
Cognitive
  overload, 67, 68, 201
  tools, 145, 321
Cognitive apprenticeship, 82, 88, 89, 98, 128, 180, 181, 260, 262, 263
Community-centered, 280
Companion, 163, 262, 264, 265
Complex skills development, 305–307, 309, 318, 321, 322
Concept, 148, 150, 151, 156
  maps, 24, 25, 29–32, 47, 48, 57, 148, 150, 151, 256, 265
Conceptual change, 63–66, 68, 70–72, 74–76, 100, 150, 179
Construction Integration Model (CI), 66, 67, 70
Constructivism, 14, 15, 17–20, 123, 192, 251

D
Data, longitudinal, 44, 55–57
Design
  enclosed, 117, 120
  evidence-centered, 24–26
  experiments, 105, 113–115, 125, 133, 134, 136
  languages, 169
Diagnosis, 30, 36, 43, 44, 48, 50–57, 82–84, 87, 100, 111, 155, 227–229, 231

E
ELM-ART, 225, 226, 229, 234, 236, 237, 239, 240, 244
Empedocles, 4, 5
Epistemology, 11, 14, 19, 20, 52, 190, 192, 207, 270
Experience, 4, 10, 11, 13–15, 17–19, 24, 64, 68, 72, 74, 75, 108, 156, 161–164, 166, 167, 213, 235, 244, 251, 254, 256, 257, 265, 306–317, 319, 321
Experiment, 44, 53, 105, 108, 110–115, 118, 119, 121–127, 129, 130, 132–136, 139, 140
  extended design, 105, 110, 125, 133
Expert system, 155
Externalization, 45, 52, 53, 155, 183, 288
Eye tracking, 63, 64, 68–71, 73, 75

G
Game, 189, 190, 192, 201–207, 307–318, 321
  design, 207, 313, 315, 317, 321

H
Higher education, 203, 277, 279, 294, 297, 309
HLM, 57
How people learn, 250, 279, 295, 297
Hypermedia, adaptive, 150, 234–236

I
Image, 11, 12, 14, 147, 149, 152, 153, 214, 215, 221
Information, 7, 8, 30, 31, 35, 38, 52–55, 58, 64, 65, 67, 68, 70–72, 76, 100, 147, 148, 155, 175–182, 184, 189, 190, 192, 194–196, 198–208, 213, 218, 220, 232, 234, 235, 237–240, 243, 252, 253, 260, 286, 287, 290, 293, 296, 306, 308, 320, 321
Inquiry, 3, 4, 6–8, 13, 20
Instruction, 105–109, 111, 112, 114, 131, 133–137, 139, 140, 161–170, 175, 176, 179, 180, 183, 185
Instructional
  design, 3–6, 8, 19, 20, 107–109, 116, 134, 135, 137, 161, 164, 165, 171, 189–191, 193, 195–200, 207, 212, 214, 217, 219, 247–250, 259, 260, 266–271, 277–281, 287, 290, 294, 297, 298
  gaming design, 207
  practice, 211, 212, 222, 305, 309, 316
  science, 4, 5, 11
  system, 3–8, 191, 267
  systems research, 6, 7
  technology, 190
Instructional design model, 171, 196
Instructional design,-model-centered, 247, 249, 266–271
Instructional design principles, 277–279, 287, 297
Instructional systems design (ISD), 247, 248, 266–269, 317
Instrument, 48, 54, 57, 82, 83, 88, 97, 110, 111, 113, 121, 131–134, 213, 259
Intelligent tutoring system, 169, 225–227, 234, 236, 243
Intervention, 8, 25, 37, 38, 64, 75, 88, 126, 130, 133, 139, 176, 178, 179–185, 203–205, 255, 259, 262, 296, 309, 321
Isomorphism, structure, 90, 99, 255

J
Jasper Woodbury, 277, 278, 281

K
KM Quest, 203–205
Knowledge, 23–26, 28–31, 35–38, 175–180, 182, 184, 189, 190, 192, 194, 196–208
  acquisition, 82, 86, 87, 190, 192, 258
  episodic, 145, 155, 156
  prior, 177–179, 182, 184
  procedural, 145, 149, 153–156
  representation, 31, 47, 81, 83, 87, 148, 162
  structural, 145, 148, 149, 151, 156
  visual, 152, 153, 156

L
Language, 10, 15–20, 47, 57, 83, 87, 138, 169, 183, 193, 219
Layer
  content, 161, 166
  data management, 170, 171
  design, 165
  media-logic, 165, 170, 171
  message, 168, 169
  representation, 165, 169, 170
  strategy, 165–167, 170
Learner-centered, 280, 284
Learning, 105–111, 113–115, 117, 119–122, 126, 127, 129–137, 139, 140, 277–299
  accommodation, 253, 254
  adult, 294
  assimilation, 252, 253
  concept, 193, 194
  environments, 199–203, 205, 206
  model-centered, 249, 251–259, 263, 268, 270, 271, 305, 309, 321
  model-based, 176–178
  problem-based, 248, 249, 259, 263, 264, 279, 285, 290
  social, 254

Learning environments, 27, 43, 52–54, 57, 107, 115, 119, 126, 200, 201, 206, 248, 262, 265, 270, 278–281, 284, 288
adaptive, 225, 226, 234–244
model-centered, 305, 309, 321
Learning goals, 107, 115, 218, 227, 236, 239, 243, 287, 290, 296, 316, 318
Learning instances, 105, 137, 139
LISP, 227–230, 233, 236, 240

M
MCLE, 305, 307, 309, 311–313, 316–318, 320, 321
Measurement of change, 44, 48–51, 56, 57
Memory organization paket (MOP), 256–258
Metaphor
barrier, 137
effect, 110, 113, 136, 140
engine, 113, 136, 137
rift, 137
Methodology, 26, 45–47, 50, 54, 82–84, 86, 87, 90, 105, 107, 110, 121, 122, 130, 131, 133
Mindtool, 145, 150, 156
MITOCAR, 47, 268
Model
causal, 45, 46, 84–89, 91, 95, 97–99
episodic learner, 227–229, 231–234, 236, 237, 244
explanation, 53, 56, 179
mental, 7, 10, 23–25, 29–31, 38, 39, 44–47, 49–58, 82–91, 93, 97–100, 145–151, 175–185, 193–195, 199, 200, 207, 212–214, 216, 218, 222, 247, 252–266, 268, 269, 271, 310, 318, 320, 321
naïve, 65, 66
overlay, 234, 236–238, 244
perceptual, 254, 255
thought, 254, 255
Modeling, 150, 155, 156, 199, 200
cognitive, 112, 113
knowledge, 239, 240
user, 225, 227
MOMBI, 175, 176, 179–185

N
Network
Bayesian, 26, 31, 32, 34
flexible belief, 31–37

O
Operationalization, 82, 97, 111, 112, 116–119, 121, 126, 128, 129, 133, 135, 136

P
Perspective, multiple, 291–294, 296
Philosophy of science, 4, 105, 125
Planned stage by stage formation of mental actions, 211, 214, 215, 217–221
Progression of mental models, 38, 44, 53, 54, 57
Psychology
cognitive, 20
development, 211, 213, 220
learning, 20, 190, 191, 196

Q
Quasi-experiments, 45, 112, 113

R
Reading processing, 66–69, 72, 73, 75
Refutational text, 71, 73, 76

S
Scaffolding, 88, 95–97, 99, 166, 175, 181, 182, 184, 185, 262, 282–284, 317
Schema, 10, 51, 52, 54, 175, 178, 181, 185, 201, 228, 253–256, 265
Schematization, 178, 181, 182
Schooling, 211–213, 219–222
Script, 253, 257, 258
Semantic, 24, 30, 31, 38, 46–48, 54, 56, 66, 67, 89, 128
relation, 151
Simulation, 37, 49, 112, 163–165, 167, 169, 189, 192, 200–207, 306, 308, 310, 313
SMART Model, 277, 279, 284, 285
SMD Technology, 47, 48, 55, 56, 57, 268
STAR.Legacy, 277, 279, 287–289, 294–296, 298
Statistic, 44, 46–50, 52, 54, 56, 57, 126, 135
Structure Lay Down Techniques, 86, 87
Systems approach, 190, 191

T
Technological revolution, 247
Technology, 47, 48, 55–57, 107, 109, 114, 134, 189–192, 195, 200, 207, 277–282, 284, 285, 287, 288, 294, 295, 298, 299

Testing, 83, 86–88, 113, 116, 119, 121–127, 132–134
Text comprehension, 63, 64, 66–72, 75, 76
Thematic organization point (TOP), 256–258
Theorem, 106, 107, 117, 122, 123, 127, 128, 133
Theory, 105–109, 111–136, 138
 cognitive, 281, 284, 298
 design, 163
 enclosed, 117, 118, 125
 learning, 249, 251, 255, 267, 316, 317
 mental model, 105, 107, 112
 model-centered learning, 249, 251, 252
 psychological, 211, 220
 transfer, 282, 283, 296
Trends of instructional design, 248

V

Validity, 25, 45, 46, 48, 56, 82–84, 86, 88, 97–100, 111, 113, 119, 130, 136