# 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).

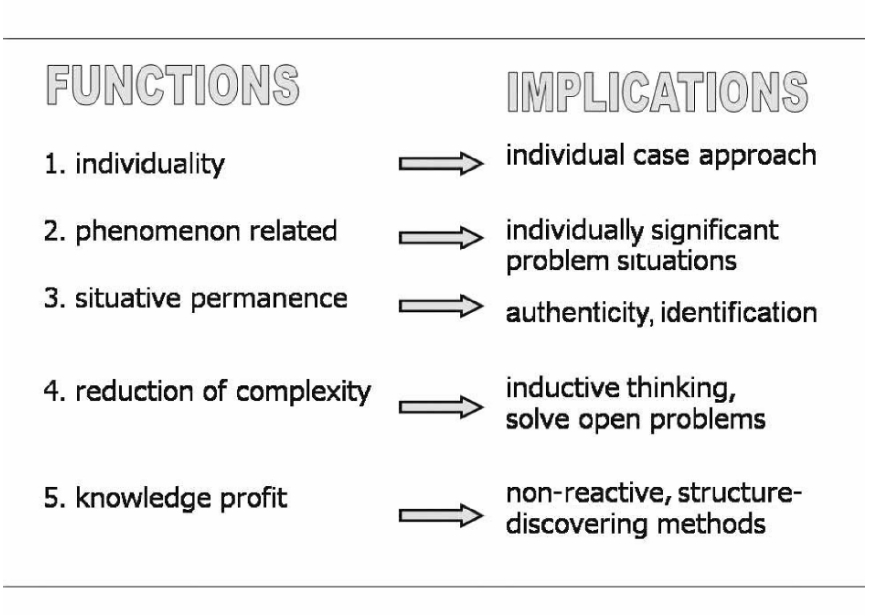


**Fig. 1.** Implications for the diagnostic of mental models.

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).

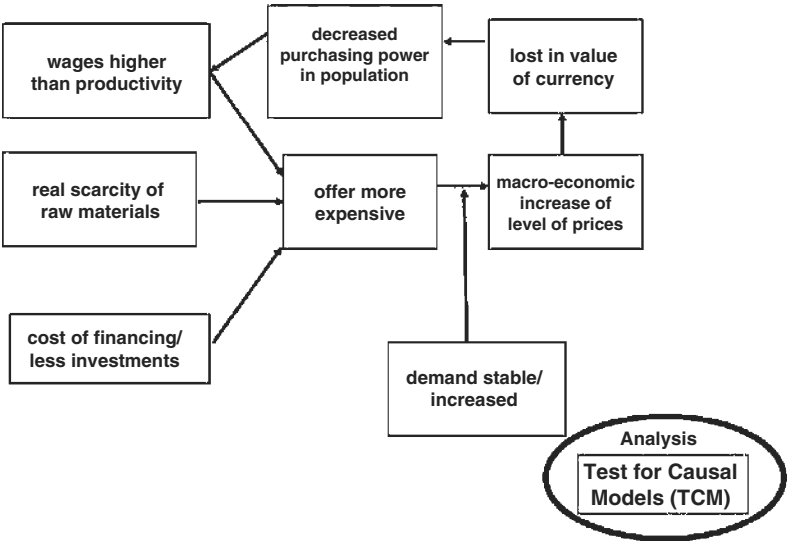


**Fig. 2.** Part of a conceptual expert model: offer caused genesis of inflation.

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

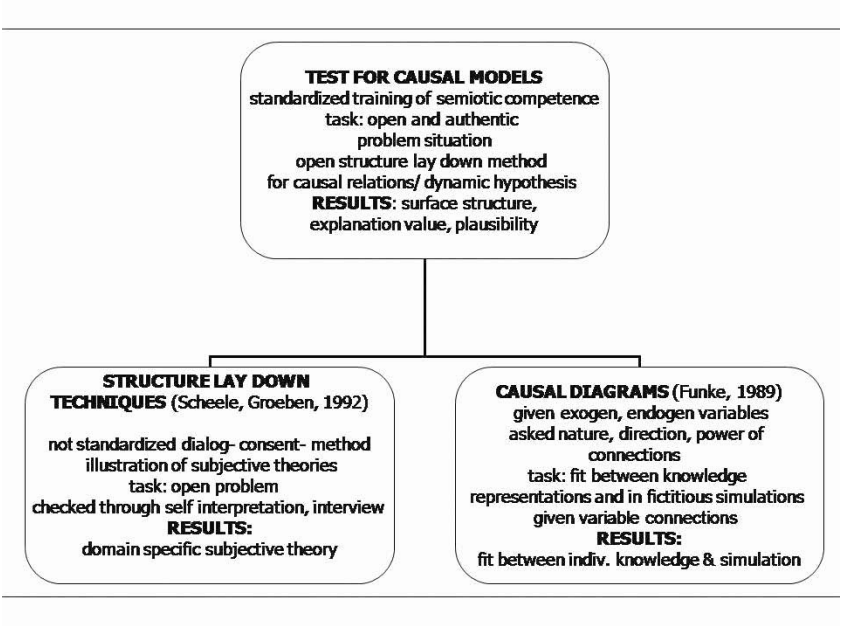


**Fig. 3.** Elements of the new developed Test for Causal Models.

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.

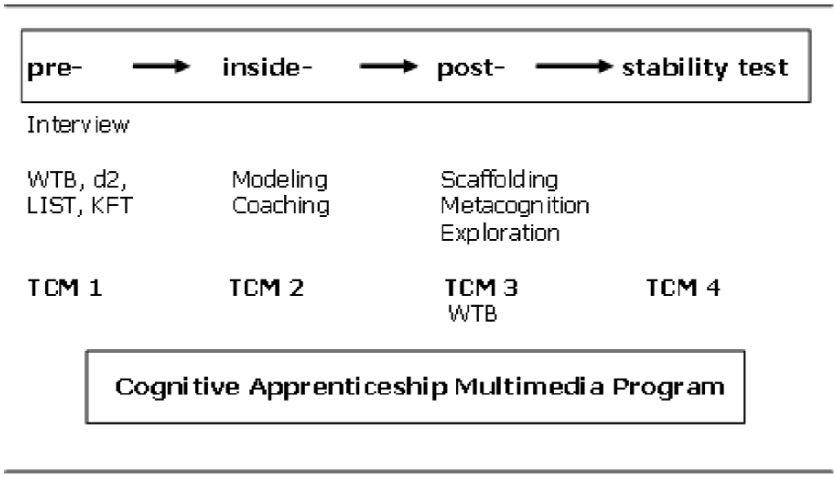


**Fig. 4.** Learning process accompanying study, N = 26 students, 11th class.

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.

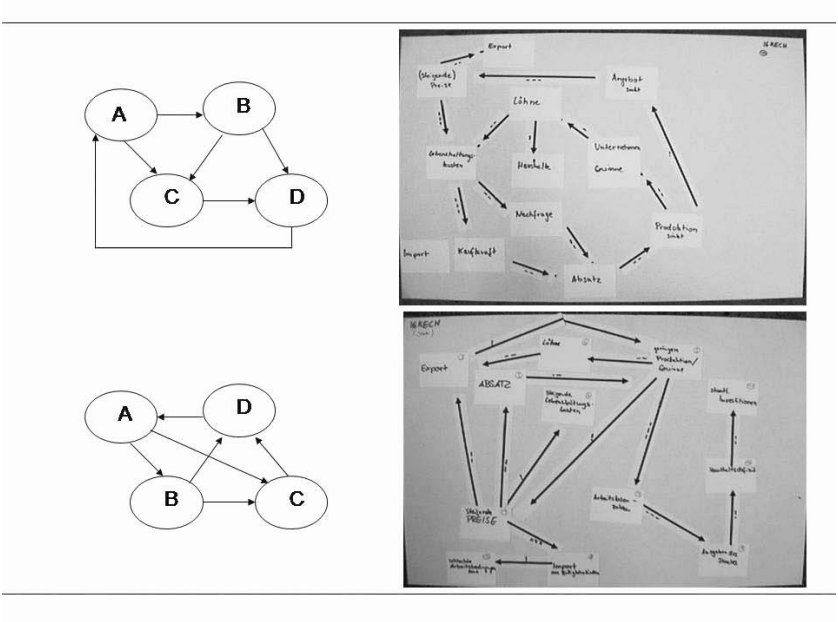


**Fig. 5.** Structure isomorphism – same or different structures? (The example on the right side is the same participant in pre- and stability test, row data).

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,

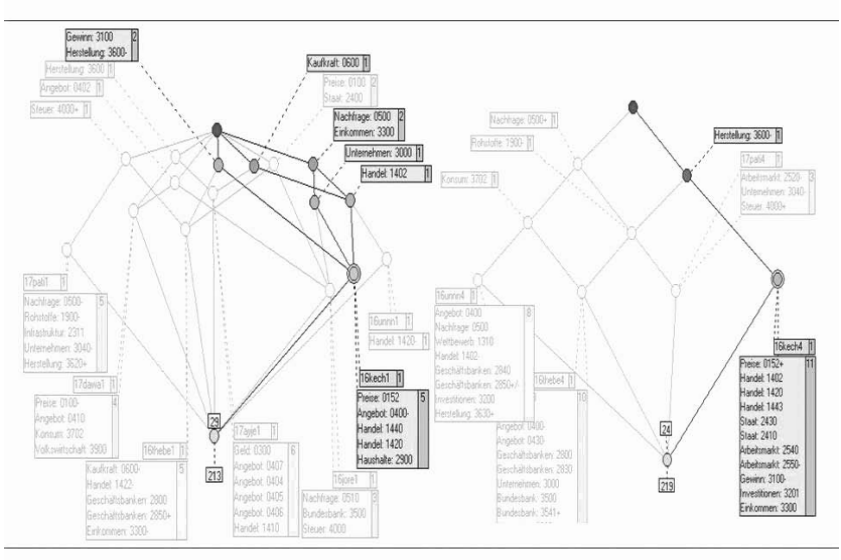


**Fig. 6.** Formal concept analysis – individual case approach on the level of terms by the same participant as Figure 5: 16kech (pre- and stability test).

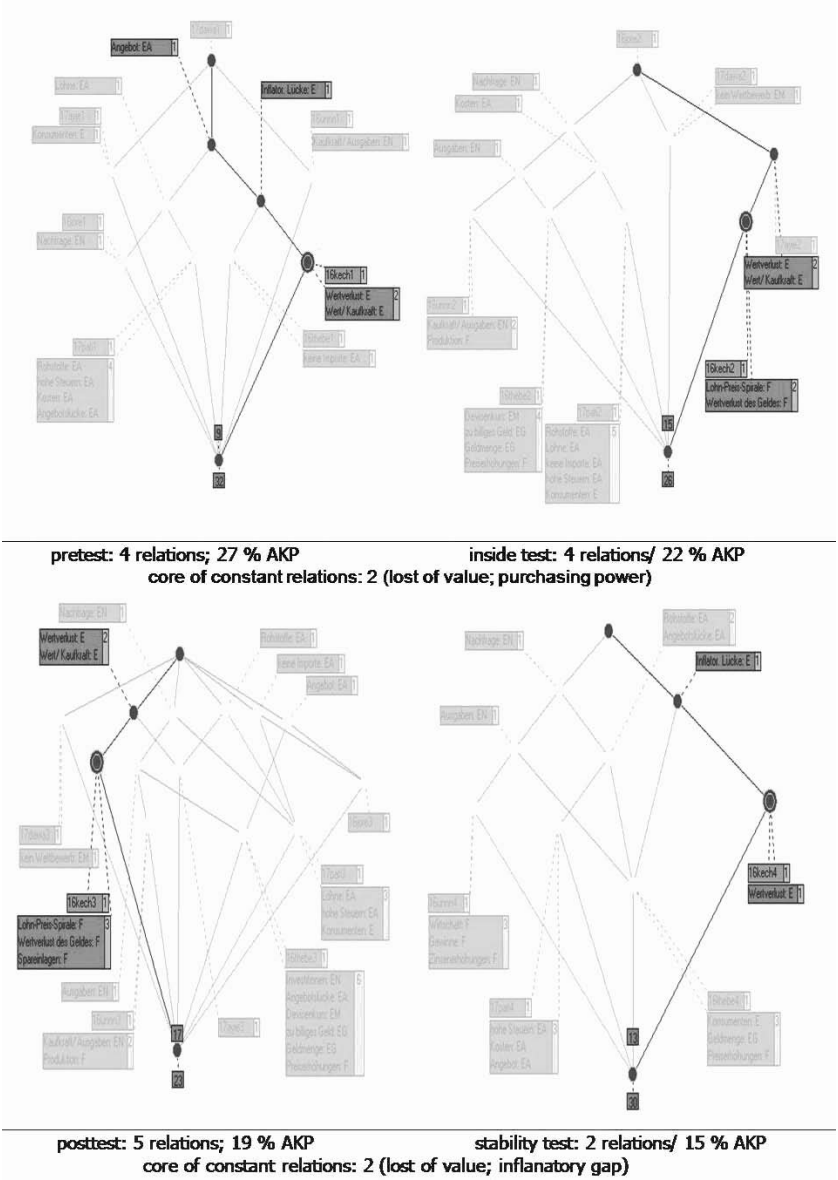


**Fig. 7.** Individual case approach – Level of causal relations by 16kech, AKP: proportion of conceptual expert model based explanatory value.

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.

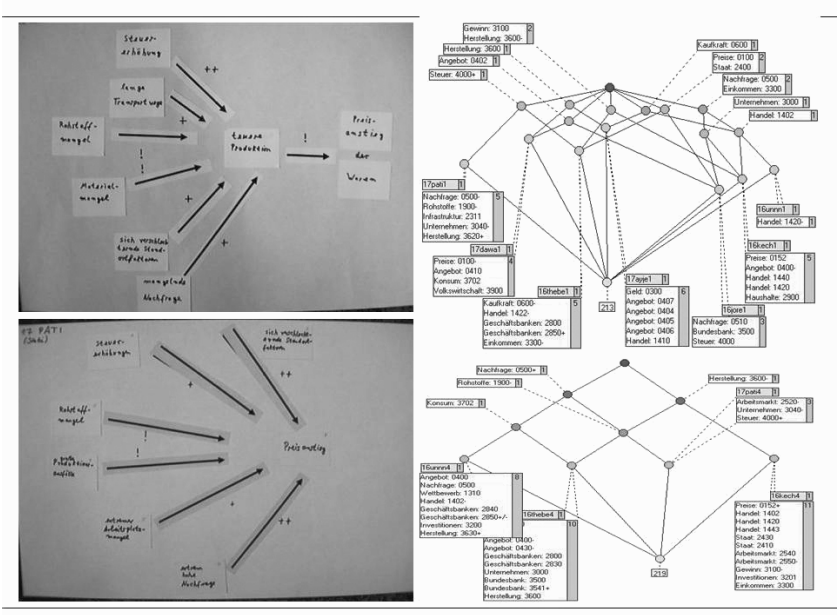

**Fig. 8.** Individual case approach 17pati – Level of row data and of terms (pre-, stability test).

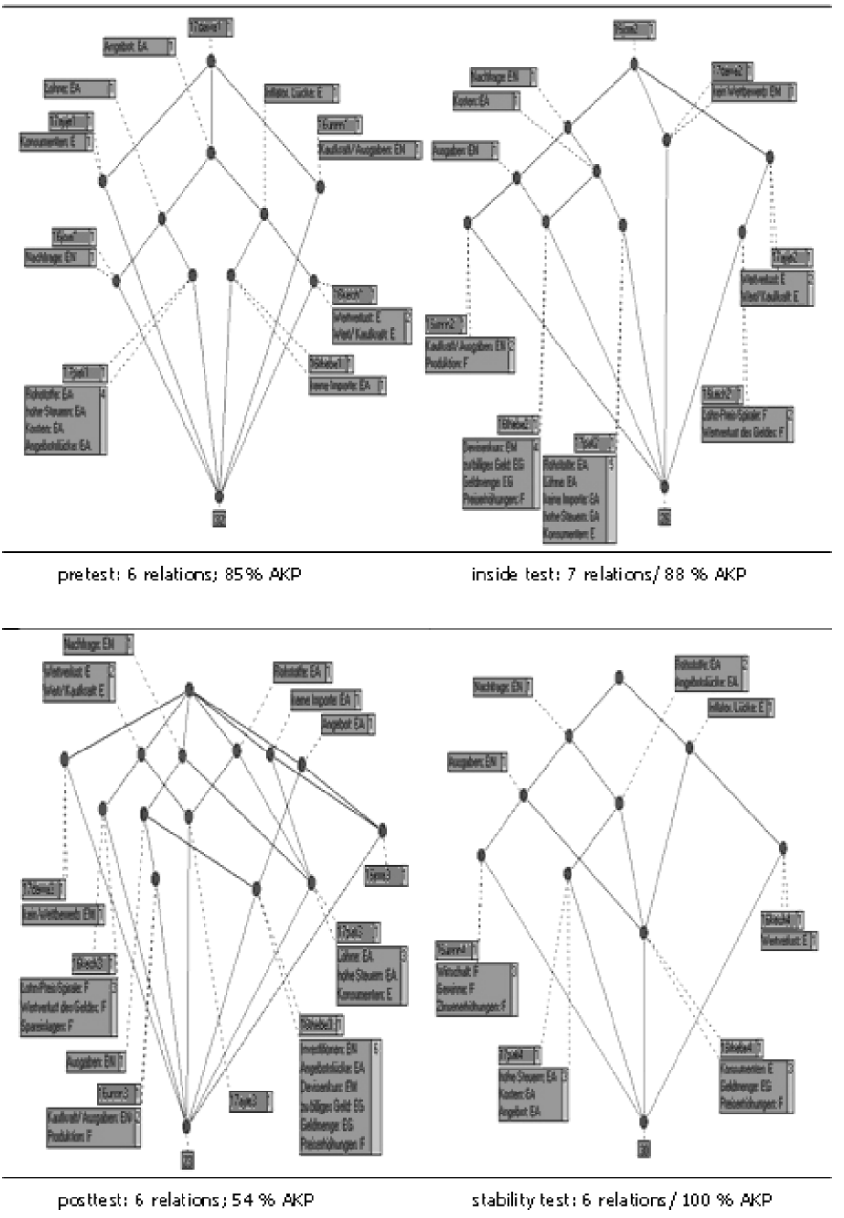


**Fig. 9.** Individual case approach 17pati – Level of row data and of terms (pre-, stability test).

## 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.

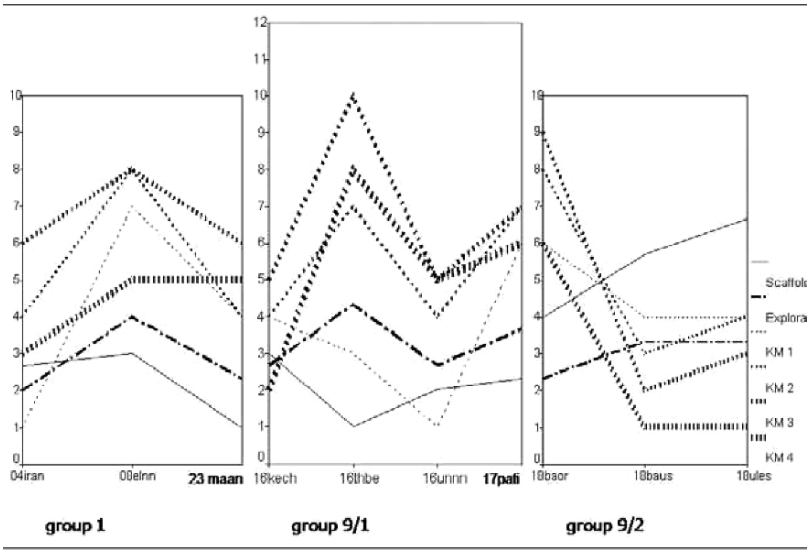


**Fig. 10.** Change of plausibility: Explanatory value absolute: Scaffolding, Exploration, Causal Models 4 times (KM), 10 single approaches.

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

| | Explanatory value | subjective content plausibility | | subjective formal plausibility | | decl. knowledge |
|---|---|---|---|---|---|---|
| | | 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* | 0.56* |
| 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

| | Explanatory value | subjective content plausibility | | subjective formal plausibility | | decl. knowledge |
|---|---|---|---|---|---|---|
| | | 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

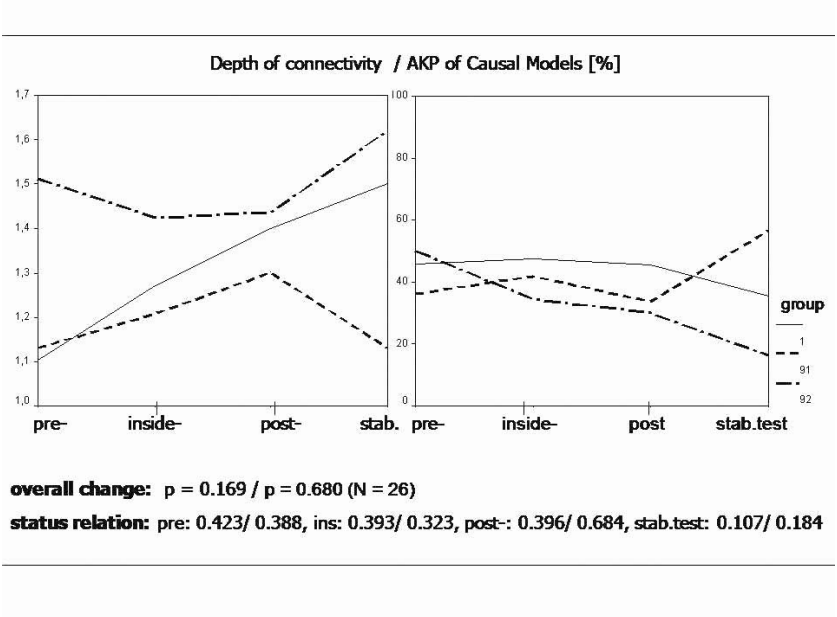


**Fig. 11.** Depth of Connectivity/AKP of Causal Models per group in 4 times.

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.

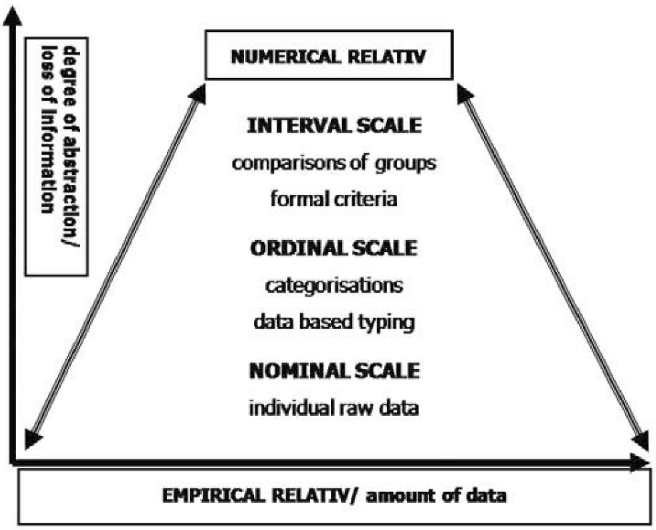


**Fig. 12.** Measuring- Validity- Dilemma by diagnosis of mental models.

[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.