Supporting Coherence Formation in Multimedia Learning

A Theoretical Framework and First Experimental Results

Abstract. Multimedia learning environments usually combine multiple forms of representations like texts, pictures, animations, tables or formulas. However, knowledge acquisition from multiple representations requires the integration of these sources in order to construct a coherent mental representation. The present paper therefore outlines a first theoretical framework for the process of coherence formation within cognitive load theory. Despite the benefits of dealing with multiple representations (e.g., multiple perspectives, elaboration, abstraction, flexibility) there is empirical evidence that learners and especially those with little prior knowledge have difficulties in integration and fail to construct a globally coherent mental representation. Different approaches to reduce the cognitive costs of coherence formation are discussed with reference to current empirical studies. Thus, the effectiveness of help for coherence formation turned out to be highly depending on learners expertise and as it imposes germane cognitive load it can only be helpful within the limits of working memory capacity.

Modern multimedia learning environments usually combine multiple forms of representations like written or spoken texts, static and animated pictures or graphs, tables or formulas (Mayer, 2001). However, knowledge acquisition from multiple representations requires different cognitive mapping processes. First, learners have to understand each of the given representations. In a second step they have to create referential connections between the different representations. Third, the acquired structures between the representations have to be linked to the subject matter which is represented. Only when learners identify all these references within and between the external representations and between the representations and the represented subject matter they can construct a coherent mental representation and come to a deeper understanding of the subject matter. The whole process of structure mapping - called coherence formation - is crucial for effectively learning with multimedia. Moreover, designing effective multimedia learning environments require instructional support of this coherence formation. The present paper therefore outlines a first theoretical framework for the process of coherence formation within cognitive load theory (Chandler & Sweller, 1991) as well as a brief review about empirical results on effective support of coherence formation.

1. A Theoretical Framework for the Process of Coherence Formation

From a cognitive point of view we can ask how learners operate while learning with multiple representations. Which cognitive processes are necessary for the different steps of coherence formation? Actually, research on learning with multiple
representations is not based on a specific theoretical framework for the process of coherence formation (Seufert, 2003a). However the crucial points in coherence formation can be seen as processes of structure mapping, which are elaborated in Gentner’s structure mapping theory (Gentner, 1983). This theory is originally formulated for learning with analogies, but can be adapted to the linking process between multiple representations (Seufert, 2003a; Bodemer, Plötzer, Feuerlein & Spada, in press):

The prerequisite for mapping between multiple representations (as well as for mapping between analogies) is to identify relevant elements in each of the presented representations. For example when a learner is given a text with a related picture he or she has to analyze the most important concepts in the text as well as in the picture. But understanding the text respectively the picture additionally requires to identify the relations between those concepts. The result of this analysis is a coherent mental representation of the text and of the picture. As this process is limited to a single representation it is called local coherence formation.

Most learning environments use representations that are partly redundant, i.e. they comprise corresponding elements and interrelations of these elements (as this is the case with analogies). Therefore learners have to map the identified concepts and local structures between the representations, in our example between text and picture. Putting the pieces together results in an integrated mental representation, that shows the global meaning of the given representations. As the focus of this process is more comprehensive we call it global coherence formation.

However, the process of structure mapping, independent from its more local or more global focus, can be carried out by the learner by two different strategies (Ainsworth, 2000; Seufert, 2003a):

- The first strategy is to rely on surface features of the given representations as a hint for correspondence, e.g. a red written word can be mapped onto a red labeled part of a picture. Corresponding surface features serve as indicators for mapping and may lead the learner to process these elements and structure more deeply. Nevertheless, it is also possible that a learner only superficially observes the representations while looking for corresponding surface features without semantic processing.

- The second strategy is to look for conceptual correspondence of elements and relations and to map them based on a semantic analysis on a deep structure level. E.g. a graphical representation of a chemical reaction is mapped onto a textual description of the same reaction because the learner identifies the chemical elements and processes - i.e. the conceptual meaning of picture and text – and realizes their semantic correspondence.

Only if learners really engage in such semantic mapping processes and analyze the deep structure level, learning with multiple representations will be effective. When learners can inter-relate the relevant concepts they are able to create a cognitively flexible knowledge structure. Multiple representations offer multiple ways to present a subject matter and they provide redundant as well as complementary information that can be used for building an elaborated knowledge structure. However, not only
elaboration but also abstraction in terms of subtraction is beneficial in learning with multiple representations: While analyzing the representations semantically, learners can grasp the underlying patterns and invariants and reflect them at a higher level of organization (Ainsworth, 1999). Therefore learners can achieve a coherent, comprehensive, flexible and more abstract knowledge structure. Given these convincing arguments we should recommend using multiple representations for every learning purpose. However, in educational research we also find arguments against using multiple representations (Ainsworth, 1999; Seufert, 2003a). E.g. the process of coherence formation is cognitively demanding and learners with insufficient prior knowledge are often unable to cope with this task (van Someren, Reimann, Boshuizen & deJong, 1998, Kozma & Russell, 1996, Seufert, 2003a, b; Yersushalmy, 1991). Consequently, they do not use different representations but rather concentrate only on one representation, often the more familiar or concrete one (Scanlon, 1998; Tabachnek & Simon, 1998). These learners only switch between representations in the case of problems in understanding the actually employed representation (Tabachnek & Simon, 1998). Therefore they do not map semantically and fail to reach the goals of elaboration, abstraction, flexibility and coherence.

The problem of the ineffective use of multiple representations can be seen as a problem of cognitive economy: provided that learning is a highly active and self-regulated process, learners can decide not to invest too much cognitive resources if the expected outcome of this effort is estimated as insufficient. Learners evaluate the possible benefit of processing additional representations and if the added value of information is not balanced with the anticipated effort for processing the integrating, the process is stopped. So, if we want learners to really benefit from the integration of multiple representations, we have to reduce the cognitive costs of processing multiple representations.

2. HELP FOR COHERENCE FORMATION

In the past few years several cost-reducing approaches have been examined. The goal of these approaches usually is to support students while integrating multiple representations, e.g. by focusing their attention or by explicitly showing them, which items can be mapped. Generally all these approaches can be assigned to two different levels of helping strategies as they are described above:

First, learners can be assisted to recognize corresponding elements and structures on a surface feature level. The best known help on surface feature level is color coding, i.e. to label corresponding elements in each representation with the same color. As the example points out, this kind of support is mostly graphical. To ensure, that learners map semantically they can be secondly supported on a deep structure level, i.e. the meaning of corresponding structures is explained and exposed more or less explicitly. This kind of help is mostly given verbally. There are numerous examples for both approaches that were empirically evaluated and discussed in educational research. Some empirical studies concerning the effectiveness of different implementations will be presented for both types of help.
2.1 Help for coherence formation on a surface feature level

In the past few years a lot of researchers developed and examined graphical help for coherence formation in learning with multiple representations.

- Color coding was successfully implemented in a study from Kalyuga, Chandler und Sweller (1998). Learners were given an electronic text with illustrations and when clicking on parts of the text they were highlighted in the same color as the corresponding parts of the illustration. Thus, learners’ attention was guided to the relevant parts of both representations and relieved cognitive capacity what in the end resulted in better learning outcomes.

- Another computer-based technique for guiding coherence formation is dynamic linking (Scanlon, 1998; Yerushalmy, 1991). Ploetzner, Bodemer und Feuerlein (2001) used dynamic linking in their computer-based statistic trainer VISUALSTAT. Learners could observe changes in a graph after they had changed the values in the corresponding formula. Scanlon (1998) also used this technique in a physical context and revealed that mapping between two dynamic representations is easier for learners than mapping between one dynamic representation (e.g. a growing graph) and a static representation (e.g. a formula or a table).

- A further type of help, inter-textual hyperlinks, was implemented in a study from our own workgroup (Brünken, Plass & Leutner, 2003a; Brünken, Seufert & Zander, subm.). Within the text there were some relevant concepts marked as a hyperlink, as we know it from Hypertexts. As learners clicked on these links they saw arrows pointing to corresponding parts in the picture. The specific guidance of attention also reduced cognitive load and thus enabled better learning outcomes for both, text and picture content.

Beside the positive effects of help on the surface feature level it is worthwhile thinking about the risk that learners process the representations only superficially. The guidance of attention may lead to a narrow focus on surface features without going deeper and analyze the representations semantically. The same phenomenon could be demonstrated in some empirical work about the effects of animations. (Lewalter, 1997; Schnotz, Böckeler & Grzondziel, 1999, Schnotz, 2002). The animation guides attention on important features of the picture but learners often only passively look at the dynamics without actively thinking and thus they fail to understand the subject matter (Schnotz, Seufert & Bannert, 2001).

Help for coherence formation therefore should guide learners’ attention to relevant parts of the representations, but however, this demands to stimulate processes of understanding the presented information on a deep structure level instead of a mere surface analysis.

2.2 Help for coherence formation on a deep structure level

Assisting learners on a deep structure or semantic level of processing can be realized in different ways. First, we can ask what aspect of the semantic analysis is aimed at,
Supporting Coherence Formation

2.2.1 What aspect of coherence formation should be assisted?
As mentioned above, coherence formation is a complex process with two major steps: local and global coherence formation. Consequently help for coherence formation can aim at these two phases. On the one hand side it can assist learners to understand the single representations as it is realized in several trainings for strategic text-processing or by guiding learners’ attention in picture processing. On the other hand side learners could be helped with the integration of at least two representations, i.e. they were helped with structure mapping between the displays. Brünken et al. (2003a) developed helping strategies for local (text and picture processing help) as well as for global coherence formation (help for integration) and tested their effects on understanding the content of the text and/or the picture. Text processing was fostered by elaborative questions about the text, picture processing was assisted by a picture labeling task, where students had to drag labels at the correct position in the picture. The integrative help was realized with inter-textual hyperlinks. Help for text processing resulted in better text performance, help for picture processing resulted in a better understanding of the picture, however only the help for integration affected both, text and picture comprehension. Bodemer et al. (in press) also revealed that an overall understanding of symbolic and pictorial representations can be significantly improved by actively integrating the given representations. They activated learners to engage in integration by a drag-and-drop task between text and picture.

2.2.2 How should help be offered?
Directive or non-directive?
Based on the discussion whether learners should be guided briefly or should be enabled to act self-regulated and active, help for coherence formation can be offered more or less directive. Directive help means that learners are given explicit hints, which items can be mapped and their attention is strongly guided by the help information. Less directive help in contrast would enable learners to find corresponding elements by themselves and they are only given hints, that there is something to map without telling them what it is. Both types of help were tested empirically and there is evidence for positive effects of directive help (Peeck, 1993; Reinking, Hayes & McEneaney, 1988; Glowalla, Rinck und Fezzardi, 1993) as well as for less directive help (Moore und Seevak, 1994, Seufert, 2003a, b). Nevertheless, there is one critical aspect of less directive help that is mentioned in several studies: the self-regulation of the mapping also needs cognitive resources and therefore competes with the learning task for limited cognitive capacities. Another result is also in favor of the less demanding directive help: because of the explicitness and
the repetition of relevant concepts of the learning material, it proved to increase especially learners recall performance (Seufert, 2003a, b).

Visually or auditively?
According to the well-known modality effect which is empirically proved in several studies in line with cognitive load theory (Brünken & Leutner, 2001; Moreno & Mayer, 2002; Tabbers, 2002, Tindall-Ford, Chandler & Sweller, 1997) it is also necessary to test, whether help should be given in visual or in auditive form. In a study from Seufert (2003a) learners revealed slightly better learning outcomes when help was offered auditively but this effect failed to reach statistically significance. On the one hand there is the advantage of auditive help that it does not overburden the visual channel of working memory, but on the other hand there is the great disadvantage of auditive help, that it is not as stable as the visual help is. Especially when the help information is complex it is important for learners to read it or hear it repeatedly, whenever or as often as they need it. Here more research is needed to analyze the specific pros and cons of both presentation modalities.

2.2.3 When should help be offered?
Generally help can be given simultaneously with the learning situation, or in contrast learners can be trained on all necessary abilities and strategies for coherence formation before learning. All helping strategies mentioned so far can be categorized as situational help strategies and in summary they turned out as beneficial (as long as they don’t require too much cognitive capacity in addition to the learning task). But there is only little empirical evidence for the effectiveness of training for coherence formation. In an explorative study we developed a training which included text processing, picture processing as well as integration strategies and tested the effects on learners’ global coherence of the subject matter. Compared to an untrained group the training turned out to be even harmful for learning. This mathemarithmetic effect (Artelt, 2000; Friedrich & Mandl, 1992) was shown in several studies of strategy training and can be explained with interference between an existing strategy of the learner with the new one. I.e. the old strategy didn’t work any longer and the new one wasn’t yet proceduralized, so that learners experienced a lack of strategy immediately after the training.

2.3 Which learner needs how much help?
One of the most important factors with respect to the effectiveness of help is the learners’ prior domain specific knowledge. With insufficient prior knowledge learners often have problems to use help information effectively because they are not able to handle more than one source of information. However, if learners’ prior knowledge is too high help also fails to success because learners don’t need any help. Only for learners with a middle level of prior knowledge help can be effective, because they are able to use it and they need to be assisted. This pattern was tested
SUPPORTING COHERENCE FORMATION

empirically in a study from Seufert (2003b) with verbal semantic help for coherence formation. As figure 2 shows there is a medium increase in performance for learners with low levels of prior knowledge (d=.69), whereas learners with middle levels of prior knowledge could strongly increase their performance when help was offered (d=1.42). Also interesting is the slight decrease in performance of the group with high levels of prior knowledge (d=.38).

![Figure 1: Comprehension scores of learners with a low, medium and high level of prior knowledge depending on different treatments.](image)

This effect is in line with current findings on the expertise reversal effect (Kalyuga, Chandler & Sweller, 2000; Kalyuga, Ayres, Chandler & Sweller, 2003). Instructional guidelines or features, like adding coherent information or an integrated instead of a separated presentation of information may be beneficial for learners with low experience but they turn out to be harmful for experts. All added information which is unneeded requires cognitive capacity because it has to be evaluated whether it has an added value or is redundant and can be ignored. In summary, the most promising strategy to help learners with insufficient prior knowledge is to teach them the most important concepts directly, while the best help for experts seems to be no help.

3. SUMMARY AND DISCUSSION

Multimedia learning environments comprise multiple representations, which have to be integrated mentally in order to construct a coherent mental representation of the
subject matter presented. This complex cognitive task turned out to be highly
demanding especially for learners with lower levels of prior knowledge.
Accordingly learners often avoid investing too much cognitive resources on the
integration process and rather concentrate on parts of the given information or on
single representations. However, effective learning can only take place when
learners really engage in mapping the corresponding elements and structures of the
multiple representations to build up a more abstract, elaborated and coherent
knowledge structure.
Consequently the costs of cognitive processing should be reduced in order to reach
the goal of global coherence. Different ways of reducing costs were presented and
discussed above. To guide learners’ attention by signaling and highlighting
converting items is one way to assist learners mapping process on what we call a
surface level. In terms of cognitive load theory this kind of support would be a
*reduction of extraneous cognitive load*. The highlighting as a design feature
facilitates necessary cognitive processes and therefore relieves learners’ working
memory. But taking into account the risk that learners only superficially process the
highlighted items it seems to be additionally beneficial to assist the integration on a
deep structure level. Also in line with cognitive load theory such kind of mapping
help can be seen as an *increase of germane cognitive load* such as it activates
learners to deal with the learning content actively and elaborate it (Bannert, 2002;
kinds of such semantic help for coherence formation have been developed and
evaluated. Thus, two crucial aspects can be pointed out: first, help for coherence
formation depends on learners’ expertise and second, support can only be helpful if
it is not too demanding and does not compete for cognitive resources with the
learning task.

**AFFILIATIONS**

*Tina Seufert, Roland Brünken  
Georg August University of Göttingen, Germany*

**REFERENCES**


Herbstschule Kognitionswissenschaften, Freiburg.


Instruction, 12*, 139-146.

during Learning with Dynamic and Interactive Representations. Learning and Instruction.
SUPPORTING COHERENCE FORMATION


