How does hypermedia support learning? The role of different representational formats and varying levels of learner control for the applicability of multimedia design principles

Maria Opfermann, Peter Gerjets
Knowledge Media Research Center (IWM-KMRC)

Katharina Scheiter
Department of Applied Cognitive and Media Psychology
University of Tuebingen

During recent years, hypermedia and web-based learning environments have become increasingly important in educational contexts. The advantages they offer compared to traditional learning methods (like books) include the possibility to access information in a nonlinear and self-controlled fashion. Additionally, information can be presented in different representational codes (e.g., text, pictures) and address different sensory modalities (e.g., visual, auditory). However, the question arises how these aspects should be combined to design a hypermedia environment that enables active, self-regulated and constructive learning and fosters knowledge acquisition. Our studies investigated whether well-established multimedia design principles apply to hypermedia as well. Results show that these principles cannot simply be transferred to hypermedia environments and that certain representational formats do not foster learning per se but that it is necessary to carefully look at the affordances that these representations provide for retrieval. These results will be presented and discussed with respect to their implications for the design of further studies.

Keywords: hypermedia learning environments, multimedia design principles, learner control

Introduction

Computers and the World Wide Web as media for information delivery, as well as for information search and communication, have gained significant influence during the last decade. Thinking of tertiary educational contexts, it is nearly impossible to imagine students and lecturers working without computers. One way of conveying information is through hypermedia learning environments that are characterized by offering a high amount of learner control. On the one hand, this means that learners have the option to select and combine different representational codes (e.g., text, static or dynamic visualizations) and address different sensory modalities (visual, auditory). On the other hand, they can access information in a linear as well as in a nonlinear fashion. Ideally, this navigational and representational freedom leads to active, constructive, and self-regulated as well as adaptive learning. However, such benefits can only take place if learners are willing and able to make the right decisions with regard to the contents they want to access as well as the rate and sequence for retrieving this content. Otherwise, hypermedia environments run the risk of leading to the assembly of suboptimal information diets, to disorientation and accordingly to cognitive overload. To avoid such disadvantages, it is therefore of pivotal importance to carefully design hypermedia environments. This requirement refers to the design of content of the environments as well as to the degree of learner control provided.

Designing the content of hypermedia learning environments

When starting our research, we found that there were hardly any recommendations that prescribe how to design hypermedia learning environments with respect to representational codes and sensory modalities. However, research on multimedia learning has extensively dealt with this topic and has provided a couple of multimedia design principles that specify how these different representational codes and sensory modalities should be combined to foster learning. Recent theories, such as the cognitive theory of multimedia learning (Mayer, 2001), the cognitive load theory (Sweller, 1999), and research on multiple representations (Ainsworth, 1999), recommend using multiple representations (Ainsworth, 1999; Mayer, 2001), presenting information in different modalities (Mayer, 2001), avoiding redundant information (Mayer, 2001) and taking into account individual differences (e.g., aptitude-treatment-interactions) in instructional design (Kalyuga, Ayres, Chandler, & Sweller, 2003; Mayer & Gallini, 1990). Although these principles have been empirically validated in controlled laboratory studies on multimedia learning, the question remained still open whether they can be simply transferred to hypermedia environments. In fact, there is already some initial evidence that specific multimedia-design principles, for instance the modality principle, are moderated by learner control (Tabbers, 2002).

Adaptive information utilization and optimal degrees of learner control

The high amount of learner control that hypermedia environments allow for means that they are capable of being explored in multiple ways, thus offering adaptive information utilization. However, the question arises whether learners really take advantage of this opportunity to select and combine optimal information diets in an adaptive way or whether they benefit from a more structured information presentation. It is also as yet unclear whether there are optimal levels and types of learner control for different kinds of learners and learning tasks. There is some evidence that the hypothesized advantages of a high level of learner control are valid for learners with high prior knowledge only. In this regard, Gall and Hannafin (1994) suggest that prior knowledge may guide learner-controlled behavior in that "individuals with extensive prior knowledge are better able to invoke schema-driven selections, wherein knowledge needs are accurately identified a priori and selections made accordingly. (...) Those with limited prior knowledge, on the other hand, are unable to establish information needs in advance, making their selections less schema-driven" (p. 222). In line with this reasoning, Clark and Mayer (2003) propose a learner-control principle that advises the use of high levels of learner control for learners with high prior knowledge or high metacognitive skills.

Experiments

The experiments we conducted addressed the following questions: (1) How should the different possible contents of hypermedia learning environments (e.g., different representational codes and sensory modalities) be designed and combined to foster efficient learning for different types of learners? (2) How much learner control should hypermedia environments allow for (depending on individual prerequisites) to optimize learning while avoiding cognitive overload?

Method

Participants
196 pupils from 6 German high schools participated in the study: 114 girls and 82 boys from grades 10 and 11 with an average of 16.55 years.

Materials and procedure
The learning environment the pupils worked with was a hypermedia environment on probability theory that aimed at conveying the basic principles of the domain by means of worked-out examples. Learners working with it had to acquire knowledge about four different categories of probability theory. The environment consisted of a personal data questionnaire, a short technical instruction, a pre-test to assess prior knowledge on probability theory, a domain introduction, an example-based learning phase with the eight worked-out examples, and a post-test.

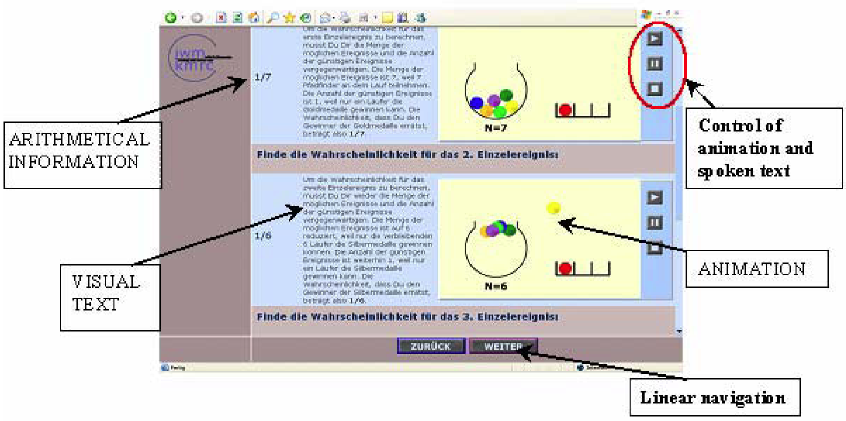
Design and dependent measures
Depending on the respective experimental conditions, different representational codes and sensory modalities were used to present the worked-out examples (Table 1). Moreover, two different levels of learner control were implemented.

Table 1: Experimental design

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="2">Control of information presentation</th></tr>
<tr><th>Low level of learner control</th><th>High level of learner control</th></tr>
<tr><td rowspan="6">Presentation format</td><td>Arithmetical only</td><td>19</td><td rowspan="6">78</td></tr>
<tr><td>+ written text</td><td>20</td></tr>
<tr><td>+ spoken text</td><td>16</td></tr>
<tr><td>+ written text + spoken text</td><td>20</td></tr>
<tr><td>+ written text + animation</td><td>25</td></tr>
<tr><td>+ spoken text + animation</td><td>18</td></tr>
</table>

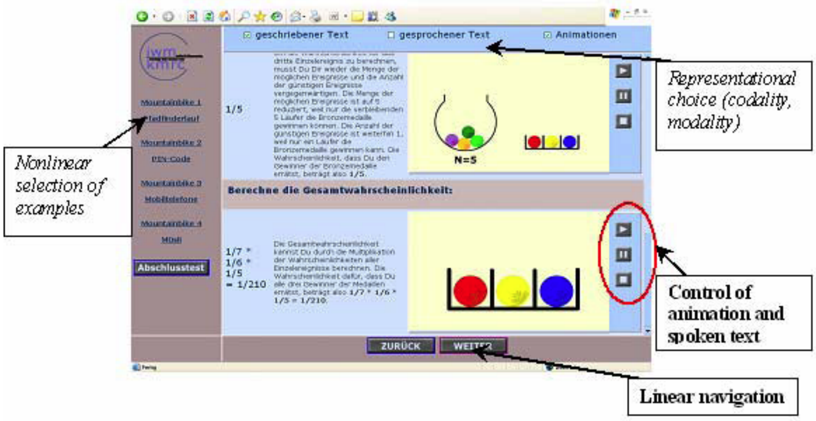
For each solution step, all experimental conditions contained arithmetical information. In the version with a low level of learner control, there were six different pre-defined formats. The "written text" presentation format provided additional visual instructional explanations of solution steps. In the "spoken text" presentation format these explanations were presented auditorily. The "written text + spoken text" presentation format contained redundant information in that it provided both types of verbal information.

In the remaining two presentation formats, written or spoken text was augmented with abstract animations. Learners who were assigned to one of these conditions could only navigate through the environment in a linear fashion by clicking the "Back" and "Next" buttons at the bottom of each page. In the condition with a high level of learner control, learners could choose the representational format by clicking the respective buttons in the upper options bar (Figures 1 and 2). They could also navigate through the environment in a nonlinear fashion by using the navigation bar on the left side of each page. Differences between the two levels of learner control therefore pertained to the selection and sequencing of worked-out examples, as well as the opportunity to choose between different representational formats, while pacing was available in all versions of the environment. Moreover, the dynamic representations were always interactive in that they were presented only on learners' demand. In all conditions (low and high levels of learner control), learners received instructions on how to retrieve the worked-out examples to make sure they were aware of the representational and navigational choices they had.



Note. Labels in capitals refer to codality and modality aspects of the environment. Labels in bold highlight control features present in both versions with low and high levels of learner control.

Figure 1: "Written text + animation" – version with a low level of learner control



Note. Labels in italics refer to the control features characteristic of the version with a high level of learner control.

Figure 2: Version with a high level of learner control

Once learners had worked through the environment at their own pace, they continued with the 44-item post-test aimed at assessing conceptual, intuitive, procedural, and situational knowledge. As dependent variables, we registered performance for the overall post-test as well as for the subcategories, learning times, and cognitive load during learning. However, due to the different amount of information provided in the experimental conditions, we expected large differences in learning times across conditions. Accordingly, using post-test performance alone as a measure for the instructional benefits of the different conditions would be grossly misleading. Thus, we calculated efficiency scores that integrated post-test performance and learning times by adapting an approach of Paas and van Merriënboer (1993). Efficiency was expressed as the difference of the performance *z*-score and the learning time *z*-score divided by the square root of 2:

$$E = (z_{\text{performance}} - z_{\text{learning time}}) / \sqrt{2}.$$

A negative score for *E* states that the relative investment of learning time exceeded the relative performance; a positive score stands for high performance scores compared to the learning time. Cognitive load was assessed with six items that assessed intrinsic (one item), extraneous (three items), germane (one item), and overall load (one item) rated on a 9-point Likert scale.

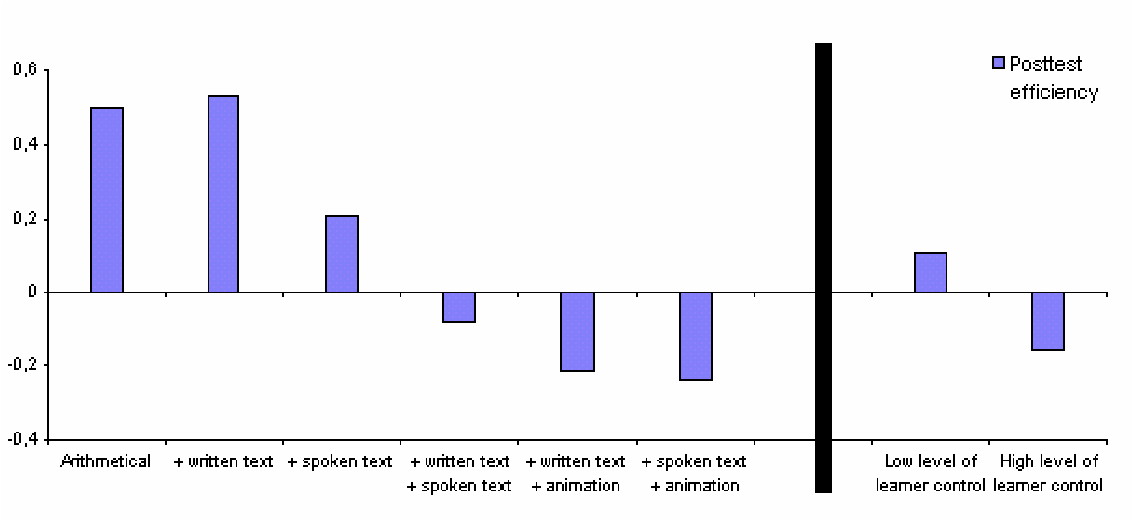
Results

Cognitive load during learning

Cognitive load was measured each time after learners had worked through one of the probability categories with two worked examples. Results showed that extraneous load did not differ across experimental conditions. Scores for intrinsic, germane and overall load were significantly higher in the "arithmetical only" than in the other conditions, where additional explanations of solution steps were provided. The latter conditions did not differ from each other regarding cognitive load. With respect to changes over time, intrinsic ($F(1,196) = 10.63$, $p < .001$) and extraneous ($F(1,196) = 19.65$, $p < .001$) load decreased significantly, while germane load ($F(1,196) = 23.92$, $p < .001$) increased significantly. These results are in line with cognitive load theory, which suggests that when extraneous and intrinsic load decrease, more working-memory capacity can be claimed by cognitive processes directly relevant for understanding, that is, germane load.

Instructional efficiency

To test the validity of the multimedia design principles, we compared the experimental conditions with a low level of learner control with regard to their instructional efficiency by means of one-way ANOVAs. The overall post-test efficiencies for these conditions can be seen in the left half of Figure 3.



Note. Due to space limitations, only overall post-test efficiencies but not efficiencies for the subcategories are depicted.

Figure 3: Efficiencies for the single experimental conditions and for low versus high levels of learner control

The *multimedia principle* was examined by comparing the "written text" with the "written text + animation" condition and the "spoken text" with the "spoken text + animation" condition. The efficiencies for "written text" were significantly higher than for "written text + animation" for all measures (overall post-test: $F(1,43) = 10.20$, $p < .01$), while there was only one significant difference in favor of "spoken text" with respect to situational knowledge ($F(1,32) = 9.91$; $p < .01$). These results clearly contradict the theory of Mayer (2001) in that they show that enriching visual or auditory text with pictures does not automatically improve performance.

The question whether *multiple representations* foster learning was addressed by aggregating the results of the "arithmetical", "written text" and "spoken text" conditions (i.e., single representations) and comparing this outcome to the aggregated results of the "written text + spoken text", "written text + animation" and "spoken text + animation" conditions (i.e., multiple representations). Learners studying single representations were consistently more efficient than learners studying multiple representations (overall post-test: $F(1,116) = 16.57$; $p < .001$). Thus, our results contradict the expectation that multiple representations foster deep conceptual understanding automatically (*cf.* Ainsworth, 1999).

The *modality principle* was examined by comparing the "written text" with the "spoken text" condition and the "written text + animation" with the "spoken text + animation" condition. The first comparison is based on the fact that the arithmetical information provided in all conditions is visual information as well. Comparing "written text" with "spoken text" revealed a marginally significant effect in favor of "written text" for conceptual ($F(1,34) = 3.52$; $p < .10$) and intuitive knowledge ($F(1,34) = 3.62$; $p < .10$). When contrasting "written text + animation" with "spoken text + animation", none of the differences were significant. These results again contradict our expectations derived from Mayer's theory (2001), who states that there should be a superiority in favor of spoken text and spoken text plus animation, respectively, because in those cases the presentation of information is distributed among a visual and an auditory processing channel.

The *redundancy principle* was examined by comparing the "written text" as well as the "spoken text" condition with the "written text + spoken text" condition. According to this principle, less material should result in better learning. Our results confirm this to a large extent, especially for the first comparison (overall post-test: $F(1,38) = 4.57$; $p < .05$). Learners receiving redundant information performed worse than learners who received less material.

In a last step, *usage of representations* was analyzed for the conditions with low levels of learner control, in which two aspects were left to learners, namely pacing and interactivity of dynamic representations. Analyzing the percentage of representations used in the different conditions revealed that learners in the three conditions with spoken text (i.e., “spoken text”, “written text + spoken text” and “spoken text + animation”) used the available dynamic representations to a significantly lower extent than learners in the “written text + animation” condition. Overall the dynamic representations were used only to a very small extent. Thus, they may only have provided minor affordances for retrieval within our environment.

Information utilization and optimal degree of learner control
The question which level of learner control is optimal for learners was investigated by comparing results of participants who received a high level of learner control with those who worked with low levels of learner control (*cf*. Figure 4). Conditions with a low level of learner control overall tended to be more efficient (overall post-test: $F(1,194) = 3.51$; $p < .10$). These results are in line with the argument that learners might not necessarily benefit from complete navigational freedom in that they can face problems in selecting and integrating relevant information like assembling suboptimal information diets, being disoriented and experiencing cognitive overload.

To test the assumption of Clark and Mayer (2003) who suggest that a high level of learner control might work for learners with high prior knowledge only, we analyzed efficiency scores by means of a 2*2 ANOVA with degree of learner control (low/high) and prior knowledge (low/high) as between-subjects factors. This revealed a main effect for prior knowledge in that high-prior-knowledge learners were significantly more efficient than low-prior-knowledge learners ($F(1,194) = 27.83$; $p < .001$). However, the expected interaction failed to reach statistical significance ($F(1,194) = 1.01$; $p > .30$). Thus, contrary to the learner-control principle advocated by Clark and Mayer (2003), low levels of learner control seem to be advisable for all learners irrespective of their prior knowledge level.

To investigate whether students’ *patterns of information utilization* can be used to distinguish different subgroups of learners, we conducted a cluster analysis. Four clusters of students could be extracted. Cluster 1 spent more time on playing dynamic representations than any other cluster; this time was used almost exclusively for a combination of spoken text and animation. Cluster 2 spent a medium amount of time on processing dynamic representations, which was also almost completely used for playing the integrated format. Cluster 3 used the dynamic representations very rarely. Cluster 4 differed from the other three in that they mostly used the animations only and not the integrated format; this usage, however, was also restricted to a rather medium frequency of retrieving animations. The other three groups, on the contrary, did not study examples in the animations-only presentation format at all. Interestingly, all four groups refrained from studying examples presented in a spoken-text only format.

Overall, similarly to the learners in the conditions with a low level of learner control, participants with a high level of learner control did not make much use of their navigational freedom and their freedom of representational choices. Besides the aforementioned lack of usage of representations, they also did not browse through the environment in a nonlinear fashion. Rather, they just clicked on the “Next” buttons once they had worked through a page. This also raises the question of how much affordance the bars for the nonlinear selection of examples and for the choice of representational formats have provided.

Discussion and conclusions

The results of the two studies reported were surprising in that they could confirm only one of the multimedia-design principles stated by Mayer (2001), namely the redundancy principle. Contrary to expectations that can be derived from the multimedia and modality principles, conditions with single representations yielded better results and were more efficient than multiple representations. Additionally, learners in the “spoken text + animation” condition, who should have been superior to learners in all other conditions sensu Mayer, even showed the lowest performance. What could have caused these unexpected findings?

Concerning the *modality principle,* we did not find performance improvements when distributing information among different sensory modalities. This is in contrast to Mayer and colleagues but in line with findings by Baggett and Ehrenfeucht (1983) or Tabbers (2002). They demonstrated that, when there was sufficient time to read the written materials, written text yielded either equal or even superior performance compared to spoken text. These aspects apply to our learning environment (e.g., our learners were given the opportunity to read the written explanations before retrieving the animation) and may have caused these results. To test whether the modality principle holds true for environments where there are time restrictions, we have been conducting a follow-up study using the "Written text + animation" and "Spoken text + animation" conditions from the learning environment; however, these conditions were administered strictly system controlled to learners this time. They could not choose whether they wanted to play the respective animation or sound file, and animation or sound started as soon as the text appeared on the screen. Participants were 39 university students. We are currently evaluating these data and will be able to report on them at the conference.

As for the *multimedia principle*, our results showed that enriching verbal instructions with animations did not improve performance. While these findings contradict Mayer and colleagues, they are in line with authors like Tversky, Bauer Morrison, and Betrancourt (2002), who claim that dynamic pictorial representations should be carefully designed in order to be more efficient for learning than static pictures or purely textual representations. Results from Schuh, Gerjets and Scheiter (2005) suggest that so called hybrid animations, which first show the transition between concrete and abstract representations and which, secondly, show the relation between symbolic expressions (e.g., text, mathematical formulas) and their pictorial representations might be better suited to convey problem-solving skills than purely abstract or concrete animations. The question arises whether the abstract animations used in the reported studies can be improved by first showing the transition between a concrete problem statement and an abstract mathematical solution procedure and by explicitly mapping symbolic expressions onto pictorial representations (e.g., a connecting line between a fraction and its representation in the animation). Explicitly showing these relations may provide affordances for learners to think about them more deeply and may thus aid learning, whereas there might have been a lack of affordances of the current material to relate pictorial and symbolic representations.

A third issue to be discussed refers to the *learner control* provided in our studies. The initial hypothesis was that learners (at least those with high prior knowledge) might benefit from a high level of learner control in that they can adapt the information presented to their needs. However, our studies showed that dynamic representations were rarely used and that learner control did not interact with prior knowledge. There are at least three possible reasons for this finding. First, the representations might not have had high affordances, because they were not designed in an optimal way. Secondly, students might have been overwhelmed by the amount of information given at the beginning of the learning phase, which could have increased their extraneous load. According to this interpretation, it may be more beneficial to expose them to more materials only after they have had more experience with domain and instructional setting. A third explanation for the lack of benefits of a high level of learner control is that students generally do not engage in suitable information utilization strategies themselves (Gerjets et al., 2000), but rather need to be prompted to use external representations (Gerjets, Scheiter & Schuh, 2005). It seems thus advisable to incorporate instructional guidance in hypermedia environments, even for learners with high prior knowledge.

Taken together, even though we could not confirm the transferability of multimedia design principles for hypermedia, future research in this area is needed. Such research should take into account: (a) time restrictions when retrieving instructional materials, (b) a comparison of differently designed animations (e.g., symbolic versus iconic versus hybrid animations) and (c) improving the affordances of representational and navigational choices especially in conditions with high levels of learner control.

References

Ainsworth, S. E. (1999). A functional taxonomy of multiple representations. *Computers and Education, 33,* 131–152.

Baggett, P., & Ehrenfeucht, A. (1983). Encoding and retaining information in the visuals and verbals of an educational movie. *Educational Communication and Technology Journal, 31*, 23–32.

Clark, R. C., & Mayer, R. E. (2003). *E-Learning and the science of instruction.* San Francisco: Jossey-Bass/Pfeiffer.

Gall, J. E., & Hannafin, M. J. (1994). A framework for the study of hypertext. *Instructional Science, 22,* 207–232.

Gerjets, P., Scheiter, K., & Catrambone, R. (2004). Designing instructional examples to reduce intrinsic cognitive load: molar versus modular presentation of solution procedures. *Instructional Science, 32,* 33–58.

Gerjets, P., Scheiter, K., & Schuh, J. (2005). Instruktionale Unterstützung beim Fertigkeitserwerb aus Beispielen in hypertextbasierten Lernumgebungen. *Zeitschrift für Pädagogische Psychologie, 19*, 23–38.

Gerjets, P., Scheiter, K., & Tack, W. H. (2000). Resource-adaptive selection of strategies in learning from worked-out examples. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings from the 22nd Annual Conference from the Cognitive Science Society* (pp. 166–171). Mahwah, NJ: Erlbaum.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38,* 23–32.

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

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

Paas, F., & Van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental-effort and performance measures. *Human Factors, 35* , 737–743.

Scheiter, K., Gerjets, P., & Catrambone, R. (2006). Making the abstract concrete: Visualizing mathematical solution procedures. *Computers in Human Behavior, 22,* 9–25.

Schuh, J., Gerjets, P., & Scheiter, K. (2005). Fostering the acquisition of transferable problem-solving knowledge with an interactive comparison tool and dynamic visualizations of solution procedures. In B. G. Bara, L. Barsalou & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 1973–1978). Mahwah, NJ: Erlbaum.

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

Tabbers, H. K. (2002). *The modality of multimedia instructions.* Maastricht: Datawyse.

Tversky, B., Bauer Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies, 57*, 247–262.

Author contact details

Maria Opfermann, Knowledge Media Research Center, Konrad-Adenauer-Str.40, Tuebingen, Germany. Email: m.opfermann@iwm-kmrc.de.

Copyright © 2006 Opfermann, M., Gerjets, P.,Scheiter, K.

The author(s) assign to ascilite and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The author(s) also grant a non-exclusive licence to ascilite to publish this document on the ascilite web site (including any mirror or archival sites that may be developed) and in electronic and printed form within the ascilite *Conference Proceedings*. Any other usage is prohibited without the express permission of the author(s). For the appropriate way of citing this article, please see the frontmatter of the *Conference Proceedings*.