Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Processing and Integrating Multimodal Material — The Influence of Color-Coding

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN

1069-7977

Authors

Folker, Sonja
Ritter, Helge
Sichelschmidt, Lorenz

Publication Date

2005

Peer reviewed

Processing and Integrating Multimodal Material – The Influence of Color-Coding

Sonja Folker (sonja.folker@uni-bielefeld.de) and Helge Ritter
Neuroinformatics Group, Faculty of Technology, Bielefeld University
P.O.-Box 100131, 33501 Bielefeld, Germany

Lorenz Sichelschmidt
Faculty of Linguistics and Literature, Bielefeld University
P.O.-Box 100131, 33501 Bielefeld, Germany

Abstract

Constructing effective learning material has always been a central issue in education, and is even more so in the time of international student assessment. In our experiment we concentrated on the processing of typical multimodal textbook material (i.e. a written text with accompanying illustrations) and on how the processing and integration of these different representational modalities can be improved by color-coding. Color-codes reduce search effort and thus produce similar effect to an increase in spatial or temporal contiguity (Kalyuga, Chandler & Sweller, 1999). In this context, contiguity is defined as the adjacency of distinct elements in general. By using eye-tracking methodology we chose a means that could provide detailed insights into the actual process of information intake. Our results affirm the differences in processing verbal and linguistic material and show the advantage of color-coded material on processing efficiency and integration. This leads us to the conclusion that the genuine difficulty of multimodal material lies in the integration of the different sources of information and that especially learners with low prior knowledge can profit from an increase in coherence and contiguity (see Mayer, 2001).

Keywords: color-coding; eye-tracking; multimodality; text and picture processing; coherence formation.

Introduction

In oral communication and instruction it is natural to use visual and linguistic sources of information simultaneously. In line with the goal to provide more naturalistic learning environments in educational contexts, multimedia approaches therefore often favor a combination of spoken text with moving or static pictures to avoid split-attention effects (Mayer, 2001).

However, the research of Guan (2003) shows that the use of different representational modes addressing only the visual modality (e.g. written text and static pictures) can prove just as efficient as a dual modality presentation where spoken text and static pictures are presented. At first glance, this might be a surprising result: in understanding an illustrated text, text and pictures can only be considered sequentially, which should complicate the integration process, while in hearing a description of a picture while looking at it, linguistic and pictorial information is accessed simultaneously. At a second glance, good reasons can be found for the use of "written-only" material: it gives the reader more *control* over his intake of information, i.e. he can switch between the sources of information at will, he even can ignore one of the sources entirely or (re)consider parts of the text and the picture in a self-determined sequence and at his own pace. The aspect of control becomes even more crucial with an increase in complexity: our brain clearly evolved to process spoken language, however, the advantage of using written language to convey complex information is undeniable. And it has often been documented that the combination of a text with illustrations has a positive effect: illustrations are known to draw the reader's attention and to ameliorate motivation, comprehension, and retention (see Carney & Levin, 2002).

While the demonstrated benefit of multimodal material is highly relevant for instructional issues, the question of how the processing of these materials is cognitively achieved remains open. The first basic assumption concerning the processing of multimodal material was made by Paivio (1986), who postulated two interconnected but functionally independent subsystems for human information processing, a verbal and a nonverbal (pictorial) one. The verbal and visual representations (logogenes and imagines) are considered to be built up separately; nonetheless referential links between corresponding logogenes and imagines can be established. The dual coding assumption is preserved in current cognitive models of multimodal information processing: Mayer (1993) explicitly presents a dual coding theory of learning from visual and verbal materials and later models of Mayer (1997) and Schnotz and Bannert (1999) also assume two clearly differentiated though in parallel working channels for building up mental representations of texts and pictures.

However, the goal of learning with multimodal material does not consist in building up separate representations of different sources of information but rather to build up referential links and to integrate the information processed in both channels. There are different views as to the cognitive demands of this integration process: Rayner, Rotello, Stewart, Keir and Duffy (2001: 219) affirm that "[p]resumably, most skilled readers are adept at alternating between text and pictures to produce a mental model of the complete message". Nevertheless it has to be kept in mind that reading (as information processing in general) depends on physiological processes and cognitive resources that are necessarily restrained. Visual information intake is limited to the area foveally and (depending on the type of stimulus material) parafoveally perceived during the fixations, i.e. the phases where the eye is kept relatively stable. Further processing is limited by attentional mechanisms (Liversedge & Findlay, 2000) and the capacity of our working memory.

These restrictions implicate that the use of written multimodal material runs the risk of asking too much of the reader, depending on the quality of the material on the one hand and on the reader's prior knowledge, his general reading capacities and his interest and motivation on the other hand. In the latter case, even "good pictures" can "fail", i.e. they might not lead to the frequently demonstrated advantage over a text without illustrations (Weidenmann, 1989, but cf. Mayer, 2001). Two possible reasons for this failure can be identified: the reader might concentrate entirely on the text, fail to process the pictures sufficiently and therefore not profit from the additional information they provide. Or else the reader might be able to process the text as well as the pictures but not have sufficient cognitive capacities to integrate the different representations.

The research by Weidenmann (1989) focuses on processing difficulties. In his view, pictures are conceived as easily accessible, as the eye quickly identifies the most informative details of the picture and adjusts the following eye movements accordingly. Thus picture processing often stops at a superficial level giving the learner the illusion of a full understanding that is in fact not being achieved. Following Weidenmann, this risk of underestimating pictures is even higher if the picture is combined with a text, as the text often is perceived as the better medium to provide the content to be learned. Drawing attention to the pictures by the means of instructions can therefore foster the learning process (Peeck, 1993) - provided that the pictures actually risk being underestimated. As Weidenmann points out, this might not be the case if the task is perceived as demanding, thus inducing a sufficient level of processing of the text and its illustrations.

But even if the two distinct sources of information are sufficiently processed, the difficulty of integrating the information remains: to build up an integrated representation, the reader has to recognize the referential connections between the two sources; he has to align the sources in order to match corresponding parts, to detect inconsistencies and to add complementary information. Even without having detailed theoretical assumptions on how the integration process is actually accomplished, the alignment can be assumed to be facilitated if the material itself underlines the existing referential connections between text and pictures. Following this view, good multimodal material therefore is characterized by a high degree of contiguity and coherence.

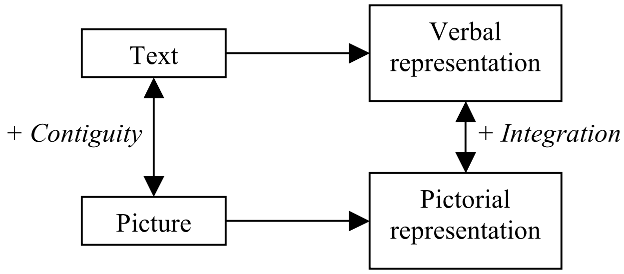


Figure 1: The influence of contiguity on integration.

Previous studies on multimodal material have indeed documented the positive influence of an increase in contiguity on different levels. As Mayer and Anderson (1992) have shown, high tempo-spatial contiguity constitutes the basis for a successful integration: pictures have to be shown in the same time and spatially close to the passages of the text they illustrate. Moreover, a high degree of contiguity should exist on the semantic level, i.e. text and pictures should be closely related to each other in regard to the content (Carney & Levin, 2002) and superfluous information should be suppressed to facilitate the integration process (Mayer, Heiser, & Lonn, 2001).

Up until now, studies on the contiguity of multimodal material mainly assessed the retention performance depending on significant changes in the degree of contiguity (i.e. showing text and pictures simultaneously vs. sequentially or suppressing superfluous information). There are few studies using eye-tracking in combination with multimodal material; they reveal reciprocal influences of picture and text processing. If the text is longer and the illustrations closely related to the text, *switches of attention* to the illustrations necessarily interrupt the text processing sequence. These switches generally occur at major syntactic and semantic boundaries (Hegarty & Just, 1989). The picture viewing process is highly *text-directed*, i.e. the pictures are only considered after having read the text and according to the information the text provides (Hegarty & Just, 1993; Carroll, Young & Guertin, 1992).

As color-codes strengthen the referential connections between separate information sources, the can create the same effects as an increase in contiguity. Parting from the idea that color can produce strong bottom-up effects on the attraction of gaze and assist higher cognitive processes such as structuring and coherence formation (Marcus, 1992), structures in the illustration corresponding to words or phrases in the text were colored identically (see Fig. 2). For example: if the word "chromosome" was colored in blue, the fibrous structures in the schematic cell representing the chromosomes as well as the corresponding labels were colored identically. The introduction of these *color codes* should have the same effect as an increase in contiguity, facilitate the alignment and matching processes, and thus lead to an enhanced integration process (see Kalyuga, Chandler and Sweller, 1999).

Experiment

The experiment we conducted was designed both to gain insight into the processing of multimodal material in general as well as to test whether the use of color-codes can foster the processing and integration of this type of material. Standard material and color-coded material was tested on two separate groups of participants (control group and color group). The data of the control group serves to discuss general processing issues while a comparison between groups should yield information about the influence of color-coding.

The changes in our material were rather subtle: content and configuration of the material remained unchanged. Moreover, the learning outcome mainly served as a control as our main interest lay on the actual processing. Therefore we decided to use eye-tracking methodology to gain a finely grained and on-line access to multimodal information processing. As eye movements are closely correlated to the reader's allocation of attention and the accompanying intake of information (Just & Carpenter, 1987), their analysis allows for conclusions to cognitive processes.

Method

Participants A total of 20 subjects with normal or corrected-to-normal vision were tested individually. The color-blind or color-impaired subjects were allocated to the normal stimulus group. All subjects were students, 14 females and 6 males, aged 20 to 34 years with a mean age of 24.3 years with a standard deviation of ± 4.0 years. They were paid for their participation in the experiment.

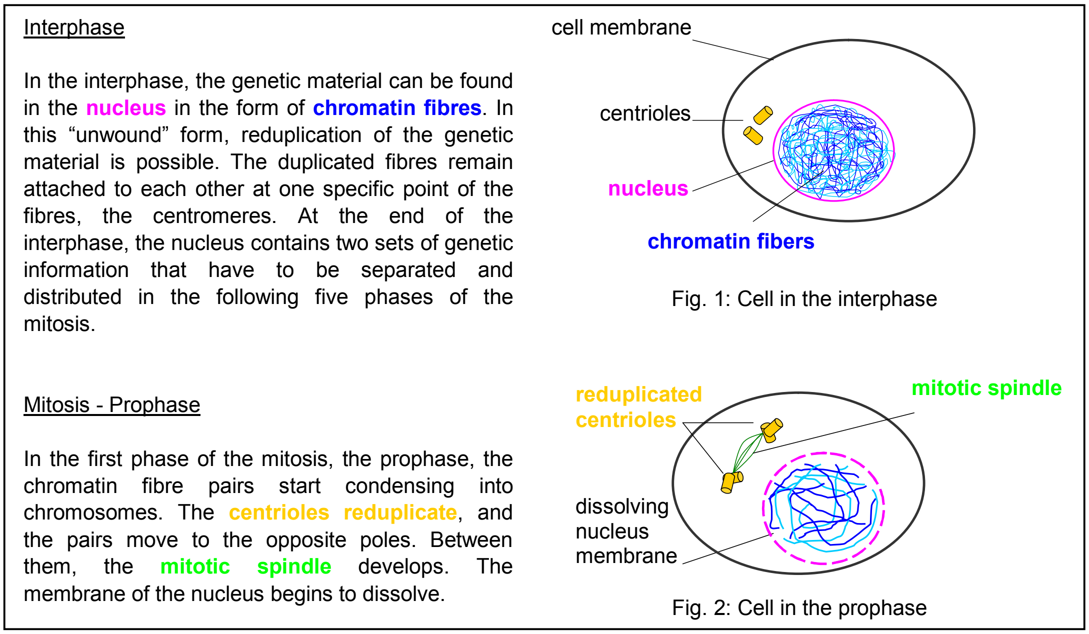
Material We designed effective multimodal material, i.e. material already displaying a high degree of contiguity. To guarantee tempo-spatial contiguity, a standardized design was chosen. A text with an average of 118.1 words was presented on the left hand side of the screen whereas the right side was divided in two equal parts, showing two pictures related to the text in chronological order (i.e. the upper picture referring to the upper part of the text and the lower picture to the lower part of the text, see Fig. 2). The pictures were schematized, suppressing superfluous information like details of cell structure (*semantic contiguity*).

The material described the function and the different phases of mitotic cell division. The choice of the topic was motivated by the following two aspects: the description of the spatial configuration of cell bodies in the different stages of mitosis necessitates the use of pictures (*semantic contiguity*). An evaluation of the material showed that the subjects indeed perceived the pictures as being important for the understanding and as highly related to the text (median of 4.0 and 4.5 on a five-point scale resp.). Moreover, the topic had the advantage of being in principle known to the subjects (the mitosis is an obligatory matter in biological education in German High school) – and at the same time not too familiar. The choice of topic and a preliminary questionnaire assessing the subject's expertise allowed for the control of *prior knowledge*, holding it at a consistently low level. This is especially important as prior knowledge can be considered as one of the most decisive factors for effective processing (Mayer, 2001).

The pictures were identically colored in both conditions. In the color condition, passages of the text corresponding to structures or labels in the picture were colored additionally (see Fig. 2). As the subjects were informed about the fact that the material was color-coded and even received an example of color-coded stimulus material before starting the trial, they could consciously use the color to direct their attention.

Recording and analysis of eye movements Eye movements were registered using the SMI EyeLink I eye-tracker. Only fixations exceeding a threshold of 100 ms were included in the analysis. In cognitive-oriented eye tracking research, a fixation threshold is usually introduced for the following reasons: due to saccadic suppression, information intake is considered to be possible only during longer fixations (Matin, 1974). In addition to this, the threshold also serves to eliminate the noise due to unsteady fixations (i.e. one long fixation is recorded as a suite of short fixations in one region).

Both eyes were tracked. It is typically assumed that left and right eye movements are conjugated (Rayner, 1998) and therefore could be expected to lead to the same experimental results. We also tested the ocular dominance of the subjects: as expected most of them showed a right eye-dominance. As



Interphase

In the interphase, the genetic material can be found in the nucleus in the form of chromatin fibres. In this "unwound" form, reduplication of the genetic material is possible. The duplicated fibres remain attached to each other at one specific point of the fibres, the centromeres. At the end of the interphase, the nucleus contains two sets of genetic information that have to be separated and distributed in the following five phases of the mitosis.

Fig. 1: Cell in the interphase

Mitosis - Prophase

In the first phase of the mitosis, the prophase, the chromatin fibre pairs start condensing into chromosomes. The centrioles reduplicate, and the pairs move to the opposite poles. Between them, the mitotic spindle develops. The membrane of the nucleus begins to dissolve.

Fig. 2: Cell in the prophase

Fig. 2: Example of color-coded stimulus material (translated).

the issue of whether to use only the dominant eye for analysis is an issue under current debate (Goldberg & Wichansky, 2003), we decided to use the right eye for analysis.

Procedure First the subjects were asked to fill in a questionnaire assessing their prior knowledge on the topic. All subjects then performed a visual acuity test with Landolt rings as well as the Ishihara test for color-blindness.

All experiment-relevant instructions were presented in written form on the computer screen in the course of the experiment. The subjects were instructed to read the presented material thoroughly in view of a test that would be given afterwards. To familiarize the subjects with the design, they first saw a stimulus example showing the same configuration as the test material; the example was unrelated in regards to the content, offering a short account of blood transfusion. Subsequently, two example questions were presented (one text related multiple choice and one picture related recall question) to give the subjects an idea of what would be asked of them in the knowledge test. During the reading of the stimulus items eye movements were recorded. Between each item a drift correction was conducted.

The reading task was followed by the test for ocular dominance, i.a. to prevent recency effects in the knowledge test. Then the subjects performed a computer-based knowledge test consisting of 24 questions referring either to the text or to the pictures or to both sources of information. The test comprised multiple-choice questions (i.e. recognition) as well as fill-in-the-blank tasks (i.e. recall) and two sorting questions.

For each answer they gave, the students had to decide whether they considered the text, the picture or their own prior knowledge as their main source of information for answering the question. The subjects always had the possibility to answer “I’m not sure” to the knowledge as well as to the information source questions to prevent them from just answering by chance.

Results and Discussion

Differences in text and picture processing

Results Statistical analysis of the control group data with pair wise t-tests yields the following results: There are significant differences in the cumulated fixation duration. On average, subjects spent 56.5 seconds on reading the text and only 20.8 seconds on the equally sized picture region (t(9)=8.496, p<.001). The pictures are processed with significantly less fixations (text average of 261.7 vs. picture 85.5, t(9)=8.961, p<.001) of a significantly longer mean duration (text 216.9 ms, picture 243.6 ms, t(9)=5.139, p<.01). If subjects received a question in the knowledge test that related to information occurring in the picture as well as in the text region, the subjects perceived the text as the main source of information, (t(9)=2.493, p<.05). Analyzing the scan path, we found that of the first 10 fixations, 7.8 are on the text, only 2.2 on the picture region (t(9)=5.979, p<.001).

Discussion In line with existing studies our data show significant differences in the attention allocation between text and pictures (see Rayner, 1998): significantly more time is spent and more fixations are made on the text region than while studying the illustrations. The results seem to support the hypothesis of Weidenmann (1989) that pictures in multimodal material are processed insufficiently and that the text is considered as the main source of information. Furthermore, this view is buttressed by the introspective data collected in the knowledge test: if the question related to information that could be found in the text as well as in the picture, the subjects mostly named the text as their most relevant source of information.

However, one can argue as well that few fixations are sufficient to understand the gist of a scene, given that, compared to reading, an increased extra-foveal intake of information is possible in scene perception (Pollatsek, Rayner, & Collins, 1984) and that the pictures only illustrate part of the information given in the text. Following this assumption, the small number of fixations would only underline the processing differences between text and pictures. The control group data alone therefore does not seem to provide grounds to argue for one of these assumptions.

In fact, some of our results confirm the typical differences in the processing of texts and pictures stated in literature: texts are processed with many short fixations, pictures with few and longer fixations. Interestingly, the difference in fixation duration is not as pronounced as in processing text or pictures alone: in his review, Rayner (1998) gives values of an average fixation duration of 330 ms during scene perception and 225 ms during reading. While our reading data approximates this value, fixation durations on pictures in a multimodal context seem to be clearly reduced compared to scene perception. The text seems to provide constraints for the picture perception process, rendering the picture viewing process more effective. The analysis of the first ten fixations underlines the directive function of the text: in spite of the property of pictures to attract attention, it is mainly the text that is considered first (Hegarty & Just, 1989).

As the discussion shows, an analysis of the control group data alone necessarily remains descriptive; it cannot yield results on the quality of the processing or the degree of integration achieved. Differences in processing times between text and pictures might simply be due to differences in information content (i.e. the text might simply be more complex than the picture); differences in processing of pictures between conditions, however, should be due to color-coding as the content of the material itself remained unchanged.

Therefore, a comparison to the data of the color group is necessary to determine if color-coding actually leads to an enhanced processing and integration of the material. As presentation in a multimodal context has a stronger influence on picture processing than on the highly constrained text processing, we also expect changes to be more pronounced in picture processing. Therefore, a comparison might provide evidence on the question of (in)sufficient picture processing as well.

Effects of color-coding

Results The subjects of the color group were significantly faster in processing the multimodal material: on average, it took them 76.6 sec to process one page of the test material while the subjects of the control group needed 97.1 sec (t(18)=2.223, p<.05). However, an analysis of variance (ANOVA) reveals that the differences in the number of fixations and the cumulated fixation duration are only approaching significance, lower values were found in the color group (284.0 vs. 347.2 fixations, F(1,18)=3.677, p<.1 and 63.5 vs. 77.3 sec. fixation duration, F(1,18)=3.371, p<.1). The ANOVA confirms the known differences in attention allocation between text and picture as to the number of fixations (F(1,18)=241.658, p<.001) and the cumulated fixation duration (F(1,18)=207.629, p<.001) but yields no significant interaction effect. Although there is a slight increase in the number of switches between the control condition and the color condition, this difference is not significant (9.47 control vs. 12.12 color, t(18)=1.518, p>.1). But the number of switches/second is significantly higher in the color condition than in the control condition (0.157 vs. 0.097, t(18)=4.344, p<.001).

The significant differences in picture and text processing together with our expectation that an increase in contiguity influences picture processing in particular motivate a more detailed between-groups analysis. In fact, the cumulated fixation duration as well as the number of fixations on the picture region are significantly reduced the color condition (15.04 sec. vs. 20.80 sec., t(18)=2.120, p<.05, 62.56 vs. 85.50 fixations, t(18)=2.182, p<.05). The comparison of the fixation duration on the text region between the two conditions, however, yields no significant results.

The two groups do not differ significantly in prior knowledge or in performance on the knowledge test (mean score prior knowledge: 23.4 (control) vs. 26.1 (color), t(18)=.489, p>.1; mean score knowledge test: 33.1 (control) vs. 34.1 (color), t(18)=.310, p>.1). However, the difference in the overall time for answering the questions is approaching significance, the color-group being faster than the standard group (337.8 vs. 434.7 seconds, t(18)=2.023, p<.1).

Aside from the attention allocation, the capability to answer the picture-related questions in the knowledge test correctly is an important indicator for sufficient picture processing. An ANOVA reveals that no significant differences can be found between the two conditions in answering the knowledge questions correctly (F(1,18)=.027, p>.1) and no interactions exists between the condition and the question type (i.e. questions relating to text, pictures or text and pictures in combination). However, a significant effect as to the type of question can be found (F(2,36)=16.505, p<.001). Subsequent pair-wise comparisons of the mean values show that no significant differences exist in the amount of correct answers to picture and text questions (57.14% text and 63.64% picture questions, t(19) =1.684, p>.1). Only the performance on questions relating to information represented in the text as well as in the pictures significantly surpasses the performance on text and picture questions (t(19)=5.605, p<0.001 and t(19)=4.383, p<.001).

Discussion The shorter processing times on the reading task show that the material in the color condition is easier to process than the standard material. Color-codes seem to provide orientation and reduce search processes, thus freeing cognitive capacities (Kalyuga, Chandler & Sweller, 1999). The differences in processing the knowledge test are approaching significance, the color group being faster in answering the questions. The time advantage is not due to a less intensive reading or superficial completion of the knowledge test, as the scores on the test do not differ between the groups. Color-coding therefore seems to facilitate the processing of multimodal instruction material and might even lead to more efficiently accessible representations.

As the text is accompanied by two different pictures, a minimum of four switches (from text to picture 1 and back and from text to picture 2 and back) should occur. Given the complexity of the text, an average of 9.47 or 12.12 switches does neither appear excessive nor represent a more or less isolated processing as found by Carroll, Young and Guertin (1992). Due to the significant differences in processing time between the two groups, the relative number of switches over time seems more appropriate to reflect differences in the efficiency of the integration process. As the number of switches per second is significantly higher in the color condition, a more integrated processing can be postulated.

Therefore our results indicate that the use of color-codes in multimodal material has similar effects as an increase in contiguity and fosters its processing and integration. The detailed analysis of the eye-tracking data documents that this amelioration is not achieved by paying more attention to the pictures. In contrast to Weidenmann's (1989) hypothesis the data shows that the number of fixations on the picture section as well as their duration is still reduced, this reduction being significant in contrast to the reductions on the text regions. Even fewer and shorter fixations than in the control condition seem to be sufficient for processing the picture.

The fact that the pictures are sufficiently processed even in the color condition is corroborated by the data from the knowledge test. Questions concerning pictures are answered as well as text questions although the amount of free recall tasks is higher than for text related questions (90.91% vs. 57.14% recall questions) and therefore should have been more difficult to answer. The higher difficulty might have been balanced out by the fact that pictorial material in general proves to be more efficiently remembered (picture-superiority effect (Paivio, 1986)) and by the fact that pictorial presentation of test items leads to better results in the knowledge test, especially if the subjects received the learning material in a pictorial form as well (Brünken, Steinbacher, Schnotz & Leutner, 2001).

The reduced fixation times are worth further discussions. The reduction can be interpreted as an increase in processing efficiency; parts of the picture are more easily identified (i.e. aligned with their verbal description). The results of Carroll, Young and Guertin (1992), however, seem to contradict this interpretation: if the caption of a cartoon is shown before the picture longer fixations can be found on the cartoon than if the material is presented in reverse order. Carroll and colleagues attributed this to a less explorative viewing pattern, indicating a more careful processing of the picture with the objective of integrating linguistic and pictorial information. In contrast to the experimental design by Carroll and colleagues, a less explorative viewing pattern can be expected in both our conditions as the subjects always tend to consider the text first. If the fixation duration indeed reflects the duration of the integration process, introducing color codes seems to accelerate it. As the performance on the knowledge test does not differ between groups, it is improbable that the shortening of the fixation duration reflects an inferior degree of integration or intensity of processing.

Conclusion

Our results underline the fact that texts and pictures are indeed processed differently. No indications of insufficient picture processing can be found, as the time allocated to the pictures can even be reduced in the color condition without deterioration of performance on the knowledge test. It is possible that the readers perceive the task as too complex and the pictures as too important not to sufficiently process them.

The data also shows that the text is perceived as the main medium for the acquisition of information. In trying to guide the reader's attention one should therefore start from the text (as color-codes do). In our opinion the importance and guiding function of the text justify the use of traditional textbook material even with a variety of other media at hand. This holds especially if the presented information is extensive and complex and an audio presentation risks to exceed working memory capacity (Kalyuga, Chandler & Sweller, 1999). The use of color-codes as a means of achieving contiguity effects and reducing search processes can also be recommended; color-codes increase the efficiency of processing for subjects with low prior knowledge and can easily be used in standard textbook material. Most importantly, they seem to lead to a more integrated processing, thus maximizing the benefit of multimodal material.

Acknowledgments

This research was funded by the *German Research Council* (Graduate Program 256, *Task Oriented Communication*).

References

Brünken, R., Steinbacher, S., Schnotz, W., Leutner, D. (2001). Mentale Modelle und Effekte der Präsentations- und Abrufkodalität beim Lernen mit Multimedia. *Zeitschrift für Pädagogische Psychologie*, 15, 16-27.

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

Carroll, P. J., Young, J. R., & Guertin, M. S. (1992). Visual analysis of cartoons. In K. Rayner (Ed.), *Eye movements and visual cognition.* New York et al.: Springer.

Guan, Y.-H. (2003). *The effects of multimedia presentations on information processing.* Doctoral dissertation, Department of Linguistics and Literature, University of Bielefeld.

Goldberg, J. H., & Wichansky, A. M. (2003). Eye tracking in usability evaluation: A practitioner's guide. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye.* Amsterdam et al.: Elsevier.

Hegarty, M., & Just, M. A. (1989). Understanding machines from text and diagrams. In H. Mandl & J. R. Levin (Eds.), *Knowledge acquisition from text and pictures*. Amsterdam et al.: Elsevier.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of memory and language*, *32*, 717-742.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn & Bacon.

Kalyuga, S., Chandler, P. & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13, 351-371.

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

Marcus, A. (1992). *Graphic design for electronic documents and user interfaces*. New York: ACM Press.

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin*, *81*, 889-917.

Mayer, R. E. (1993). Comprehension of graphics in texts: An overview. *Learning and Instruction*, *3*, 239-245.

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

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

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation. *Journal of Educational Psychology*, *84*, 444-452.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning. *Journal of Educational Psychology*, *93*, 187-198.

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

Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction*, *3*, 227-238.

Pollatsek, A., Rayner, K., & Collins, W. E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General*, *113*, 426-442.

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

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information. *Journal of Experimental Psychology*, *7*, 219-226.

Schnotz, W., & Bannert, M. (1999). Einflüsse der Visualisierungsform auf die Konstruktion mentaler Modelle beim Text- und Bildverstehen. *Zeitschrift für Experimentelle Psychologie*, *46*, 217-236.

Weidenmann, B. (1989). When good pictures fail. In H. Mandl & J. R. Levin (Eds.), *Knowledge acquisition from text and pictures*. Amsterdam et al.: Elsevier.