Learning from text and graphs: The role of annotations and sensory modality

Cengiz Acarturk [a], Christopher Habel [a], Kursat Cagiltay [b], Ozge Alacam [c]
[a] University of Hamburg, Computer Science
[b] Middle East Technical University, Computer Education and Instructional Technology
[c] Middle East Technical University, Cognitive Science
ACARTURK@INFORMATIK.UNI-HAMBURG.DE

Statistical graphs are important means for communication as well as for visual analysis of data in science and scientific education. In this paper we focus on graphs co-presented with verbal material in visual and auditory sensory modality. In particular we study the role of annotation use by investigating the learning outcomes (i.e., visual and verbal recall and transfer) as well as eye movement characteristics in the use of text-graph documents. Results of the experimental study favoured the use of annotated graphs with accompanying narrative verbal material over other constellations. Additionally, other parameters, which could affect learning, are proposed for future research.
Keywords: Text-graph documents, annotations, sensory modality, eye movements.

Introduction[1]

Research on the interaction between specific representation types (e.g. text and pictures, diagrams, graphs) and communication modes (visual, auditory etc.) is currently in a premature state due to high variability in representation types and communication modes. Compared to pictorial representations, diagrams, charts and graphs are novel means not only of data visualization but also education and communication. More specifically, statistical graphs (e.g. line graphs) provide researchers with the opportunity to visually analyze data in scientific settings. Furthermore, graphs are used in educational settings for the purpose of conveying information in concise and readily comprehensible form and reflecting conceptual relationships by highlighting their essential situations. According to Newcombe & Learmonth (2005) graphs and diagrams can be used to good effect in education and their usefulness in qualitative instruction can begin early.

Learning from multimodal documents containing text and pictorial representations (e.g., graphs, maps, diagrams, pictures, etc.) is based on the learner's ability to comprehend text as well as to comprehend these pictorial representations (Schnotz, 2005). In contrast to many other types of pictorial representations, graphs are representational artifacts, which possess internal syntactic structures. Thus, a syntactic analysis of graphs, kindred to syntactic analysis of language, is fundamental for succeeding processes of semantic and pragmatic analyses in graph comprehension (Pinker, 1990). In integrated comprehension of text and graphs both modalities have to contribute to a common conceptual representation. As described in Habel and Acarturk (2007), the interaction between information graphics and language is mediated by common conceptual representations, in particular a set of basic spatial concepts, which is fundamental for the conceptualization of graphs as well as for the conceptualization of the specific domain in which an individual graph is applied. In particular, the layer of common conceptual representations is the place where *co-reference links* among conceptual entities introduced by various modalities are constructed (Bodemer & Faust, 2006), and where *inter-* and *intra-representational coherence* is established (Seufert, 2003). In the design and construction of diagrams and graphs, several cues, such as *annotations*[2] can be used for introducing technical terms, taking attention to the key concepts (Lorch, Lorch & Klusewitz, 1995) and stressing important information (Rickards et al., 1997) for the purpose of improving inter- and intra-representational coherence. Annotations, on the one hand serve the integration of graphical elements

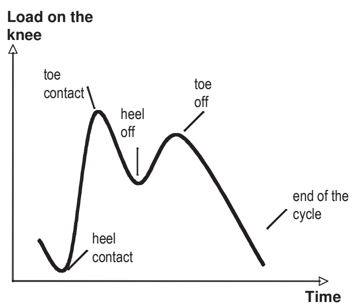
[1] The research reported in this paper has been partially supported by DFG (German Science Foundation) in ITRG 1247 'Cross-modal Interaction in Natural and Artificial Cognitive Systems' (CINACS). We thank Ozgur L. Ozcep for his support in evaluation of the results and two anonymous reviewers for their helpful comments.

[2] In this study, we use the term 'annotation' to mean the combination of textual elements (e.g. phrases) and a connecting line (i.e., 'annotation icon') which on the one end points the textual element, on the other end points a specific location on the graph.

(e.g. part of a graph line which represents an increase in the domain value) and domain entities via annotation text in a graph proper. On the other hand, annotations play the role of bridging the graph region and main text (e.g. paragraphs) of a text-graph document (Acarturk et al., in press). Our purpose in this study is to investigate the *spatial contiguity* effect (Mayer, 2001) and the *modality* effect (Tindall-Ford, Chandler & Sweller, 1997) in graph-text documents. The systematic analysis of common conceptual representations investigated in this study will be performed in a future study.

The Experiment

Thirty-six graduate and undergraduate students at the Middle East Technical University were paid to participate in the experiment. All material was presented on computer screen. Four screens that include a line graph with accompanying verbal material (Figure 1) in four different domains were presented to the subjects in random order. The design was 2x2, with two independent within-subjects parameters. The first independent parameter was *sensory modality*: the accompanying verbal material (i.e. paragraphs) was presented either visually as written text next to the graph (conditions G1 and G3) or as auditory narrative (conditions G2 and G4). The second independent parameter was presence or absence of *annotations*: the graph either included four or five annotations (conditions G1 and G2) or did not include any annotation (conditions G3 and G4). The verbal information and graphs included redundant as well as complementary information.



How does the load on the knee change during gait?

The single step of human leg, especially during lateral walking is a periodically repeated motion. For this purpose we can consider the steps as a cycle which is repeated continuously. The human gait is specified in terms of the load on a single knee in Kinematics and shown by the gait cycle graph. At the beginning of the gait cycle, together with the contact of the heel to the ground, the load starts to increase. The load decreases slightly with the toe contacting the ground. The load which increases again with the release of the heel from the ground, decreases with the release of the toe from the ground and the cycle is completed.

Condition	Modality	Annotation
G1	Text	Present
G2	Narration	Present
G3	Text	Absent
G4	Narration	Absent

Figure 1 : *Sample material from the experiment: A sample screen with the G1 condition (left); experimental conditions (right). Translated from Turkish by the first author.*

The dependent parameters were subjects' answers to post-test questions (transfer and visual/verbal recall questions) and eye movement parameters. The four domains were melting curve of a solid, working principles of a toilet tank, human gait cycle during literal walking, and working principles of a diving pressure regulator. The domain information was retrieved from journal papers, course books and online sources, and modified by the experimenter for consistency in appearance and length of the verbal material. Eye-tracking data were recorded by a 50 Hz remote eye tracker. The subjects attended the experiment in single sessions. The material was presented for a limited duration. The reading speed of the narration was set to slow reading speed. After the presentation of each screen, the subjects answered four post-test questions on paper: one visual, one verbal recall question and two transfer questions. The visual recall question asked the subjects to explain the presented material by drawing the graph on paper. The transfer questions required subjects to draw graphs for problems which require transfer of knowledge gained in the presented material to real-world situations. The verbal recall question was a multiple choice fill-in-the-blanks question, which included the paragraphs of the presented material with missing words.

Results

Answers to the post-test questions. The answers to verbal recall questions were evaluated by the correct number of answers to the recall questions. The answers to visual recall questions were evaluated by the number of correctly drawn graph lines. Among the four domains presented to the subjects, the two domains were reported as familiar by the subjects in a domain-knowledge questionnaire. Due to the potential influence of prior domain knowledge on learning (see Cook 2006 for a recent review), the two familiar domains were not included into the analysis of the post-test questions. The results (Figure 2) showed that the subjects had higher verbal recall test scores in conditions in which the annotations were present on the graph [mean (*M*) = 4.87, standard deviation (*SD*) = .29 for G1, *M* = 4.92, *SD* = .23 for G2] than the conditions in which there were no annotations on the graph [*M* = 4.45, *SD* = .62 for G3, *M* = 4.54, *SD* = .62 for G4; 2x2 ANOVA with main effect of *annotation,* $F(1, 68) = 13.12$, $p < .01$]. Furthermore, the subjects had higher visual recall test scores in conditions in which the accompanying verbal material was presented in the narrative form [*M* = 4.83, SD = .51 for G2, *M* = 4.65, *SD* = .70 for G4] than the conditions in which the verbal material was presented as written text [*M* = 4.33, *SD* = 1.46 for G1, *M* = 4.18, *SD* = 1.26 for G3; 2x2 ANOVA with main effect of *modality*, $F(1, 68) = 3.81$, $p < .05$]. The answers to transfer questions, evaluated by two researchers (Cronbach's alpha = .86), indicated no significant differences between the conditions.

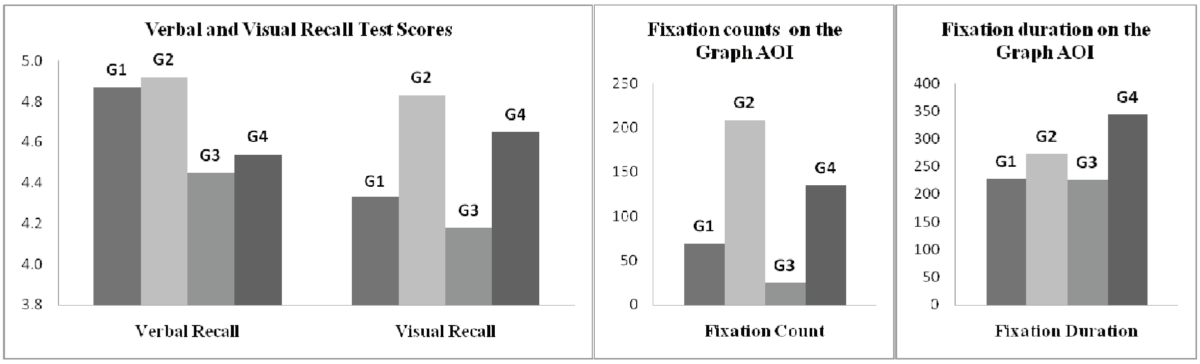


Figure 2 : *Results of the experiment: Posttest score answers and eye movement parameters.*

Eye movement parameters. For the analysis of eye movement parameters, namely fixation count and fixation duration in this study, two AOIs (Area of Interest) were specified: the *Paragraphs AOI*, which covered the paragraphs of the accompanying written text, and the *Graph AOI*, which covered the graph region, i.e. the graph proper and the annotations. The Paragraphs AOI was applicable to the conditions G1 and G3 whereas the Graph AOI was applicable to all four conditions. Results from 24 subjects were included into the analysis, due to technical calibration problems in 12 subjects. Figure 2 shows the fixation counts and fixation durations on the previously specified AOIs. The results showed that mean fixation counts on the Graph AOI were higher in the presence of annotations [*M* = 69.1, *SD* = 27.2 for G1, *M* = 208.7, *SD* = 87.6 for G2] than the ones in the absence of annotations [*M* = 24.6, *SD* = 12.5 for G3, *M* = 135.1, *SD* = 39.8 for G4] for both written text and narrative conditions, $F(3, 67) = 29.49$, $p < .01$. Furthermore, mean fixation durations on the Graph AOI were longer in the narration conditions [*M* = 272.9 ms, *SD* = 119.6 for G2, *M* = 342.8 ms, *SD* = 181.2 for G4] than the ones in the written text conditions [*M* = 226.5 ms, *SD* = 65.9 for G1, *M* = 225.2 ms, *SD* = 77.7 for G3] for both in the presence or absence of annotations, $F(3, 67) = 3.75$, $p < .05$.

Conclusions and Future Work

In this study, two aspects of learning with text and graphs were investigated: the role of annotations and the role of sensory modality. In general, our findings support the framework of theories that propose two separate, though connected, memory systems (Winn, 2004). The results of our experimental study showed that the presence of annotations supports learning, by increasing verbal recall of the material. This finding is parallel to previous research on learning from text and pictures, more specifically the *spatial contiguity effect* proposed by Mayer (2001). In addition, the presentation of the material in narrative form supports visual recall of the material, compared to the presentation of the verbal elements in the form of written text. This finding is parallel to the findings concerning the effect of sensory modality, the *modality effect* (Tindall-Ford, Chandler & Sweller, 1997; Mayer, 2001). Furthermore, longer fixation durations in the narrated conditions, specifically in the absence of annotations point to processing difficulties in the absence of annotations.

Although these findings point to an optimum configuration of multimodal documents that include annotated graphs with accompanying narrative verbal material as instructional material, the role of presence or absence of annotations and the role of sensory modality on transfer need further investigation. The non-significant effects found in answers to transfer questions might have been due to the characteristics of the functional relations between the verbal material and the graph (Ainsworth, 2006), as well as annotation text, annotation icon and graphical elements. In addition, individual differences, such as verbal/spatial abilities as well as prior domain knowledge and graph expertise are important parameters (Scheiter, Wiebe & Holsanova, submitted), which will be investigated in future studies.

References

Acarturk, C., Habel, C., Cagiltay, K., & Alacam, O. (in press). Multimodal comprehension of language and graphics: Graphs with and without annotations. To appear in *Journal of Eye Movement Research*.

Ainsworth, S. E. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16, 183-198.

Bodemer, D. & Faust, U. (2006). External and mental referencing of multiple representations. *Computers in Human Behavior, 22*. 27-42.

Habel, C., & Acarturk, C. (2007). *On reciprocal improvement in multimodal generation: Co-reference by text and information graphics.* Paper presented at the Workshop on Multimodal Output Generation (MOG 2007), Aberdeen, United Kingdom.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, 90(6), 1073-1091.

Lorch, R. F., Jr., Lorch, E. P., & Klusewitz, M. (1995). Effects of typographical cues on reading and recall of text. *Contemporary Educational Psychology, 20*, 51-64.

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

Newcombe, N. S., & Learmonth, A. (2005). Development of spatial competence. In P. Shah & A. Miyake (Eds.), *Handbook of visuospatial thinking* (pp. 213-256): Cambridge University Press.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial Intelligence and the Future of Testing* (pp. 73-126). Hillsdale, NJ: Erlbaum.

Rickards, J. P., Fajen, B. R., Sullivan, J. F., & Gillespie, G. (1997). Signaling, notetaking and field independence-dependence in text comprehension and recall. *Journal of Educational Psychology, 89*(3), 508-517.

Scheiter, K., Wiebe, E., & Holsanova, J. (submitted). Theoretical and instructional aspects of learning with visualizations. In R. Zheng (Ed.), *Cognitive effects of multimedia learning*. USA: IGI Global.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R.E. Mayer (Ed.), *Cambridge handbook of multimedia learning.* (pp. 49-69). Cambridge: Cambridge University Press.

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction, 13*. 227-237.

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied, 3*, 257-287.

Winn, W. (2004). Cognitive perspectives in psychology. In D. J. Jonassen (Ed.), *Handbook of research on educational communications and technology, 2nd ed* (pp. 79-112). Mahwah, NJ: Lawrence Erlbaum Associates.