

Surface and deep structures in graphics comprehension

Wolfgang Schnotz · Christiane Baadte

Published online: 3 December 2014
© Psychonomic Society, Inc. 2014

Abstract Comprehension of graphics can be considered as a process of schema-mediated structure mapping from external graphics on internal mental models. Two experiments were conducted to test the hypothesis that graphics possess a perceptible surface structure as well as a semantic deep structure both of which affect mental model construction. The same content was presented to different groups of learners by graphics from different perspectives with different surface structures but the same deep structure. Deep structures were complementary: major features of the learning content in one experiment became minor features in the other experiment, and vice versa. Text was held constant. Participants were asked to read, understand, and memorize the learning material. Furthermore, they were either instructed to process the material from the perspective supported by the graphic or from an alternative perspective, or they received no further instruction. After learning, they were asked to recall the learning content from different perspectives by completing graphs of different formats as accurately as possible. Learners' recall was more accurate if the format of recall was the same as the learning format which indicates surface structure influences. However, participants also showed more accurate recall when they remembered the content from a perspective emphasizing the deep structure, regardless of the graphics format presented before. This included better recall of what they had *not* seen than of what they really *had* seen before. That is, deep structure effects overrode surface effects. Depending on context conditions, stimulation of additional cognitive processing by instruction had partially positive and partially negative effects.

Keywords Comprehension of graphics · Perspectives · Structure mapping · Schemata · Surface structure of graphics · Deep structure of graphics

Visual communication with graphics plays an increasingly important role in education. In the course of growing reading literacy, students are expected to learn from texts that include graphics such as realistic pictures, diagrams, or graphs, and integrate information from these external representations into coherent knowledge structures.

Text comprehension has received much attention during the last decades. Generative linguistics suggested that sentences have a surface structure and a deep structure (Chomsky, 1965). The surface structure is the outward form of a sentence which is actually spoken and heard (or written and read). It includes phonemic or graphemic features as well as lexical and syntactic characteristics. The deep structure of a sentence is a theoretical construct which makes the underlying logical and semantic relations explicit and from which the actual form of the sentence (i.e., its surface structure) is derived. Theorists such as Fillmore (1968) and Chafe (1970) considered the deep structure a semantic construct that expresses the meaning of a sentence. Psycholinguistic studies demonstrated that texts are indeed mentally represented in a deep structure rather than a surface structure format (Sachs, 1967). Further research on text comprehension refined these concepts. It is now broadly accepted that text comprehension includes the formation of multiple mental representations. Readers are assumed to construct a mental representation of the text surface structure and a representation of its semantic deep structure. The latter, often referred to as the text base, consists of propositions representing the ideas expressed in the text. These propositions serve as a data base for constructing a mental model of the text content (cf. Graesser, Millis, & Zwaan, 1997; Kintsch, 1998; McNamara, 2007; van Dijk, & Kintsch, 1983; Weaver, Mannes, & Fletcher, 1995).

In view of the fact that far less research has been invested into comprehension of graphics (cf. Cleveland, 1985; Glenberg & Langston, 1992; Schnotz & Kulhavy, 1994), this

W. Schnotz (✉)
Faculty of Psychology, University of Koblenz-Landau, Fortstrasse 7,
76829 Landau, Germany
e-mail: schnotz@uni-landau.de

C. Baadte
Institute of Psychology, University of Mainz, Binger Str. 14-16,
Mainz 55122, Germany

paper aims to analyze how people mentally represent instructional graphics. More specifically, it addresses the following research questions:

- Is the distinction between surface and deep structures also applicable to the comprehension of graphics?
- If yes, is the distinction associated with differences in cognitive processing which result in different recall of the presented information?
- To what extent can these differences be manipulated by instruction?

We will first describe a theoretical framework of graphics comprehension. Thereafter we will report on two experiments targeting these questions.

Theory

Structural features of graphics

A graphic is a two-dimensional object with a visuo-spatial structure that represents some content based on structural commonalities between the representing graphic and the represented content. In order to understand a graphic, a learner needs to perceive it. However, this is not sufficient for understanding. Understanding requires the learner to grasp the meaning of the graphic by constructing a mental model of its content (Kosslyn, 1994; Lowe, 1996; Pinker, 1990). Accordingly, one can distinguish between a surface structure and a deep structure also with regard to graphics. Similar to the surface structure of texts, the surface structure of graphics is the outward form of the graphic, which can be actually perceived. It includes graphic elements such as dots, lines, areas, and their visual features (Bertin, 1981) as well as the spatial relations between these elements (Gentner, 1989; Schnotz, 1993). In contrast, the deep structure of a graphic is a semantic construct which expresses the meaning of the graphic. In other words, the deep structure directly represents the content of the graphic.

The difference between surface and deep structure of graphics can be illustrated by the following example. Let us assume that people's voting behavior during an election should be displayed. In most Western countries, voters primarily choose a candidate who conforms to their own political orientation. Thus, political orientation is a major influence on voting behavior. In contrast, the influence of religion is considerably weaker, although voters may also tend to vote for a candidate with the same religion. Thus, religion is a minor influence on voting behavior. Fig. 1 shows a fictitious example of voters with different political orientations (party A versus party B) and with different religions (religion x versus religion y). The three-dimensional (3D) bar chart reveals the

percentages of votes for the party B candidate (who has religion y) within voter categories Ax, Ay, Bx and By.

These data can be displayed from different perspectives. Data displayed from the party perspective results in two two-dimensional (2D) graphs. These graphs compare voters with party A affiliation and voters with party B affiliation, one graph referring to religion x voters and the other to religion y voters. These graphs are shown at the bottom right of Fig. 1. They are henceforth called "party graphs." Data display from the religion perspective results in two other 2D graphs. These graphs compare voters with religion x and voters with religion y, one graph referring to party A's affiliation and the other to party B's affiliation. The corresponding graphs are shown at the bottom left of Fig. 1. They are henceforth called "religion graphs." The party graphs and the religion graphs look different and, thus, have different surface structures. However, they convey exactly the same data and, thus, are informationally equivalent (Larkin & Simon, 1987). They can therefore be assumed to have the same deep structure, which in this case is characterized by the dominance of political orientation rather than religion with respect to voting behavior.

Perceptual and cognitive processing of graphics

We assume that similar to text comprehension, graphics comprehension includes the formation of multiple mental representations based on perceptual and cognitive processing (Kosslyn, 1994; Lowe, 1996; Sadoski & Paivio, 2001). Perceptual processing is largely pre-attentive, bottom-up, and data-driven (Neisser, 1976; Winn, 1994). Incoming visual data are organized by highly automated visual routines (Ullman, 1984) according to the gestalt laws (Wertheimer, 1938). This processing results in organized perceptual representations (i.e., visual images) which include figure-ground distinctions (e.g., dark bars on a white background) as well as spatial relations between figures (e.g., one bar being higher than another bar). Cognitive processing, on the contrary, is more attentive and concept driven. It depends on the learner's intentions as well as his/her prior knowledge and it includes both bottom-up processes and top-down processes driven by conceptual schema which comprise the individual's prior knowledge (Shah, Mayer & Hegarty, 1999). This processing results in mental models that are assumed to possess an inherent structure corresponding to the structure of the subject matter (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991). In the election example mentioned above, the mental model would represent the dominance of political orientation with regard to voting behavior. Mental models are more abstract than visual images because they can integrate information from different modalities.

Graphics and mental models have in common that they represent some content via structural commonalities. Accordingly, mental model construction via graphics comprehension

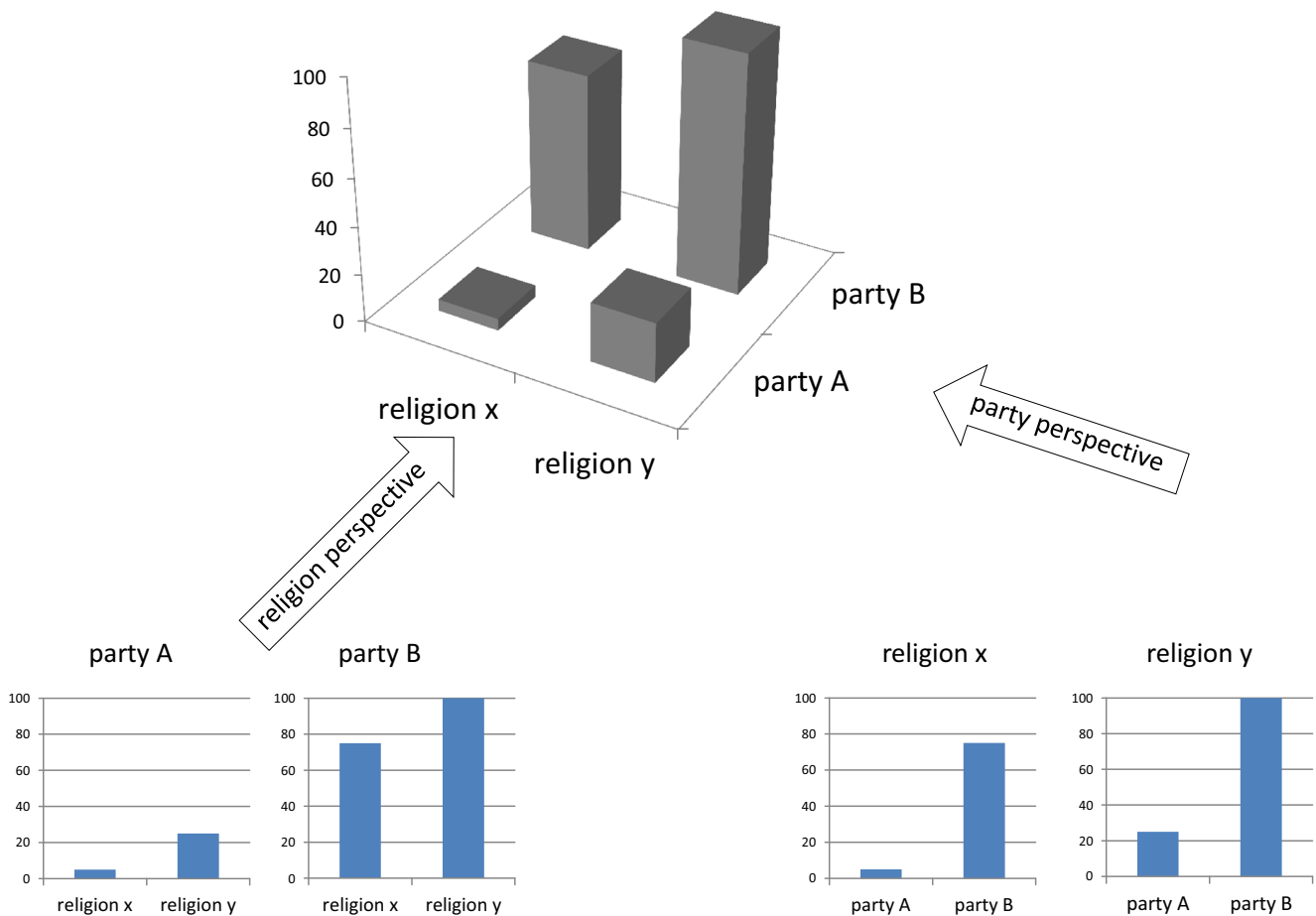


Fig. 1 Visualization of the voting behavior in an election presented from a party perspective (resulting in party graphs) or from a religion perspective (resulting in religion graphs)

is a process of structure-mapping between an external graphic and an internal mental model: graphical entities (e.g., bars) are mapped onto semantic entities (e.g., shares of votes) and spatial relations (e.g., differences between bars) are mapped onto semantic relations (e.g., differences between shares of votes) within the mental model (Falkenhainer, Forbus & Gentner, 1989/90; Gentner, 1989; Schnotz, 1993, 2014; Schnotz & Bannert, 2003).

We assume that structure-mapping is mediated by the activation of cognitive schemata. Cognitive schemata are viewed as the conceptual building blocks of cognition (Rumelhart, 1980). They are active generative cognitive units that represent typical instances and configurations within a domain based on previous experiences. Schemata are selective and organizing devices: incoming information (via bottom-up processing) is selected by schemata according to its relevance and then organized into coherent mental representations (Brewer & Nakamura, 1984). Schemata which fit better to the data will more likely be activated than others. For instance, in the elections example shown in Fig. 1, a party graph is more likely (*ceteris paribus*) to activate the party schema than the religion schema.

Schemata are based on previous experiences within a domain, but frequencies of these experiences can vary. In a specific context, a certain schema can therefore be more familiar than another one. Regarding political elections, for example, people from the Western world might be more familiar with considering the party affiliation than the religion of a candidate. One can assume that more frequently used schemata are “better trained” and, thus, can be activated more easily and require fewer cognitive resources than less frequently used schemata. In the election example, one could therefore expect that people from the Western world can activate their party schema in an election context more easily needing fewer resources than their religion schema. In another culture, however, things could be inverted if religion had the main influence and political orientation a minor influence on voting behavior.

To summarize, graphics comprehension is considered as a process of mental model construction by a schema-mediated process of structure mapping in which graphic structures are mapped onto semantically meaningful patterns of an emerging mental model (Gentner, 1989; Schnotz, 1993). The schemata activated in this mapping process act as scaffolds for mental

model construction. The mediating function of schemata in mental model construction means that they shape the emerging mental model by imposing their inherent structure on the model according to their degree of activation.

Transparency of deep structures

As the deep structure of a graphic expresses the graphic's meaning, it corresponds to the structure of the mental model to be established. However, the deep structure may become more or less transparent depending on the congruence or incongruence with the surface structure. Tversky has proposed the congruence principle for the design of instructional graphics (Tversky, Heiser, Mackenzie, Lozano & Morrison, 2008; Tversky, Morrison & Betrancourt, 2002). Accordingly, graphics should match the structure of the content and of the desired mental representation in order to be effective for learning. In terms of the present conceptual framework, the surface structure should be attuned to the underlying deep structure. If surface and deep structure are congruent, the surface structure emphasizes the major features of the content. In this case, transparency of the deep structure is high. If surface and deep structure are incongruent, the surface structure emphasizes the minor features and, thus, obscures the major ones. In this case, transparency of the deep structure is low. In the elections example shown in Fig. 1, political orientation is the major influence on voting behavior. Accordingly, party graphs are congruent with the underlying deep structure as they emphasize the major influence on voting behavior. Religion graphs, on the contrary, are incongruent as they emphasize minor influence. If religion were the major influence on voting behavior, however, religion graphs would be congruent and party graphs incongruent with the underlying deep structure.

Hypotheses regarding deep structure effects

Various studies have shown that the surface structure of graphics has an impact on the mental model construction of learners (Ainsworth & Loizou, 2003; Tversky et al., 2008). Shah, Mayer and Hegarty (1999) showed that perceptual grouping of information affects visual pattern association, which makes interpretation of graphics easier or more difficult. Keehner, Hegarty, Cohen, Khooshabeh and Montello (2008) found that participants' performance was better when they had seen objects from a task-relevant perspective than from another perspective. In other studies, graphics enhanced comprehension only if the content was visualized in a task-appropriate way. Otherwise, the graphics interfered with the construction of a task-appropriate mental model (Rasch & Schnotz, 2009; Schnotz & Bannert, 2003; Schnotz & Kürschner, 2008). Various scholars have developed design

principles for graphics as practical guidelines (Cleveland, 1985; Tufte, 1983, 1990, 1997; Wainer, 1997).

Up to now, however, no attempts have been made to our knowledge to demonstrate deep structure effects on graphics comprehension beyond surface structure effects. Differences in mental representations generally manifest themselves in different achievements with specific tasks. The fact that a mental model has been affected by the deep structure of a graphic becomes apparent if learners remember the learned information better in a format making the deep structure transparent than in another format, provided that surface structure effects have been controlled for. In the above-mentioned election example, one can hypothesize that individuals remember the learned information better in a party graph format than a religion graph format even if party graphs and religion graphs have been presented during learning with the same frequency. To put it into more general terms:

- (H1) Individuals remember the learned information in a graph format associated with the major influence better than in a format associated with the minor influence, even if both formats have been presented with the same frequency during learning.

Deep structure features of a mental model become especially apparent if learners remember the information better in a format making the deep structure transparent although they have not seen it before than in a format they actually have seen. In the above-mentioned election example, one can hypothesize that individuals remember the learned information better in a party graph format than a religion graph format even if they have actually seen religion graphs before. To put it into more general terms:

- (H2) Individuals remember the learned information in a graph format associated with the major influence better than in a format associated with the minor influence, even if they have actually seen graphs with the minor influence format before. In other words, they paradoxically remember better what they have not seen than what they actually have seen.

Hypotheses regarding instructional effects

Cognitive processing, contrary to pre-attentive perceptual processing, is influenced by the individual's intentions which can in turn be affected by instruction. As for text processing, Pichert and Anderson (1977) demonstrated that readers instructed to process text (e.g., about a house) from different perspectives (e.g., burglar's vs. broker's perspective) constructed different mental representations and had different recall. It should be noted that processing perspectives are

closely related to cognitive schemata. Displaying content from a specific perspective means that some features are brought to the foreground whereas others are moved to the rear. Likewise, a cognitive schema emphasizes specific features and ignores others. A burglar's schema, for example, will emphasize valuables in a house but ignore the quality of the roof, whereas a broker's schema will do the opposite.

The question arises whether instructions to process information from specific perspectives will have similar effects in learning from graphics. In the above-mentioned elections example, an instruction to focus on voters with different political orientations should activate the party schema leading to a mental model construction from the party perspective. Conversely, an instruction to focus on voters with different religions should activate the religion schema leading to a mental model construction from the religion perspective. However, activation of schema might be moderated by additional influences. First, schemata associated with a more familiar perspective might be activated more easily by instruction than other schemata. Second, because the cognitive capacity for mental model construction is limited (Gyselinck, Meneghetti, De Beni & Pazzaglia, 2009), one can assume that a schema will be activated at the expense of the activation of another competing schema. This leads to a trade-off between schema activations due to interference. Third, as schemata associated with less familiar perspectives might require more mental capacity (i.e., impose a higher cognitive load on working memory; Chandler & Sweller, 1991), they might cause higher interference than schemata with more familiar perspectives. Fourth, a schema will only be activated by a corresponding instruction if it has not been activated yet by other influences. Fifth, a schema can only be inhibited by interference from another schema if it has been activated by other influences before. Taking these moderating effects into account, one can make the following predictions:

- (H3) A cognitive schema will be inhibited by incongruent instruction if it has been already activated and if the perspective suggested by the instruction is unfamiliar to the learner.
- (H4) A cognitive schema will be activated by a congruent instruction if it has not been activated yet and if the schema corresponds to a familiar, frequently used perspective.

In order to test these hypotheses about deep structure mappings and about instructional effects in graphics comprehension, two experiments were conducted. The experiments were complementary because features of major importance in Experiment 1 became minor in Experiment 2, and vice versa.

Experiment 1

Participants

One hundred and fifty-seven students (116 females) from different faculties of a university in Germany participated in this study. Their average age was 23.8 years ($SD = 5.0$). Participants were paid 10 Euros for participation. They were randomly assigned to six different treatment groups to receive different learning material with different instructions. The groups did not differ significantly in terms of their proportion of gender, $\chi^2(5) = 7.21, p = .21$. They did also not differ significantly in terms of age, $F(5,150) = 1.306, p = .26, \eta^2 = .04$.

Learning material

Learning content was chosen with a view that its structure should be simple and participants should have no prior knowledge about it. Students were asked to learn about the voting behaviors of voters with different political orientations and religions in the US presidential elections in 1956 and 1960 (adapted from Tufte, 1983). One half of the students received two party graphs for 1956 and two party graphs for 1960 as shown in Fig. 2. For each year, one graph referred to Catholic and the other to Protestant voters, each comparing the voting percentages of Republican voters and Democratic voters for the Democratic candidate. The other half of the participants received two religion graphs for 1956 and two religion graphs for 1960 as shown in Fig. 3. For each year, one graph referred to Republican voters and the other to Democratic voters, each comparing the voting percentages of Catholic voters and Protestant voters for the Democratic candidate. For each year, the party graphs and the religion graphs were informationally equivalent as they conveyed exactly the same data. As explanatory background information preceding the graphs, all participants received the following text (which included 168 words in the German version):

In the presidential elections in the United States of America two factors are especially important for a person's vote: (1) his/her party affiliation, and (2) his/her agreement between his/her own religion and the religion of the candidate. This can be illustrated with regard to the elections in 1956 and 1960. Voters with an affiliation to the Democrats preferred the Democratic candidate to the Republican candidate. Conversely, voters with an affiliation to the Republicans preferred the Republican candidate to the Democratic candidate. Furthermore, a Catholic voter will show a preference for a Catholic candidate if the other candidate is a Protestant. Conversely, a Protestant voter will prefer a Protestant candidate if the other candidate is a Catholic.

The difference between voters' behavior in 1956 and in 1960 can be explained by the fact that in 1956 the Democrat candidate was a Protestant and the Republican candidate was a Catholic, whereas in 1960 both candidates were Catholics. The corresponding voting percentages are shown in the following graphics. (Adapted from Tufte, 1983).

The two groups of participants were further subdivided into three subgroups which received different instructions in order to trigger different kinds of processing. The first subgroup received no instruction. The second subgroup received a party instruction (congruent to party graphs and incongruent to religion graphs):

How much did voters with a Democrat affiliation differ from those with a Republican affiliation in their preference for the Democrat candidate?

A third subgroup received a religion instruction (congruent to religion graphs and incongruent to party graphs):

How much did Protestant voters differ from Catholic voters in their preference for the Democrat candidate?

These questions were expected to be applied to each of the four presented graphs.

Procedure

Pretest phase Because graphics-based mental model construction from different perspectives might require spatial skills, participants were tested for their spatial cognitive abilities with the 3D cube test of Gittler (1990) as a control variable. The treatment groups did not differ significantly in terms of their spatial abilities, $F(5,151) = 1.713, p = .14, \eta^2 = .05$. Furthermore, participants were interviewed regarding their prior knowledge about the US Presidential Elections in 1956 and 1960. None of them reported any knowledge about these elections.

Learning phase All participants received the 168-word text presented above combined either with party graphs or religion graphs – either with no instruction, party instruction, or religion instruction. Accordingly, the study followed a 2×3 design with the between factors *graph* (party graphs/religion graphs) and *instruction* (no instruction, party instruction, or religion instruction). The participants were requested to understand and memorize the learning material. They were not allowed to take notes but asked to concentrate on

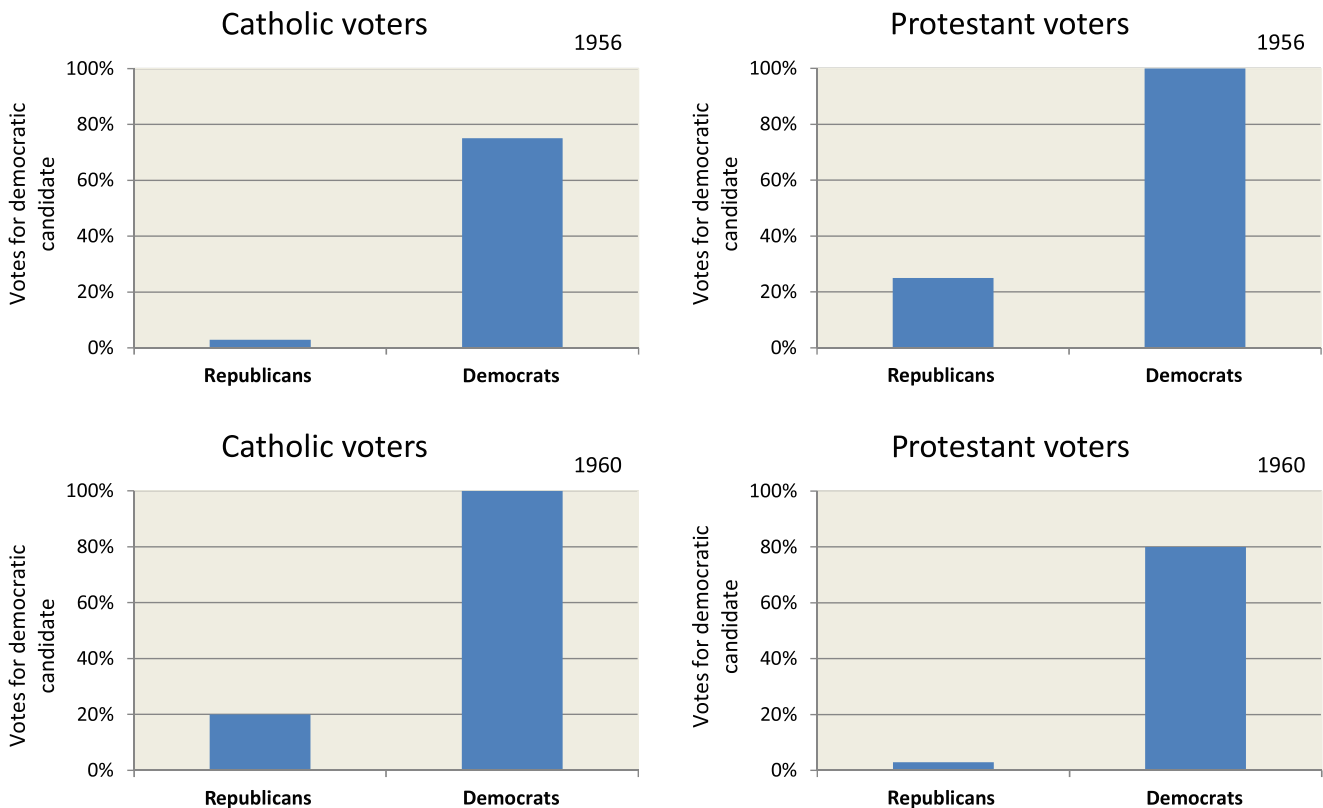


Fig. 2 Party graphs concerning the US Presidential Elections in 1956 and 1960 in Experiment 1 (In Experiment 2, these party graphs became religion graphs: Catholic voters and Protestant voters were replaced by

Progressive voters and Conservative voters, respectively. Republicans and Democrats were replaced by Sunnites and Shiites, respectively)

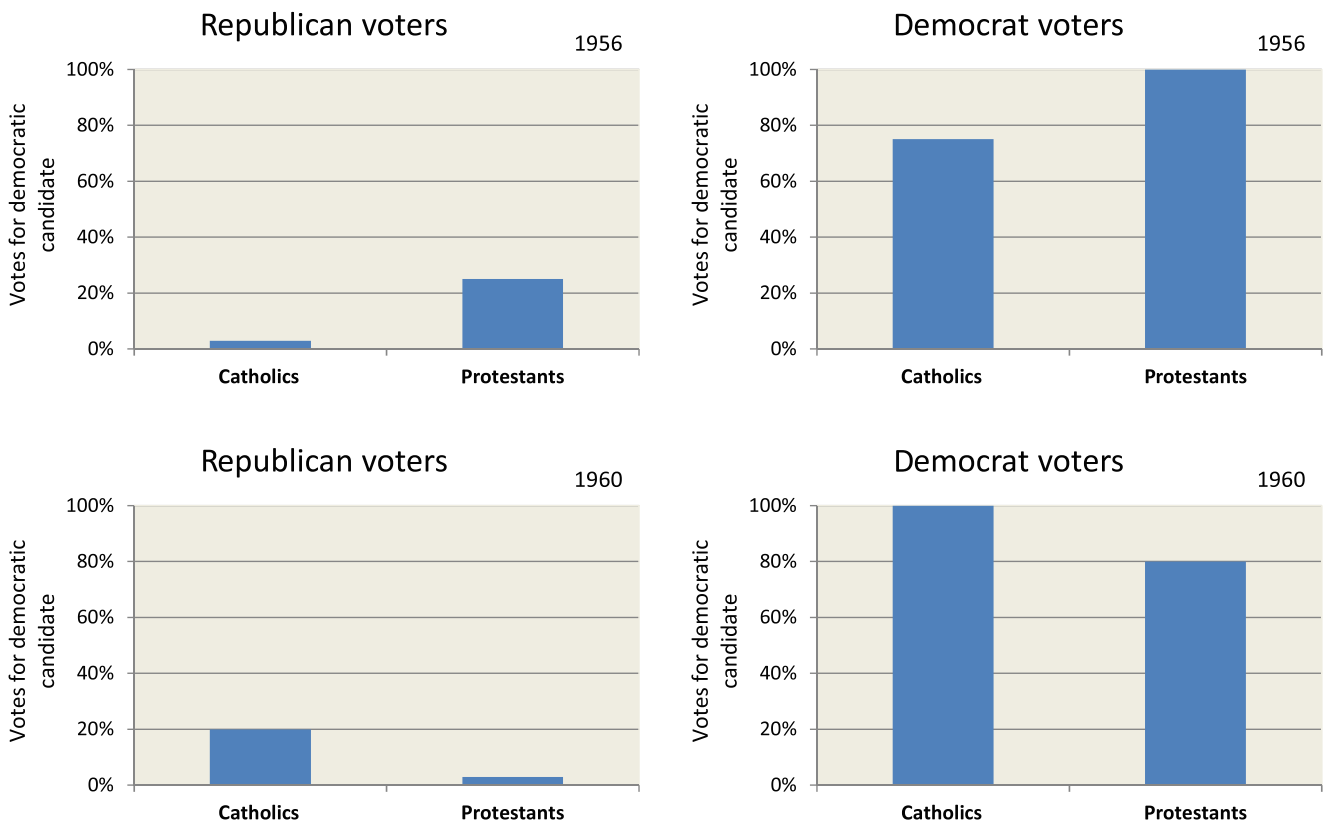


Fig. 3 Religion graphs concerning the US Presidential Elections in 1956 and 1960 in Experiment 1 (In Experiment 2, these religion graphs became party graphs: Catholic voters and Protestant voters were replaced by

Progressive voters and Conservative voters, respectively. Republicans and Democrats were replaced by Sunnites and Shiites, respectively)

the presented subject matter and to respond to any additional instruction (party or religion) only mentally. Although the text was relatively short, a pilot study had shown that learners need considerable time to make sense of the four graphs and the explanatory text, that is, to understand the overall situation including the specifics of the election years. Based on this experience, the maximum learning time was set at 30 minutes. Participants were free to terminate the learning phase whenever they wanted and to move on to the post-test.

Post-test phase Participants were tested for their knowledge about the learning content immediately after learning. They received eight blank graphs including only captions and labels. Four graphs called “party items” had a party format. The other four graphs called “religion items” had a religion format. Participants had to draw the missing bars into the eight graphs as accurately as possible. Due to the two kinds of items, participants had to recall the content from two perspectives, one being congruent and the other being incongruent with the previously seen graph format. There were no time constraints in the post-test phase.

Scoring For each participant, accuracy of recall was determined for the party items and for the religion items in the following way. For each item, the effect denoted by the participant (i.e., the difference between the two columns filled in by the learner) was compared to the correct effect (the correct difference between the two columns). Over- or underestimations of the effect were considered as inaccuracies. The absolute values of inaccuracies (no matter whether there was an over- or an underestimation) were added up across the party items resulting in a party items sum of inaccuracies. The same was done for the religion items leading to a religion items sum of inaccuracies. In order to get scores of accuracy rather than of inaccuracy, the sum of party item inaccuracies was subtracted from 100, which led to a score of party items accuracy. The corresponding procedure for the religion items led to a score of religion items accuracy. An analysis of internal consistency revealed a Cronbach's alpha of .71 for party items accuracy and .67 for religion items accuracy.

Results

Participants invested on average 14.6 minutes ($SD = 3.13$) for learning. The treatment groups did not differ significantly with regard to learning times, $F(5,151) = 1.066, p = .38, \eta^2 = .03$.

According to hypothesis H1, participants in Experiment 1 were expected to remember the learned information on the average better with party items than with religion items. According to hypothesis H2, they were expected to remember the information with party items better even if they had seen religion graphs in the learning phase.

Means and standard deviations of participants' recall accuracies with party items and with religion items are shown in Table 1. A mixed-design $2 \times 3(\times 2)$ ANOVA with the between factors *graph* (party graphs/religion graphs) and *instruction* (no instruction/party instruction/religion instruction) and the within factor *item* (party items/religion items) led to the results presented in Table 2. A significant main effect was found for factor *item*: as predicted by hypothesis H1, participants showed more accurate recall with party items ($M = 47.9$; $SD = 55.2$) than with religion items ($M = 22.6$; $SD = 46.2$). Regarding hypothesis H2, there was a tendency in the expected direction: party items were answered more accurately ($M = 38.5$; $SD = 62.7$) than religion items ($M = 25.6$; $SD = 37.9$) even when participants had seen religion graphs in the learning phase. However, this difference did not reach the 5 % level of significance and can at the most be considered as marginally significant, $t(78) = 1.826$; $p = .072$; $d = 0.25$. The significant interaction *graph* \times *item* in Table 2 was due to the fact that participants who had seen party graphs during learning showed more accurate recall with party items ($M = 57.4$; $SD = 44.8$) than those who had seen religion graphs ($M = 38.5$; $SD = 62.7$), whereas participants who had seen religion graphs showed more accurate recall with religion items ($M = 25.6$; $SD = 37.4$) than those who had seen party graphs ($M = 19.7$; $SD = 53.4$). This interaction mirrors the uncontroversial fact that the graphics' surface structure is memorized at the perceptual or cognitive level.

The significant interaction *instruction* \times *item* in Table 2 gives a general hint to differential effects of instruction on item-specific recall. Regarding these kinds of effects, Experiment 1 allowed for a test of hypothesis H3 which predicted that a cognitive schema will be inhibited by incongruent instruction if it has already been activated and if the perspective suggested by the instruction is unfamiliar to the learner. Thus, an already activated party schema should be inhibited by a religion instruction because the religion schema requires many cognitive resources due to its unfamiliarity. The party schema might have been activated by the surface structure of party graphs or by the graphs' deep structure (which means, even when religion graphs were seen). In fact, learners who had studied party graphs with a religion instruction were significantly less accurate in answering party items ($M = 36.9$; $SD = 59.5$) than learners who had studied party graphs without additional instruction ($M = 66.9$; $SD = 35.9$), $t(151) = 2.033$; $p = .022$; $d = 0.54$). Similarly, learners who had studied religion graphs with a religion instruction were significantly less accurate in answering party items ($M = 18.9$; $SD =$

Table 1 Recall accuracies with party items and religion items in Experiment 1

		Experiment 1 (USA)		
		Accuracy scores		
		<i>M</i>	<i>SD</i>	<i>n</i>
Party graphs				
Total				
	Party items	57.44	44.81	78
	Religion items	19.67	53.37	78
No instruction				
	Party items	66.92	35.89	26
	Religion items	28.08	43.73	26
Party instruction				
	Party items	68.46	26.37	26
	Religion items	9.23	66.69	26
Religion instruction				
	Party items	36.92	59.53	26
	Religion items	21.71	47.14	26
Religion graphs				
Total				
	Party items	38.54	62.66	79
	Religion items	25.57	37.88	79
No instruction				
	Party items	47.69	41.53	26
	Religion items	26.15	41.38	26
Party instruction				
	Party items	51.00	53.23	25
	Religion items	25.60	38.63	25
Religion instruction				
	Party items	18.93	81.26	28
	Religion items	25.00	35.12	28

81.3) than learners who had studied religion graphs without additional instruction ($M = 47.7$; $SD = 41.5$),

Table 2 Results of the $2 \times 3(\times 2)$ ANOVA with the between-factors *graph* (party graphs/religion graphs) and *instruction* (no instruction/party instruction/religion instruction) and the within-factor *item* (party items/religion items) in Experiment 1

Factor	<i>F</i>	<i>p</i>	η^2
Graph ^a	0.918	.34	.01
Instruction ^b	2.487	.09	.03
Item ^a	29.54	<.001	.16
Graph \times instruction ^b	0.208	.81	.00
Graph \times item ^a	6.519	.006	.04
Instruction \times item ^b	5.591	.005	.07
Graph \times instruction \times Item ^b	0.273	.076	.00

^a $F(1,151)$; ^b $F(2,151)$

$t(151) = 1.985$; $p = .025$; $d = 0.52$). Accordingly, hypothesis H3 was supported by the results.

Discussion

As a whole, the results provided limited support for the assumption that the graphic's deep structure is mapped on the structure of the emerging mental model. In this experiment, the deep structure reflected the fact that the political orientation had a major influence and religion a minor influence on voting behavior. As predicted by hypothesis H1, learners recalled the content with party items significantly more accurately than with religion items, although the required item responses were informationally equivalent and although both kinds of graphs were presented during learning with the same frequency. Hypothesis H2 predicted that even participants who had seen religion graphs in the learning phase would have more accurate recall with party items than with religion items. This effect was not statistically significant. There was only a difference in the expected direction which could at most be considered as marginally significant. Furthermore, the results confirmed the noncontroversial assumption that the surface structure of graphics is also mentally presented (Ainsworth & Loizou, 2003; Glenberg & Langston, 1992; Schnotz & Bannert, 2003; Schnotz & Kürschner, 2008; Tversky et al., 2008). Participants who had learned with party graphs showed on average more accurate recall with party items than with religion items. Those who had learned with religion graphs showed on the average more accurate recall with religion items than with party items.

Given the limited support for the abovementioned hypotheses, the findings should at this point be interpreted very carefully. They can be considered as a tentative hint that the deep structure of graphics is mentally represented due to a process of structure mapping. This is noteworthy insofar as recall took place in this study immediately after learning, which should favor mental representations of the surface structure compared to mental representations of the deep structure. From a broader point of view, the findings suggest that graphics comprehension is not a mechanical mapping of surface features onto a mental model but a process of active sense-making (Mayer, 2009; Wittrock, 1989). Learners engage in active cognitive processing to construct coherent knowledge structures in the form of an appropriate mental model based on information including the graphic's surface structure as well as its deep structure.

Regarding instructional effects, the data supported the assumption that a cognitive schema will be inhibited by an incongruent instruction if it has already been activated and if the perspective suggested by the instruction is unfamiliar to the learner (hypothesis H3). In the present study, the party schema might have been activated anyway by the graphics' deep structure, and the religion perspective might have been

unfamiliar in the context of elections. Accordingly, learners who were instructed to adopt the unfamiliar religion perspective showed significantly lower recall accuracy with party items than learners without instruction. It seems that the unfamiliar religion perspective interfered with the activation of the party schema. Additional cognitive activation does obviously not necessarily result in better learning (cf. Spiro, Feltovich, Jacobson & Coulson, 1991). Further moderating factors have to be taken into account in order to predict instructional effects. Such factors might include the current amount of schema pre-activation. It might also include the familiarity of the corresponding perspective, which implies that some schemata can be activated more easily than others. There is obviously a possibility that activation of one schema inhibits the activation of another schema. A schema representing a less familiar perspective might cause more interference as it requires higher cognitive effort than a schema with a more familiar perspective (cf. Chandler & Sweller, 1991). These topics need further investigation.

The participants of Experiment 1 came from a Western culture. In the context of elections, a party perspective might have been more familiar to them than a religion perspective. The learning content of this study also implied that the political orientation was of major influence on voting behavior. In order to control for this confounding factor, a second experiment was conducted in which the graphics' deep structure was manipulated in a way that the roles of the religion perspective and the party perspective were reversed.

Experiment 2

Participants

One hundred and thirty-four students (96 women) from different faculties of a university in Germany participated in this study. Their average age was 23.0 years ($SD = 4.45$). Participants were randomly assigned to six treatment groups. The groups did not differ either in terms of gender proportion or age significantly; $\chi^2(5) = 3.18$, $p = .67$ and $F(5, 128) = 0.419$, $p = .84$, $\eta^2 = .02$, respectively. Students were paid 10 Euros for participation.

Learning material

Participants were required to learn about voting behavior during the elections in 1956 and in 1960 in a fictitious foreign country, the Republic of Ustan, where the religious orientation (Schiites vs. Sunnites) is the most important influence on voting behavior, whereas the party affiliation (conservatives versus progressives) has only secondary influence. One half of the participants received two religion graphs for 1956 and two

religion graphs for 1960. These graphs were identical with those shown in Fig. 2, except that Republicans and Democrats were replaced by Sunnites and Shiites, and Catholic voters and Protestant voters were replaced by Progressive voters and Conservative voters, respectively. So, the party graphs of Experiment 1 became religion graphs in Experiment 2. The other half of the participants received two party graphs for 1956 and two party graphs for 1960. These graphs were identical with those in Fig. 3, except for the abovementioned replacements. So, the religion graphs of Experiment 1 became party graphs in Experiment 2. All participants received the following text (which included 171 words in the German version) preceding the graphs as background information:

In the presidential elections in the Republic of Ustan, two factors are especially important for a person's vote: (1) his/her agreement between his/her own religion and the religion of the candidate, and (2) his/her party affiliation. This can be illustrated with regard to the elections in 1956 and 1960. Shiite voters preferred the Shiite candidate to the Sunnite candidate. Conversely, Sunnite voters preferred the Sunnite candidate to the Shiite candidate. Furthermore, a voter with an affiliation to the Progressive party shows a preference for the Progressive candidate, if the other candidate is Conservative. Conversely, a voter with an affiliation to the Conservative party prefers the Conservative candidate, if the other candidate is Progressive. The difference between voters' behavior in 1956 and in 1960 can be explained by the fact that in 1956 a conservative Shiite candidate competed with a Progressive sunnite candidate, whereas in 1960 the Shiite candidate and the Sunnite candidate were both members of the Progressive party. The corresponding voting percentages are shown in the following graphics.

Just as in Experiment 1, both groups were further subdivided into three subgroups. The first subgroup received no instruction, the second subgroup received a religion instruction, and the third subgroup received a party instruction.

Procedure and scoring

The procedure of Experiment 2 was exactly the same as in Experiment 1, except that participants were not asked about their prior knowledge, because the content was fictitious. Scoring was also done in exactly the same way as in the previous study.

Results

Participants invested on average 18.4 minutes ($SD = 5.08$) for learning. The treatment groups did not differ significantly with

regard to learning times, $F(5,128) = 0.244$, $p = .95$, $\eta^2 = .01$. According to hypothesis H1, learners in Experiment 2 were expected to remember the information on average better with religion items than with party items. According to hypothesis H2, they were expected to remember the information with religion items better even if they had seen party graphs in the learning phase.

Means and standard deviations of participants' recall accuracies with party items and with religion items are shown in Table 3. A mixed-design $2 \times 3(\times 2)$ ANOVA with the between factors *graph* (party graphs/religion graphs) and *instruction* (no instruction/party instruction/religion instruction) and the within factor *item* (party items/religion items) led to the results presented in Table 4. A significant main effect was found for factor *item*: as expected according to hypothesis H1, participants showed more accurate recall with religion items ($M = 62.3$; $SD = 47.7$) than with party items ($M = 32.6$; $SD = 35.5$).

Table 3 Recall accuracies with party items and religion items in Experiment 2

		Experiment 2 (Ustan)		
		Accuracy scores		
		<i>M</i>	<i>SD</i>	<i>n</i>
Party graphs				
Total				
	Party items	37.93	24.53	67
	Religion items	58.58	52.21	67
No instruction				
	Party items	41.30	24.18	23
	Religion items	56.30	58.95	23
Party instruction				
	Party Items	40.00	24.88	22
	Religion items	63.64	45.18	22
Religion instruction				
	Party items	32.32	24.69	22
	Religion items	55.91	53.36	22
Religion graphs				
Total				
	Party items	27.31	43.33	67
	Religion items	65.97	42.89	67
No instruction				
	Party items	25.22	41.76	23
	Religion items	69.35	30.87	23
Party instruction				
	Party items	40.00	20.24	22
	Religion items	56.14	60.32	22
Religion instruction				
	Party items	16.82	58.10	22
	Religion items	72.27	31.46	22

Table 4 Results of the $2 \times 3(\times 2)$ ANOVA with between-factors graph (party graphs/religion graphs) and instruction (no instruction/party instruction/religion instruction) and the within-factor item (party items/religion items) in Experiment 2

Factor	<i>F</i>	<i>p</i>	η^2
Graph ^a	0.076	.78	.00
Instruction ^b	0.315	.73	.01
Item ^a	46.645	<.001	.27
Graph \times iInstruction ^b	0.042	.96	.00
Graph \times item ^a	4.215	.021	.03
Instruction \times item ^b	1.679	.19	.03
Graph \times instruction \times item ^b	2.116	.13	.03

^a $F(1,128)$; ^b $F(2,128)$

As predicted by hypothesis H2, recall accuracy was significantly higher with religion items ($M = 58.6$; $SD = 52.2$) than with party items ($M = 37.9$; $SD = 24.5$) even when participants had seen party graphs in the learning phase, $t(66) = 3.83$; $p < .001$; $d = 0.51$. In other words, these students paradoxically showed better recall of what they had *not* seen than of what they really *had* seen before.

The significant interaction *graph* \times *item* in Table 2 mirrors the fact that participants who had seen religion graphs showed on average more accurate recall with religion items ($M = 66.0$; $SD = 42.9$) than participants who had seen party graphs ($M = 58.6$; $SD = 52.2$), whereas participants who had seen party graphs showed more accurate recall with party items ($M = 37.9$; $SD = 24.5$) than participants who had seen religion graphs ($M = 27.3$; $SD = 43.3$). Once again, this interaction suggests that besides graphics' deep structures, their surface structures are also mentally represented.

Experiment 2 allowed for a test of hypothesis H4 which predicted that a cognitive schema will be activated by a congruent instruction if it has not been activated yet and if the schema corresponds to a familiar, frequently used perspective. In the present study, where religion had a major effect on voting, religion graph learners had no reason to activate their party schema. Under these conditions, a party instruction should be effective because the party perspective is familiar to the learners. In fact, religion graph learners with party instruction showed more accurate party items recall ($M = 40.0$; $SD = 20.2$) than religion graph learners without instruction ($M = 25.2$; $SD = 41.8$). However, this difference did not reach the 5 % level of significance and can despite a moderate effect strength only be considered as marginally significant, $t(32.1) = 1.521$; $p = .069$; $d = 0.42$).

Discussion

In this experiment, the roles of religion and party affiliation were reversed compared to the previous Experiment 1:

religion became the major effect, whereas party affiliation was of minor importance for voting behavior. Although the participants still came from a Western culture, where a party perspective might be more familiar in the context of elections than a religion perspective, the result pattern changed fundamentally. As predicted by hypothesis H1, learners recalled the content with religion items significantly more accurately than with party items, although the required item responses were informationally equivalent and although both kinds of graphs were presented with the same frequency during learning. Hypothesis H2 had predicted that even participants who had seen only party graphs in the learning phase would nevertheless have more accurate recall with religion items than with party items. This hypothesis was also confirmed by a highly significant effect. That is, students showed better recall of what they had *not* seen than of what they really *had* seen before. Furthermore, the findings showed once again that besides their deep structure, the surface structure of graphics affects the mental representation too (cf. Ainsworth & Loizou, 2003; Schnotz & Bannert, 2003; Schnotz & Kürschner, 2008; Tversky et al., 2008).

As a whole, the findings support the assumption that during graphics comprehension, the graphic's deep structure affects the structure of the emerging mental model. At a general level, the findings suggest once again that graphics comprehension is an active process of sense-making based on surface and deep structure information rather than a mechanical mapping of surface features onto a mental model (Mayer, 2009; Wittrock, 1989).

A tendency was found only according to hypothesis H4, which had predicted that religion graph learners who were instructed to adopt the (highly familiar) party perspective would show more accurate recall with party items than those who were not. There was only a marginally significant effect which should be interpreted very cautiously. As there was no significant main effect of instruction, one can once again suspect that additional cognitive activation *per se* does not generally result in better learning. However, it would be premature to draw definite conclusions on the basis of the current data regarding the moderation of instructional effects by schema activation, familiarity and competing schemata. Further research is needed on this issue.

General discussion

After decades of intensive text processing research, comprehension of graphics receives increased interest in multimedia research. Graphics have usually been viewed as a complement to texts: they provide an additional code (Paivio, 1986; Sadoski & Paivio, 2001), allow elaborate conjoint processing (Kulhavy, Stock & Caterino, 1994), or enable the construction

of an additional mental model (Mayer, 2009). The structural aspect of graphics, however, received little attention so far (cf. Glenberg & Langston, 1992).

In this article, we have argued that graphics have a perceptual surface structure and a semantic deep structure. When different graphics convey the same information in different ways they look different and have therefore different surface structures. Because they are informationally equivalent (Larkin & Simon, 1987), however, they possess the same deep structure. Graphics comprehension is considered a process of structure mapping from a graphic onto a corresponding mental model based on surface structure as well as deep structure characteristics (cf. Falkenhainer et al., 1989, 1990; Gentner, 1989; Schnotz, 2014; Schnotz & Bannert, 2003). The mapping process is assumed to take place under the guidance of cognitive schemata as building blocks of cognition (Brewer & Nakamura, 1984; Rumelhart, 1980). Activated schemata serve as scaffolds for mental model construction (cf. Eitel, Scheiter & Schüler, 2012).

The idea that graphics comprehension implies perceptual surface structure and semantic deep structure mappings is in line with the research by Knauff and Johnson-Laird (2002) who showed that mental models can differ from visual images and that different brain areas are involved in creating visual images and spatially organized mental models (Knauff, Fangmeier, Ruff & Johnson-Laird, 2003; Knauff, Mulack, Kassubek, Saligh & Greenlee, 2002). Evidence for a distinction between visual and spatial components in processing of verbal and pictorial information was also found by Gyselinck and her colleagues (Gyselinck, Ehrlich, Cornoldi, de Beni and Dubois, 2000; Gyselinck, Cornoldi, Ehrlich, Dubois, & de Beni, 2002).

Two experiments were conducted in order to test the assumptions about deep structure mappings in graphics comprehension. Surface structure as well as deep structure characteristics were systematically varied across the two studies. The cultural background of participants was kept constant. It could therefore not account for different results between the experimental conditions. Both studies confirmed that the structure of the emerging mental model during graphics comprehension is affected not only by the graphic's surface structure but also by its deep structure. In both studies, participants recalled the learned information significantly more accurately if the item format was congruent (rather than incongruent) with the deep structure, although the required item responses were informationally equivalent and although the different kinds of graphs were presented with the same frequency during learning. If surface structure and deep structure were incongruent, participants had nevertheless more accurate recall with item formats congruent to the deep structure than item formats congruent to the surface structure. In this case, students showed paradoxically better recall of what they had not seen than of what they really had seen before during

learning. This effect was only marginally significant in the first experiment, but highly significant in the second experiment. It should be kept in mind that the participants had to recall the information immediately after learning, which should favor recall of the surface rather than the deep structure.

These findings rule out the idea of graphics comprehension as a mechanical one-on-one mapping of surface graphical features onto features of the mental model. Instead, it corroborates the view that humans are active sense makers (Mayer, 2009; Wittrock, 1989) who engage in constructing coherent mental representations from the available information. Accordingly, graphics comprehension is an adaptive process of mapping perceptual surface structures as well as semantic deep structures onto an emerging mental model in which learners can compensate to some extent for inappropriate visualization formats.

In both studies, some participants were instructed to answer a question from a specific perspective in order to trigger additional cognitive processing. The findings did not support the view that this leads generally to a more elaborated mental representation. Asking questions was admittedly only a weak instructional manipulation. Nevertheless, it seems that the activation of cognitive schemata which guide the process of mental model construction is affected by multiple factors that have to be taken into account. The activation of a schema by instruction can interfere with the activation of another schema, whereby the amount of interference would depend on the cognitive load imposed by the interfering schema (Chandler & Sweller, 1991). Schemata representing a more familiar perspective might be easier to activate, whereas less familiar perspectives might impose a higher cognitive load on working memory resulting in stronger interference effects. These topics deserve further investigation.

The two experiments varied both the perceptual surface structure and the semantic deep structure of graphics, but confined themselves to one graph format: bar graphs. Accordingly, future investigations should also deal with other types of graphics. Furthermore, learners' prior knowledge was not systematically varied in the two experiments. This might also be an important issue for further research, because learners do not necessarily possess all relevant cognitive schemata for understanding graphs (Pinker, 1990).

As for the practical implications, the above findings corroborate the importance of adequate design of graphics in terms of congruence between semantic content, visualization format, and intended usage of the emerging mental representation (Tversky et al., 2008). It is not sufficient to deliver correct information via graphics. It is also important to choose an appropriate perceptual format for the display of information corresponding to a perspective that makes the intended semantic deep structure as transparent as possible (cf. Ainsworth & Loizou, 2003; Rasch & Schnotz, 2009; Schnotz

& Bannert, 2003). In other words, one has to avoid that perceptually salient but thematically irrelevant features dominate cognition via perception (Lowe, 1996).

The above findings suggest that enhancing graphics comprehension by visual design and learners' cognitive activation induced by instruction is not a matter of simple rules of thumb. It seems to be a matter of complex interactions between perceptual surface structures, semantic deep structures, perspectives of different familiarity, cognitive schemata associated with these perspectives, and interference between schemata, whereby interference depends on the cognitive load imposed by the interfering schema. All these interactions seem to co-determine the process of construction of mental models in graphics comprehension. A deeper understanding of these processes will improve our chances to develop adequate guidelines of using graphics for visual knowledge communication.

Acknowledgments We are grateful to Christian Kürschner for his conceptual input and his support for both experiments. We also thank Tina Seufert for her help in preparing Experiment 1 and Alexander Frey, Niccola Soccodato, and Angelika Sommer for their support in running the experiments. Furthermore, we want to thank Amy Johnson for her critical remarks and suggestions based on a former version of this paper.

References

- Ainsworth, S., & Loizou. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science*, 27, 669–681.
- Bertin, J. (1981). *Graphics and graphic-information-processing*. Berlin/New York: Walter de Gruyter.
- Brewer, W. F., & Nakamura, G. V. (1984). The nature and functions of schemas. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 119–160). Hillsdale, N.J.: Erlbaum.
- Chafe, W. L. (1970). *Meaning and the structure of language*. Chicago: University of Chicago Press.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293–332.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge MA: MIT Press.
- Cleveland, W. S. (1985). *The elements of graphing data*. New York: Chapman and Hall.
- Eitel, A., Scheiter, K., & Schüler, A. (2012). The time course of information extraction from instructional diagrams. *Perceptual and Motor Skills: Learning & Memory*, 115(3), 677–701.
- Falkenhainer, B., Forbus, K.D., & Gentner, D. (1989/90). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, 41, 1–63.
- Fillmore, C. J. (1968). The case for case. In E. Bach & R. Harms (Eds.), *Universals in linguistic theory* (pp. 1–88). New York: Holt, Rinehart & Winston.
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 197–241). Cambridge: Cambridge University Press.
- Gittler, G. (1990). *Dreidimensionaler Würfeltest (3DW)*. Ein Rasch-skaliertes Test zur Messung des räumlichen Vorstellungsvermögens. [3D cube test. A Rasch-scaled test for measuring spatial imaginative ability.] Weinheim: Beltz Test GmbH.
- Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language*, 31, 129–151.
- Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, 48, 163–189.
- Gyselinck, V., Cornoldi, C., Ehrlich, M.-F., Dubois, V., & de Beni, R. (2002). Visuospatial memory and phonological loop in processing texts and illustrations. *Applied Cognitive Psychology*, 16, 665–685.
- Gyselinck, V., Ehrlich, M.-F., Cornoldi, C., de Beni, R., & Dubois, V. (2000). Visuospatial working memory in learning from multimedia systems. *Journal of Computer Assisted Learning*, 16, 166–176.
- Gyselinck, V., Meneghetti, C., De Beni, R., & Pazzaglia, F. (2009). The role of working memory in spatial texts processing: What benefit of imagery strategy and visuospatial abilities? *Learning and Individual Differences*, 19, 12–20.
- Johnson-Laird, P. N. (1983). *Mental Models. Towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.
- Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hillsdale, NJ: Lawrence Erlbaum.
- Keehner, M., Hegarty, M., Cohen, C. A., Khooshabeh, P., & Montello, D. R. (2008). Spatial reasoning with external visualizations: What matters is what you see, not whether you interact. *Cognitive Science*, 32, 1099–1132.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Knauff, M., & Johnson-Laird, P. N. (2002). Visual imagery can impede reasoning. *Memory & Cognition*, 30(3), 363–371.
- Knauff, M., Fangmeier, T., Ruff, C. C., & Johnson-Laird, P. N. (2003). Reasoning, models, and images: Behavioral measures and cortical activity. *Journal of Cognitive Neuroscience*, 15(4), 559–573.
- Knauff, M., Mulack, T., Kassubek, J., Salih, H. R., & Greenlee, M. W. (2002). Spatial imagery in deductive reasoning: A functional MRI study. *Cognitive Brain Research*, 13, 203–212.
- Kosslyn, S. M. (1994). *Image and brain*. Cambridge, MA: MIT Press.
- Kulhavy, R. W., Stock, W. A., & Caterino, L. C. (1994). Reference maps as a framework for remembering text. In W. Schnotz & R. W. Kulhavy (Eds.), *Comprehension of graphics* (pp. 153–162). Amsterdam: Elsevier Science B.V.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65–99.
- Lowe, R. K. (1996). Background knowledge and the construction of a situational representation from a diagram. *European Journal of Psychology of Education*, 11, 377–397.
- Mayer, R. E. (2009). *Multimedia learning*. New York: Cambridge University Press.
- McNamara, D. S. (Ed.). (2007). *Reading comprehension strategies: Theories, interventions, and technologies*. New York: Lawrence Erlbaum.
- Neisser, U. (1976). *Cognition and reality*. San Francisco: Freeman.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.
- Pichert, J. W., & Anderson, R. C. (1977). Taking different perspectives on a story. *Journal of Educational Psychology*, 69, 309–315.
- 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.
- Rasch, T., & Schnotz, W. (2009). Interactive and non-interactive pictures in multimedia learning environments: Effects on learning outcomes and learning efficiency. *Learning and Instruction*, 19, 411–422.
- Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, N.J.: Erlbaum.
- Sachs. (1967). Recognition memory for syntactic and semantic aspects of a connected discourse. *Perception and Psychophysics*, 2, 437–442.

- Sadoski, M., & Paivio, A. (2001). *Imagery and text. A dual coding theory of reading and writing*. Mahwah NJ: Erlbaum.
- Schnotz, W. (1993). On the relation between dual coding and mental models in graphics comprehension. *Learning and Instruction*, 3, 247–249.
- Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (2nd ed., pp. 72–103). Cambridge: Cambridge University Press.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction*, 13, 141–156.
- Schnotz, W., & Kulhavy, R.W. (1994). *Comprehension of graphics*. Volume in the series Advances in Psychology. Amsterdam: Elsevier Publishers.
- Schnotz, W., & Kürschner, C. (2008). External and internal representations in the acquisition and use of knowledge: Visualization effects on mental model construction. *Instructional Science*, 36, 175–190.
- Shah, P., Mayer, R. E., & Hegarty, M. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology*, 91(4), 690–702.
- Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1991). Knowledge representation, content specification, and the development of skill in situation-specific knowledge assembly: Some constructivist issues as they relate to cognitive flexibility theory and hypertext. *Educational Technology*, 31(9), 22–25.
- Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire CT: Graphics Press.
- Tufte, E. R. (1990). *Envisioning information*. Cheshire CT: Graphics Press.
- Tufte, E. R. (1997). *Visual explanations. Images and quantities, evidence and narrative*. Cheshire CT: Graphics Press.
- Tversky, B., Heiser, J., Mackenzie, R., Lozano, S., & Morrison, J. (2008). Enriching animations. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation. Research implications for design* (pp. 263–285). New York: Cambridge University Press.
- Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, 57, 247–262.
- Ullman, S. (1984). Visual routines. *Cognition*, 18, 97–159.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Wainer, H. (1997). *Visual revelations. Graphic tales of fate and deception from Napoleon Bonaparte to Ross Perot*. Mahwah, NJ: Lawrence Erlbaum.
- Weaver, C. A., III, Mannes, S., & Fletcher, C. R. (Eds.). (1995). *Discourse comprehension*. Hillsdale, NJ: Erlbaum.
- Wertheimer, M. (1938). *Laws of organization in perceptual forms in a source book for Gestalt Psychology*. London: Routledge & Kegan Paul.
- Winn, W. D. (1994). Contributions of perceptual and cognitive processes to the comprehension of graphics. In W. Schnotz & R. Kulhavy (Eds.), *Comprehension of graphics* (pp. 3–27). Amsterdam: Elsevier.
- Witrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist*, 24, 345–376.