THE ROLE OF SEMANTIC CONGRUENCY IN THE DESIGN OF GRAPHIC ORGANIZERS

William A. Kealy
University of South Florida

Abstract: Thirty-one undergraduates read a 204-word passage describing social groupings of fish and their color, depth, size and diet. Subjects concurrently viewed a graphic organizer (GO) that presented this conceptual information in a matrix-like format. Half the GOs were semantically congruent with the text and organized according to social grouping while the remaining ones were spatially organized by fish depth. Subjects answered comparison and inference-type questions immediately after the study period and again two days later. Analysis of variance showed semantic congruity influenced GO efficacy but only for delayed recall of material related to the main topic of the text.

INTRODUCTION

More than ever before, people need assistance in absorbing and making sense of the enormous flow of data generated by the Information Age. Consequently, there has been increased interest in the potential of graphics for assimilating and mentally processing knowledge. In some cases, graphics facilitate acquisition of new knowledge in a supplemental way by embellishing or illustrating an accompanying text passage (Levie & Lentz, 1982; Levin & Lesgold, 1978). Alternatively, graphics may offer a meaningful context that is critical to how well an accompanying text is interpreted and remembered (Bransford & Johnson, 1972; Schallert, 1980).

Sometimes a graphic display may be preferred over text for delivering information simply because the former possesses a cognitive capacity (Kozma, 1991) lacking in the latter. For example, Larkin and Simon (1987) persuasively demonstrated that a simple diagram of a pulley was vastly more effective than its sentential counterpart in helping one to understand the physics principles being presented. Because, unlike text, the processing of diagram-based information is not serially constrained (Bauer & Johnson-Laird, 1993), it offers a “computational efficiency” that assists

• **William A. Kealy**, Ph.D., Associate Professor of Instructional Technology, Department of Secondary Education, College of Education, University of South Florida, 4202 East Fowler Avenue, EDU 162, Tampa, FL 33620-5650; phone: (813) 974-7952; E-mail: wkealy@tempest.coedu.usf.edu; Web: www.coedu.usf.edu/kealy

The Quarterly Review of Distance Education, Volume 1(3), 2000, pp. 205–214 ISSN 1528-3518

in determining the relationships of components depicted.

Another type of display that may be an efficient alternative to text is the graphic organizer or GO (pronounced "Gee-Oh"). This is a tabular array of categorical information that is indexed over one or more variables. Consider, for example, a text passage about types of fish and the characteristics of each such as size, color, social grouping (i.e., solitary, small, or school), swimming depth, and diet. In this case, a corresponding GO might arrange fish names in a row according to swimming depth while listing, as a column below every name, the social grouping, color, size, and diet of each fish. Such an arrangement aids visual reasoning (Gattis & Holyoak, 1996) by using space to organize factual data in a way that facilitates quick comparisons between types of fish across several different variables. Because a GO reveals how two or more characteristics (e.g., size and color) vary with respect to one another, learners can more readily draw inferences based on conceptual interrelationship (e.g., larger fish tend to be lighter in color).

Robinson and Schraw (1994) studied the aforementioned GO and the 204- word text from which it was derived. Subjects read the text and then studied either the GO, an outline of the text, or completed a second reading of the text. They then judged the accuracy of 68 separate statements related to the text, either comparing two types of fish on a particular characteristic or inferring the relationship between two characteristics. Significantly more correct responses were made by subjects studying a graphic organizer compared to those rereading the text a second time.

While GOs have been shown to be effective adjuncts to text, like other types of displays, their value can probably be further improved through improvements in design. Merely rearranging the verbal elements of a display can have a profound influence on how it is used to analyze the information presented. Wainer (1992), for instance, demonstrated that use of proximity to spatially cluster the labels of a table can evoke deeper processing of the display's data. Similarly, Schuch, Tada, von Eberstein, and Kealy (1999) found that spatially clustering the elements of a GO according to categorical membership facilitated significantly better comparative judgments about the material studied.

Another possible influence on GO efficacy, the focus of the current study, is the "semantic congruency" between a text and its corresponding GO when both are studied concurrently. Just as prose has a semantic structure--paragraphs begin with a main or topic sentence followed by ones subordinate in meaning--so too does a graphic display. How one "reads" both tables and diagrams is determined by the visually prominent display of information that cues viewers on how the display is conceptually organized. In both the Robinson and Schraw (1994) and Schuch, Tada, von Eberstein, and Kealy (1999) studies, the topic of the experimental text was the social groupings of fish. By contrast, the corresponding GO was primarily organized according to the depth at which the six types of fish lived: 200, 400 or 600 feet. This contradiction raises the question of whether such a semantic incongruity between text and GO causes interference during mental processing along with a resulting decrement in learning. Since the linguistic structure of text has a powerful effect on what readers recall (Johnson, 1970), it was reasoned that contradictions between the apparent semantic structures of the text and its accompanying GO may impede learning performance. By contrast, when the superordinate conceptual organization of the GO, reflected by the categories used for the top row of the display, were consistent with the structure of the corresponding text, we speculated that the GO would yield superior learning.

METHOD

Design and Subjects

The study consisted of a 2 Semantic Organization (Congruent vs. Incongruent) x 2

Depth of Fish (ft)

	200		**400**		**600**	
Fish:	Hat	Lup	Arch	Bone	Tin	Scale
Grouping:	solitary	small	solitary	school	small	school
Color:	black	brown	blue	orange	yellow	white
Size (cm):	30	30	45	45	70	70
Diet:	shrimp		krill		prawn	

Groupings of Fish

	solitary		**small**		**school**	
Fish:	Hat	Arch	Lup	Tin	Bone	Scale
Depth (ft):	200	400	200	600	400	600
Color:	black	blue	brown	yellow	orange	white
Size (cm):	30	45	30	70	45	70
Diet:	shrimp	krill	shrimp	prawn	krill	prawn

FIGURE 1
Semantically Congruent (Top) and Incongruent (Bottom) Graphic Organizers Used in the Study

Question Type (Comparison vs. Inference) x 2 Recall (Immediate vs. Delayed) x 5 Topic (Color vs. Depth vs. Feeding vs. Grouping vs. Size) factorial design. Semantic Organization was varied between-subjects while the remaining variables served as repeated measures. Thirty-one undergraduates participated in the study, receiving course credit for doing so.

Materials

Using the same text as that used in the Robinson and Schraw (1994) study, two versions of a GO were created: one organized by fish depth (200, 400, and 600 feet) and the other by social grouping (solitary, small, and school). The resulting GOs, shown at the top and bottom of Figure 1, respectively, were incorporated into two separate computer programs developed using Authorware. When run, the programs presented subjects with instruction on graphic organizers, showed samples of the comparison and inference questions they would later complete, and provided practice on the experimental task. The programs also collected and stored the responses by subjects to test items as well as the latencies of these responses.

Computer displays were presented on 15-inch monitors at SVGA resolution (600 x 800 pixels). All computer screens were designed with a light green background and dark blue

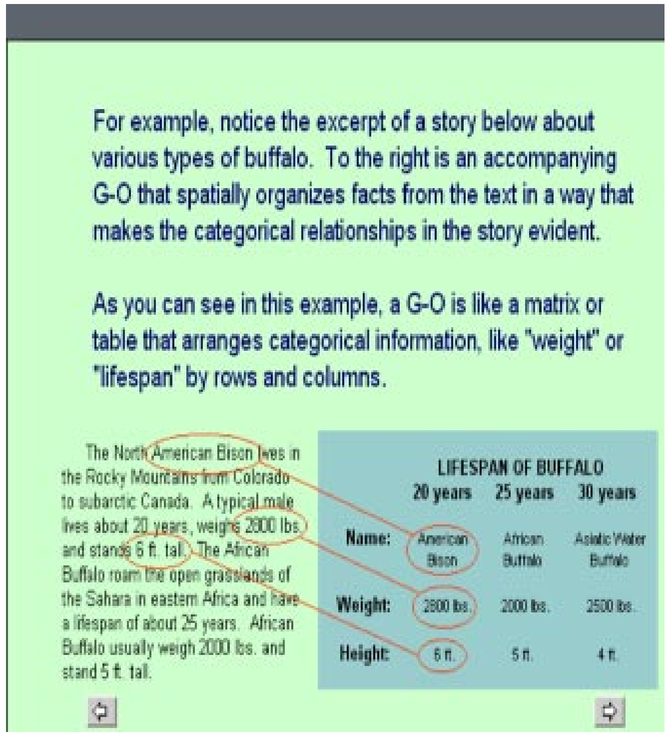

FIGURE 2
Example of Instruction on the Use of GOs Contained in the Computer-based Experimental Treatments

lowercase 14-point Ariel typeface with the exception of instructional examples and GO labels that used black 10-point lettering.

Since some subjects would likely be unfamiliar with graphic organizers, the experimental programs contained six screens that explained the nature and purpose of a GO and showed an example using facts about different buffalo (see Figure 2). By directing the viewer to specific points in the GO and text, the instruction illustrated how the GO could be especially useful for making comparisons and inferences from the information. During the instructional phase, subjects could progress or review the information by using the mouse to click the forward and back buttons, respectively, at the bottom of the screen.

The instruction explained that, while reading the text about fish types that would follow, subjects could view the accompanying GO at any time by holding down the mouse button. This allowed subjects to read the text while periodically consulting the GO whenever they wished to do so. In the Robinson and Schraw (1994) study, subjects first read the text passage, then studied the GO only after the text had been removed. By contrast, we chose to give subjects the means to study the text and GO concurrently, reasoning that this approach created a learning environment that more closely approximated the way people actually use a graphic organizer.

Another feature of the current study that was a departure from the Robinson and Schraw (1994) experiment was the nature of the criterion measure used. In the Robinson and Schraw study, subjects responded to 30 comparison and 30 inference statements, half within each category true and the rest false, to which they responded by pressing either of two keypads. However, a replication of this methodology by Schuch, Tada, von Eberstein, and Kealy (1999) showed only chance differences between subject responses to false comparison and inference statements.

The current study used a different strategy whereby subjects saw comparison and inference items in a two-item multiple-choice question format. An example of a comparison question used was, “Which is darker in color?” (Tin or Arch) while a sample inference-type question was, “Smaller groupings of fish tend to be _______ in size” (bigger or smaller). Although this approach also involved a 50% chance or making a correct response, it cut the total number of question in half thereby reducing the possibility of subject fatigue while potentially evoking greater semantic processing through a use of a more authentic recall task.

Both the pool of 15 comparison items and the group of 15 inference items consisted of three questions on each of five fish characteristics depicted in the text and GO: color, depth, social grouping, diet, and size. For a given characteristic such as color (e.g., black, brown, blue, orange, yellow, white), an attempt was made to select attributes that were not mentioned in any of the remaining questions. While not exhausting all the possible questions that could have been formed, this procedure nevertheless provided a broad and balanced coverage of the content studied by subjects. The entire pool of 30 comparison and inference questions thus constructed were built into the computer-based treatments so that they were presented in a separate random order to each subject during the experiment.

Procedures

Subjects participated in an experimental session in groups of about ten. Upon arrival for the session, each subject was randomly assigned to a desktop computer that contained one of the two experimental conditions. This resulted in a total of 18 subjects assigned to the Congruent treatment and 13 assigned to the Incongruent treatment.

Once all subjects were seated in front of their blank computer screens, the experimenter explained that they would be examining a short text passage and related display and then asked questions about what they studied. After all procedural questions were answered, subjects pressed the TAB key to start the program which prompted them to type their name and press the ENTER key.

Subjects then viewed the first of the six instructional screens which repeated that they would have five minutes to study a 200-word text about different types of fish. To assist them in monitoring the passing time, the instruction stated, a small clock-like icon similar to the one currently on their screen would appear in the upper-left corner of the text. Proceeding at their own pace, subjects read through the remaining five instructional screens. The last screen again informed them that after studying the text and GO about fish they would answer comparison and inference types of questions about what they read. This time, however, they were shown examples of a hypothetical comparison and inference question based on the text about buffalo. Finally, subjects were instructed to raise their hand if a question remained about the experimental task or, if not, to wait for further instructions. When all questions had been answered, they were told to click on the word “please” in the text to progress to the screen containing the experimental text and graphic organizer.

Immediately after the five-minute study period, three two-column simple addition problems appeared on the screen. The computer asked subjects to confirm the accuracy of each sum by typing a “Y” if correct or an “N” if incorrect. This brief interpolated task was designed to clear the working memory of subjects before administering the criterion measure. Upon completion of the three problems, the computer prompted them to look up from their screen to indicate they were done. When all subjects indicated completion of their arithmetic, the experimenter told them to click on the word "look" to proceed to the next task.

The screen then showed subjects an example of a hypothetical comparison- type question dealing with a fish characteristic not mentioned in the text they studied earlier. At the lower half of the screen appeared two shaded rectangles, side-by-side and of equal size, with a one-word answer printed in the center of each. Subjects were told to respond to the question shown ("Which fish typically weighs more?") by clicking on the box with the correct answer (either "Cod" or "Dolphin"). Once clicked, both shaded box and enclosed word flashed in an inverse tone to signal that that choice had been made. Similarly, subjects were next shown an inference-type close question ("Fish that weigh less tend to have a ______ lifespan") and two possible answers each printed in a separate box ("longer" and "shorter").

The program informed subjects that an asterisk would appear in the center of the screen for two seconds just before each comparison or inference question appeared. They were encouraged to make their responses as soon as possible and then told to press the TAB key to practice sample comparison and inference questions. After finishing the practice items, subjects were informed that these questions were similar to the ones they were about to complete. Subjects were encouraged to pay attention, work quickly, and do their best while performing the task. When it was clear that they fully understood what they were being asked to do, the experimenter told them to click on the word "best" and to begin the test phase of the study. Within both groups of 15 questions, three questions dealt with one of the five categories of fish characteristics (e.g., depth, grouping, diet, color, and size) reported in the text and GO. The 30 questions thus formed were presented during the testing phase one at a time and in random order. Once the last person in the room completed answering the questions, all the subjects were dismissed from the study as a single group.

Two days later, subjects reported for a different experimental session (i.e., one they believed unrelated to GOs) and assigned to one of the desktop computers. When ready, they pressed the TAB key, starting the computer program. Subjects then entered their name and viewed instructions informing them that they would be answering the same questions completed two days earlier. Subjects pressed the TAB key once again, completed the 30 inference and comparison questions and were dismissed.

RESULTS AND DISCUSSION

During the testing phase of the study, the experimental program scored a one if a subject's response to an item was correct and a zero if the response was wrong. To facilitate data entry and analysis, the computer-based treatments calculated a mean score for each cluster of three test items involving the same fish characteristic for both comparison and inference questions. This produced ten mean proportional scores, five for comparisons and five for inferences, representing all the five categories of fish characteristics mentioned by the text and GO. Hence, a value of zero, .33, .67, or 1.0 was possible for each mean proportional score thus calculated. Following the recommendations of Winer (1971), an arcsine transformation was performed on each of the proportional scores prior to statistical analysis. An alpha level of .05 was used for all statistical tests.

Accuracy of Recall Performance

Data on the accuracy of subjects' responses to comparison and inference questions were entered in a 2 Semantic Organization x 2 Mental Operation x 2 Recall x 5 Topic repeated measures ANOVA. The analysis revealed a significant main effects for the Mental Operation, $F(1, 29) = 4.89$, $p = .035$, and Topic, $F(4, 116) = 3.12$, $p = .018$, variables and a significant Mental Operation x Topic interaction,

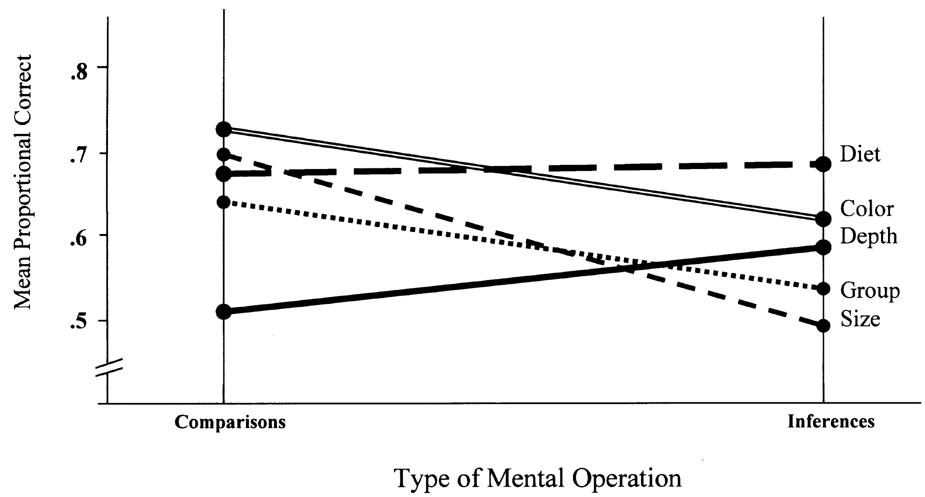

FIGURE 3
Influence of Text Topic on Performance of Subjects in Making Cojmparative and Inferential Judgments

$F(4, 116) = 4.81$, $p < .001$, that accounted for most of the variability ($\chi^2 = .14$) in the dependant measure. Figure 3 depicts the mean proportional scores involved in the Mental Operation x Topic interaction. This illustrates the superior performance by subjects on comparison items compared to inference questions and replicates the findings of previous research by Robinson and Schraw (1994). Both the role of Topic on performance and its interaction with Mental Operation is less clear. Evidently subjects performed poorer on questions related to fish depth but only when the questions dealt with comparisons.

It was puzzling that subjects' performance showed no significant decrement between the first experimental session and the second trial two days later. To examine this further and improve interpretability of the data, a separate ANOVA was performed for each of the sessions. This revealed that differences between performance on comparison and inference items were significant for immediate, $F(1, 29) = 4.71$, $p = .038$), but not delayed, recall. The Topic variable, on the other hand, only had a significant influence on performance occurring two days after study of the GO. The analysis of delayed recall showed an interaction between Semantic Organization (i.e., the type of GO studied) and Topic that was not significant, $F(4, 116) = 2.09$, $p = .063$), by conventional standards, but one that was nevertheless interesting. As Figure 4 illustrates, the mean proportional score on comparisons and inference questions was notably lower for items dealing with fish grouping when the GO used was spatially organized by depth.

While this was the expected outcome, it raises the question of why such a decrement in performance by subjects studying a semantically incongruous GO only occurred two days after the display was studied? One conceivable explanation is that, due to the dominance of subjects' verbal processing abilities, the deleterious effect of an incongruous graphic display had little impact on immediate recall. However, two days later when memory of the information was diminished, subjects began to more actively use the spatially-encoded GO as a secondary retrieval cue for related information in verbal storage (see Kulhavy, Stock, & Kealy, 1993 for a discussion on the value of maps as secondary retrieval cues). Hypothetically, if the GO brought into working memory was semantically incongruous with the text, it proactively interfered (Osgood, 1949) with

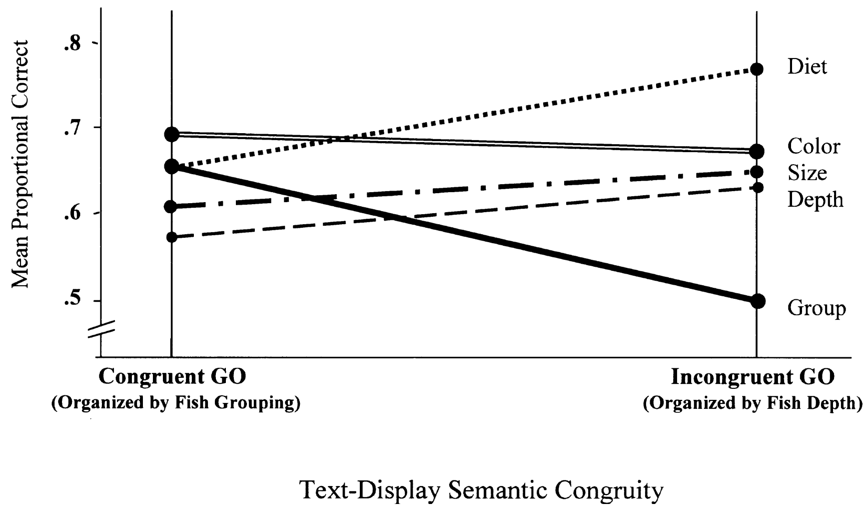

FIGURE 4
Relationship Between Text Topic and the Semantic Congruity of a Graphic Organizer on Delayed Recall

verbal recall--especially information contained in the macrostructure of the text (i.e., facts about depth).

Latency of Recall Performance

Latencies of subjects' responses to questions were entered into a 2 Semantic Organization x 2 Mental Operation x 2 Recall x 5 Topic repeated measures ANOVA. This revealed a significant main effect for Mental Operation, $F(1, 29) = 24.61$, $p < .000$, which accounted for nearly half of subjects' variability ($\chi^2 = .46$) in scores. ANOVA also identified significant main effects for Topic, $F(4, 116) = 4.29$, $p = .003$, and, unlike the analysis of test performance, Recall $F(1, 29) = 13.67$, $p < .001$.

Additionally, ANOVA indicated significant two-way interactions for Semantic Congruity x Mental Operation, $F(1, 29) = 4.93$, $p = .034$, Semantic Congruity x Topic, $F(4, 116) = 2.67$, $p = .035$, and Mental Operation x Recall, $F(1, 29) = 8.14$, $p = .008$. Figure 5 depicts the relative effect of Semantic Congruity on latency of performance for comparison and inference questions. As expected, subjects took much less time in responding to the comparison questions compared to the relatively more difficult inference items. On the other hand, the dramatic drop in latency of response among subjects who studied the semantically incongruous GO, may be an indication of mental processing that was simply more superficial rather than more efficient.

This hypothesis is supported, in part, by Figure 6 which shows the differential effects of GO type on speed for processing questions according to the Topic variable. The slower processing speed evident for questions related to grouping, the main topic of the text, points to the dominant status of verbal processing among adult learners. The semantic structure of text provides a strong interpretive mechanism which, in the present case, signaled that the main topic of the passage was the socialization of fish. Given the absence of explicitly stated educational objectives, learners are likely to use the semantic structure of a studied text to help determine the instructional task they will be asked to perform. The relationship between instructional objectives and task expectations is especially interesting in the context of graphic organizers. It is intriguing, for instance, to imagine what effect study of the GO prior to viewing the relevant text

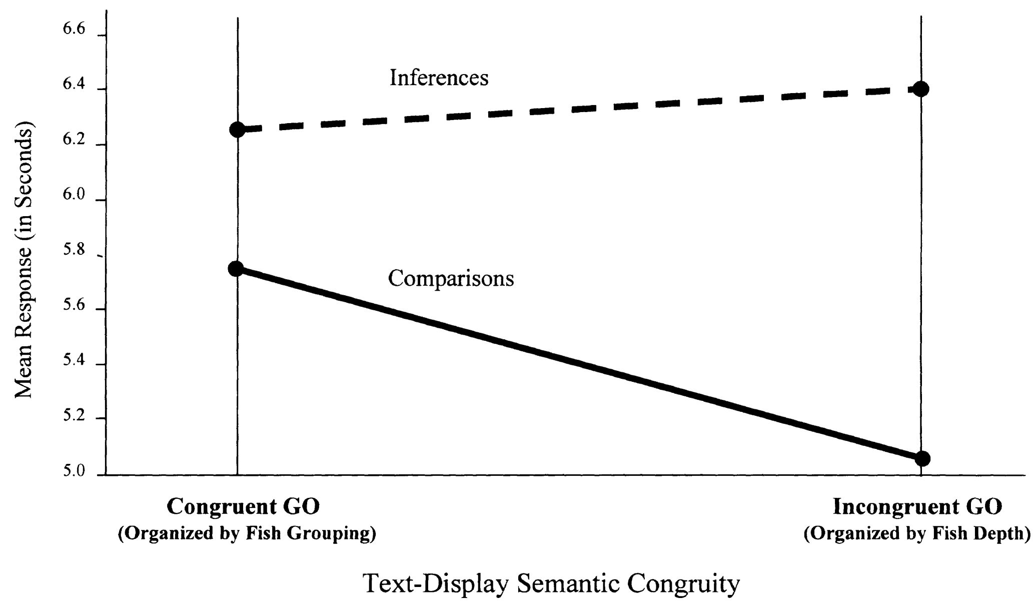

FIGURE 5
Effectg of Text-to-Display Semantic Congruence on Response Latencies for Answering Comparison and Inference Questions

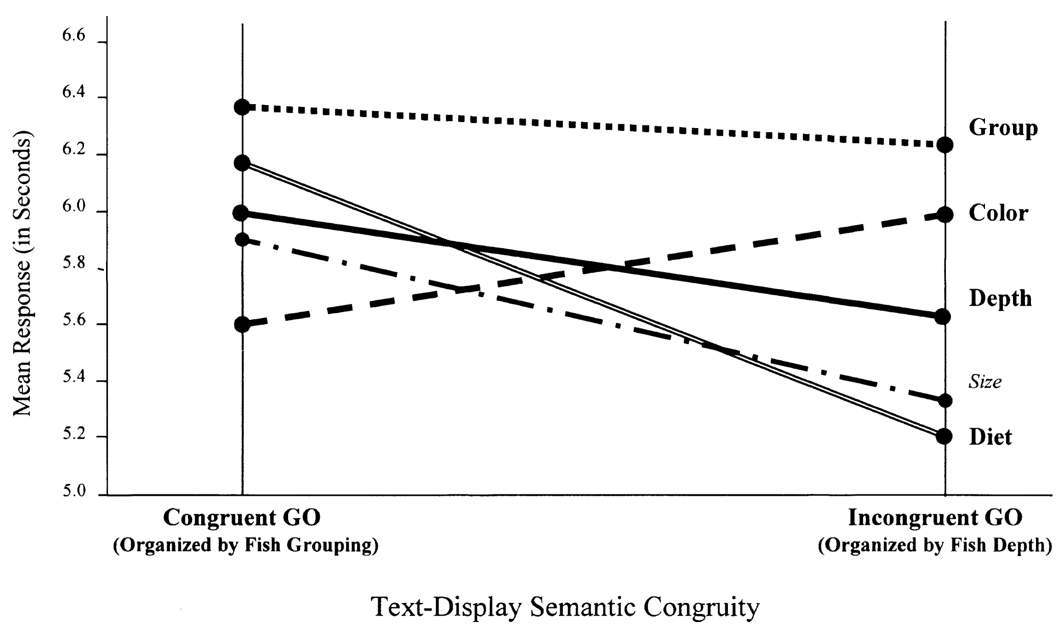

FIGURE 6
Effect of Text-to-Display Semantic Coingruency on Response Latencies for Answering Questions According to Topic

would have had on subjects' task expectancy as well as processing precedence.

Studies on how to effectively integrate cognitive tools, such as graphic organizers, with instructional design practices exemplify the kind of "prescriptive" research (Clark, 1989) that could contribute much to the field of educational technology.

Graphic organizers are important adjuncts for enhancing what people recall from a related text. In many cases they yield insights on the relationship between facts in a story even better than the story itself. To be effective, however, such learning tools should be designed so that they are semantically congruent with the texts they accompany. Doing so will maximize the mental processing of both.

REFERENCES

Bauer, M.I., & Johnson-Laird, P.N. (1993). How diagrams can improve reasoning. *Psychological Science, 4*(6), 372-378.

Bransford, J.D., & Johnson, M.K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*, 717-726.

Clark, R.E. (1989). Current progress and future directions for research in instructional technology. *Educational Technology Research and Development, 37*(1), 57-66.

Gattis, M., & Holyoak, K.J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 231-239.

Johnson, R.E. (1970). Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Verbal Behavior, 9*, 12-20.

Kozma, R.B. (1991). Learning with media. *Review of Educational Research, 61*(2), 179-211.

Kulhavy, R.W., Stock, W.A., & Kealy, W.A. (1993). How geographical maps increase recall of instructional text. *Educational Technology Research and Development, 41*(4), 47-62.

Larkin, J.H., & Simon, H.A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65-99.

Levie, W.H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *Educational Communications and Technology Journal, 30*, 195-232.

Levin, J.R., & Lesgold, A.M. (1978). On pictures in prose. *Educational Communications and Technology Journal, 26*(3), 233-243.

Osgood, C.E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review, 56*, 132-143.

Robinson, D.H., & Schraw, G. (1994). Computational efficiency through visual argument: Do graphic organizers communicate relations in text too effectively? *Contemporary Educational Psychology, 19*(4), 399-415.

Schallert, D.L. (1980). The role of illustrations in reading comprehension. In R.J. Spiro, B.C. Bruce, & W.F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education.* Hillsdale, NJ: Erlbaum.

Schuch, D., Tada, N., von Eberstein, A.M., & Kealy, W.A. (1999, February). *Design variables in the effectiveness of graphic organizers*. Paper presented at the Annual Convention of the Association for Educational Communications and Technology, Houston.

Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher, 21*(1), 14-23.

Winer, B.J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.

Winn, W.D., & Holliday, W.G. (1982). Design principles for diagrams and charts. In D.H. Jonassen (Ed.), *The technology of text* (pp. 277-299). Englewood Cliffs, NJ: Educational Technology Publications.