The Effect of the Spatial Arrangement of Simple Diagrams on the Interpretation of English and Nonsense Sentences

William Winn
Cliff Solomon

In two experiments, subjects received problem sets consisting of a simple diagram accompanied by two sentences and were asked to select the sentence that best expressed the meaning of the diagram. Each diagram showed a relationship between two concepts. The relationships were category membership, possession of a property, and causality. The relative placement of the concepts was varied. In the first experiment, the concepts were given nonsense names. Significant biases in response frequencies were found, suggesting that, in the absence of semantic content, the interpretation of the diagrams was largely determined by syntactic rules of English. In the second experiment, similar biases were found for diagrams whose concepts were given English names. However, the biases were less evident when the diagrams violated the rules of English syntax. This suggested that conflict between the sense of English and a diagram's spatial arrangement makes consistent interpretation unlikely. Conclusions are drawn concerning the relationships between the spatial arrangement of diagrams and the meaning of relationships among concepts stated in sentences. Suggestions for diagram design are offered.

Research has consistently demonstrated that diagrams can improve text comprehension. This is true whether the diagrams accompany text in instructional materials (Guri-Rozenblit, 1988a, 1988b, 1989; Holliday, Brunner, & Donais, 1977; Koran & Koran, 1980), or whether students construct diagrams to map the structure of the text they are studying (Armbruster & Anderson, 1982, 1984; Berkowitz, 1986; Geva, 1983; Johnson, Pittelman, & Heimlich, 1986; Prater & Terry, 1988; Ruddell & Boyle, 1989).

Diagrams make text easier to understand because they describe abstract and often implicit relationships among concepts explicitly and in two-dimensional space (Schewel, 1989; Sinatra, Stahl-Gemake, & Berg, 1984; Winn & Holliday, 1982). Two advantages are claimed for such spatial representations. First, they reveal the structure of the text, which helps the reader assimilate new knowledge to existing memory schemata (Holley & Dansereau, 1984; Hughes, 1989). Second, they require the reader to perform less translation of the material in the text before arriving at a satisfactory conceptual structure of the information the text contains (Breuker, 1984). This second claim has been supported by Larkin and Simon (1987) and by Winn, Li, and Schill (1991), who have demonstrated that, relative to text, diagrams make it easier to search through information and answer questions because information relevant to these tasks is given in the way concepts are arranged on the page.

However, research has not yet established how people understand *specific* spatial arrangements among concepts in diagrams. Moreover, instructional designers using diagrams in their materials often need specific prescriptions for the spatial arrangements of objects. Following from the work of Larkin and Simon (1987), the present study therefore examined how specific variations in the spatial arrangement of diagrams affected interpretation. In particular, we set out to discover whether people interpret identically arranged diagrams in consistent ways, and whether interpretation is influenced by the meaning of what diagrams represent. Answers to these questions form the basis for principles of diagram design.

Spatial arrangements for the study were selected from those commonly used in text-mapping procedures. Holley and Dansereau (1984) and Sinatra, Stahl-Gemake, and Morgan (1986) require the construction of diagrams in which concepts are arranged in hierarchies with superordinate categories above subordinate concepts, linked by lines. Armbruster and Anderson (1982) use boundaries around concepts to express category membership and the possession of a property. They also map causality by placing the cause to the left of the effect and joining them with an arrow pointing from the cause to its effect. Similar arrangements are used in other text-mapping procedures. (See Holley & Dansereau, 1984, and Hughes, 1989, for descriptions of many such procedures and for evidence of their effectiveness.)

It therefore appears that important features of diagrams that show relationships among concepts are: (1) the placement of one concept to the left or right of, or above or below, another, (2) connecting related concepts with lines or arrows, and (3) surrounding them with a common boundary. It is also apparent that there is a fairly consistent mapping between these features and the types of relationship that diagrams of text typically show. For example, causes are usually placed to the left of or above effects, subordinate categories are placed below or inside superordinate categories, and properties are placed inside or to the right of the objects that possess them.

The first purpose of this study was to verify empirically that the conventional spatial arrangements of diagrams illustrating relationships among concepts are interpreted consistently. Guri-Rozenblit (1989) has pointed out that the creation and interpretation of diagrams are not governed by strict grammatical rules in the same way that writing and understanding text are. This observation implies a lack of consistency in the design of diagrams and ambiguity in their interpretation. Diagrams seem to adhere more to what Salomon (1979, p. 30) has called "conventions of coherence" than to "rules of prescription." Nonetheless, if diagrams communicate well because of the arrangement of their components, the conventions they employ must be sufficiently consistent to act with at least some of the force of grammatical rules. Biases in the way spatial arrangements are interpreted would be evidence for this consistency and would form the basis for design prescriptions.

The second purpose of the study was to examine the role of spatial arrangements in establishing meaning. Boxes, lines, and locations have no semantic component of their own. Interpreting a word to the left of another as a cause rather than an effect relies on a purely syntactic convention. Yet diagrams that accompany text succeed because they express semantically correct relationships among concepts. This implies that the arrangements of concepts in diagrams must also accurately reflect the semantic content of the text. Failure to meet this requirement will lead to confusion and inconsistent interpretation. Ambiguity in the interpretation of diagrams *not* conforming to the semantic content of text would be evidence for the importance of this requirement.

Two experiments were designed to test these general assumptions. The first addressed the question of whether spatial arrangements in diagrams are interpreted consistently. The second examined relationships between spatial arrangement and the semantic content of text.

EXPERIMENT 1

In the first experiment, subjects selected from two sentences the one they felt best expressed

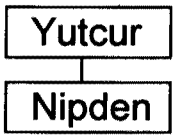
the meaning of a simple diagram. Each diagram showed the relationship between two concepts whose names were placed in boxes. In order to control for the effect of sentence meaning on interpretation, the concepts were given nonsense names. Three common spatial arrangements for showing relationships, described above, were used in the diagrams: concept A above concept B (Vertical format); A to the left of B (Horizontal format); and A in its box inside the box containing B (Included format). These formats are shown in Figure 1. Three relationships frequently associated with these formats were also selected for study: category membership (A is B), the possession of a property (A has B), and causality (A causes B).

The pairs of sentences from which subjects chose their interpretation were manipulated in two ways. In some pairs, as in the first three examples in Figure 1, the relationships were the same, but the concept's role as subject or object of the sentence was reversed. For example, the choice would be between "Kolfab cause vednex" and "Vednex cause kolfab," as in the third example in Figure 1. In other pairs, the syntactic functions of the concepts were the same, but the relationship was different. The choice would then be between "Purfep are fomgaf" and "Fomgaf cause purfep," as in the fourth example in Figure 1. In this way, all possible interpretations could be compared. However, only some comparisons were relevant to the research questions. First, each sentence was compared with its reverse in order to determine response consistency as a function of the relative positions of concepts. For example, is "kolfab cause vednex" selected more often than "vednex cause kolfab" for a given spatial arrangement? Second, comparisons were made of all pair combinations of the three relationships across the three formats to determine whether one format was consistently judged to be better than the others at expressing a particular relationship. For example, do the relative frequencies with which "purfep are fomgaf" and "fomgaf cause purfep" are selected differ when the diagram is presented in the Vertical format as compared to the Horizontal format?

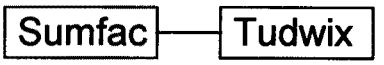
Hypotheses formed pertaining to the three relationships selected for the study were:

1. *Category membership:* A concept located above, to the left of, or outside another would be interpreted as the category to which the other belonged.
2. *Possession of a property:* A concept located below, to the right of, or inside another would be interpreted as an attribute of the other.

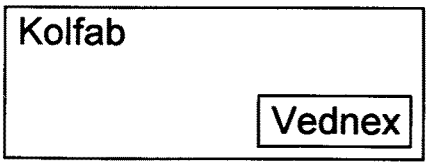
FIGURE 1 □ Diagram Arrangements and Examples of Sentences Used in Experiment 1. Subjects chose the sentence that best expressed the meaning of the diagram. The sentences marked "a" signify "A is B," "A has B," etc., in the text. The "b" sentences are "B is A," etc. Although only four sets are shown, all possible permutations of arrangements and sentence pairings were presented to subjects.



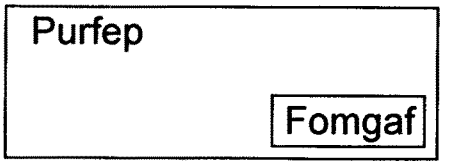
a. A yutcur is a nipden
b. A nipden is a yutcur



a. Sumfac have tudwix
b. Tudwix have sumfac



a. Kolfab cause vednex
b. Vednex cause kolfab



a. Purfep are fomgaf
b. Fomgaf cause purfep

3. *Causality:* A concept located above, to the left of, or outside another would be interpreted as the cause of the other.

Significant differences between the frequencies with which subjects selected sentences that reflected these conventions would confirm these hypotheses.

Method

Subjects

Sixty-seven graduate students in education volunteered to take part in the study. None was included whose first language was not English. Subjects' academic maturity was taken to indicate that they would have no difficulty reading the sentences.

Materials

A pool of nonsense words was created. Each word was six letters long and had two syllables, each of which consisted of a consonant, a vowel, and another consonant. Within these constraints, the words were created randomly by computer. The words used to construct the diagrams were selected from this pool on the condition that they could be pronounced easily and that they did not resemble any known words or acronyms. Each diagram was constructed by randomly selecting two words. In the Horizontal format, the words were placed in boxes beside each other and joined by a line. In the Vertical format, one word was placed above the other. In the Included format, one word was placed in a box and then placed inside a larger box containing the other word. Examples are shown in Figure 1.

Placed below each diagram were a pair of sentences from which subjects were to select the one that best expressed the meaning of the diagram. Each sentence expressed a relationship, either category membership ("A yutcur is a nipden"), possession of a property ("A yutcur has a nipden"), or causality ("A yutcur causes nipden"). With the three reverse forms of these sentences ("A nipden is a yutcur," "A nipden has a yutcur," and "A nipden causes yutcur"), a total of six interpretations was possible for each diagram. Responses to all possible combinations of these six interpretations were collected. This meant that choices from 15 pairs of sentences were required for each of the three diagram formats, which gave a total of 45 diagram/sentence-pair combinations.

The 45 diagrams and sentence pairs were randomly ordered and printed five to an 8½ by 11 inch sheet. A cover sheet was prepared on which instructions were written and three practice items provided.

Because no "favorite" was predicted for some sentence pairs, a five-point scale was placed beside each item so that subjects could rate their confidence in their choices.

Procedure

Subjects completed the task in class. They were instructed to mark the sentence in each pair that, for them, best expressed the meaning of the diagram. They were told that there were no right answers. They were also told to use the five-point scale beside each item to rate their confidence that the sentence they selected was the best choice. They worked through three practice items and then continued with the other 45 items. There was no time limit for completion of the task, but in all cases the task was completed in less than 15 minutes.

Data Analysis

The selection frequency for each sentence was determined. Appropriate nonparametric tests, described below, were applied to the data. Alpha was set at .05 for all analyses.

One-sample chi-square analysis was conducted to determine whether subjects selected a particular sentence more frequently from among the pairs "A is B" and "B is A," "A has B" and "B has A," and "A causes B" and "B causes A." This analysis was performed for diagrams in each of the three formats. One-sample chi-square analysis was also performed on sentence selection frequency for all combinations of "A is B," "A has B," and "A causes B," and of "B is A," "B has A," and "B causes A." The McNemar test was used to compare sentence selection across formats. For exam-

ple, the relative frequencies of sentences selected for "A has B" and "A is B" in the Horizontal format was compared to the corresponding relative selection frequencies for the Vertical and Included formats. Likewise, the relative frequencies for the Vertical format were compared to those for the Included format.

Finally, the confidence ratings for each response were compared across sentences in each pair. The Mann-Witney test was used for this analysis.

Results

In the following sections, "A" refers to the word to the left of "B," above "B," or in the outside box in the Included format (see Figure 1).

Table 1 shows the frequency with which each sentence was selected when the choices were between sentences stating the same relationships but with concepts in different orders. Chi-square analysis showed that subjects chose "A has B" significantly more than "B has A" and "A causes B" more than "B causes A" for all three diagram formats. For the "A is B" and "B is A" comparisons, there was only a significant difference for the Included format. In this case, subjects chose "B is A" more frequently than they chose "A is B."

Table 2 shows the sentence selection frequencies when each sentence described a different relationship. Chi-square analysis showed that all differences were significant except for the pairs "A is B" and "A has B" and "A is B" and "A causes B" in the Horizontal format, "A has B" and "A causes B" and "B has A" and "B causes A" in the Vertical format, and "B is A" and "B has A" in the Included format.

The McNemar test was used to compare relative selection frequencies across formats. In Table 2, within each row, the pairs of frequencies for the Horizontal format were compared to the pairs of frequencies for the Vertical and Included formats. The pairs of frequencies for the Vertical format were compared to those for the Included format. For example, a contingency table was constructed from the selection frequencies for "A yutcur is a nipden" and

TABLE 1 ☐ Frequencies of Selection of Sentences in Experiment 1

		HORIZONTAL FORMAT			VERTICAL FORMAT			INCLUDED FORMAT		
Sentence		*Sentence*			*Sentence*			*Sentence*		
1	*2*	*1*	*2*	p	*1*	*2*	p	*1*	*2*	p
A is B	B is A	38	29	–	33	34	–	20	47	<.001
A has B	B has A	57	10	<.001	59	8	<.001	56	11	<.001
A causes B	B causes A	63	4	<.001	58	9	<.001	54	13	<.001

Note: In this and subsequent tables, "A" refers to the word to the left of, above, and outside "B," i.e., sentence 1 corresponds to the sentences labeled "a" in Figure 1 and sentence 2 to the sentences labeled "b".

TABLE 2 ☐ Frequencies of Responses Across Relationships in Experiment 1

		HORIZONTAL FORMAT			VERTICAL FORMAT			INCLUDED FORMAT		
Sentence		*Sentence*			*Sentence*			*Sentence*		
1	*2*	*1*	*2*	p	*1*	*2*	p	*1*	*2*	p
A is B	A has B	32	35	–	15	52	<.001	20	47	<.001
A is B	A causes B	26	41	–	14	53	<.001	24	43	<.05
A has B	A causes B	15	52	<.001	38	29	–	57	10	<.001
B is A	B has A	60	7	<.001	45	22	<.01	40	27	–
B is A	B causes A	60	7	<.001	52	15	<.01	46	21	<.05
B has A	B causes A	22	45	<.001	35	32	–	14	53	<.01

Note: For analysis across format, 2 × 2 contingency tables were constructed for both sentences in all pairwise combinations of columns. Thus, for the first row, the tables were 32/35 with 15/52, 32/35, with 20/47, and 15/52 with 20/47.

"A yutcur has a nipden" in the Horizontal format, and the same pairs of frequencies in the Vertical format. All comparisons across formats were significant except: "A is B" and "A has B" when the Vertical and Included formats were compared; "A is B" and "A causes B" when the Included format was compared with the Horizontal and Vertical formats; "B is A" and "B has A" and "B is A" and "B causes A" for the Vertical and Included formats; and "B has A" and "B causes A" when the Horizontal format was compared to the Vertical and Included formats.

The mean confidence ratings for all the sentence pairs discussed above are shown in Table 3. The higher scores indicate subjects' greater confidence in the choices that they made. The maximum score was 5.

Results of the Mann-Witney test produced very few significant differences between confidence ratings for sentence pairs. However, in all of the cases where the differences were significant, the choice made by the majority of subjects received the higher confidence rating. Variation in the absolute values of mean confidence ratings suggests that subjects found some choices more difficult to make than others.

Discussion

Results support the hypotheses concerning the consistency of sentence selection. Significant selection biases were found for sentences that interpreted a cause as being to the left of, above, and outside an effect; a property being to the right of, below, and inside the object possessing it; and for superordinate categories being outside subordinate ones. Not only were these selection differences significant, inspection of Table 1 shows them to be large as well.

The two exceptions were the diagrams in the Vertical and Horizontal formats that expressed category membership. Here, the roughly even split in sentence selection frequencies suggests that there was no clear-cut preference for one interpretation over the other. An explanation for this is that subjects' interpretations were influenced as much by the syntax of English as they were by the diagrams, even though the materials contained nonsense words. The sentences used in this experiment were in the active voice, meaning that readers would interpret the first concept they read as the subject. There is evidence from cross-cultural studies that people's interpretation and production of graphics is influenced by the way their native language is written (Tversky, Kugelmass, & Winter, 1991; Wong & Kao, 1991). For English speakers, this means that diagrams are "read" from left to right and top to bottom (Winn, 1982, 1983). In the case of the diagrams in this experiment, the word to the left of, above, or outside the other would be read first in a diagram and interpreted as the subject. In the cases of diagrams showing causality and the possession of a property, causes or objects possessing properties would be read first and interpreted as subjects

TABLE 3 ☐ Mean Confidence Ratings in Experiment 1

		HORIZONTAL FORMAT			VERTICAL FORMAT			INCLUDED FORMAT		
Sentence		*Sentence*			*Sentence*			*Sentence*		
1	*2*	*1*	*2*	p	*1*	*2*	p	*1*	*2*	p
A is B	B is A	3.34	3.47	–	3.33	3.82	–	3.00	4.02	<.01
A has B	B has A	3.65	2.80	–	3.97	3.75	–	3.95	3.36	<.05
A causes B	B causes A	3.92	3.25	–	4.10	3.67	–	3.13	2.85	–
A is B	A has B	3.75	3.71	–	3.00	3.92	<.05	3.50	3.69	–
A is B	A causes B	3.92	3.90	–	3.93	3.96	<.05	2.38	2.98	–
A has B	B causes A	3.33	3.87	–	3.66	4.10	–	3.87	3.90	–
B is A	B has A	3.60	3.43	–	3.38	3.45	–	3.78	3.22	<.05
B is A	B causes A	3.48	3.00	–	3.21	3.13	–	3.61	2.76	<.05
B has A	B causes A	2.05	2.42	–	2.43	2.81	–	2.36	3.36	<.01

just as they would in a sentence. One sentence could therefore be selected unambiguously.

In the case of category membership, however, the subordinate category was the subject and was read first in the sentence, while in the Horizontal and Vertical diagrams the subordinate category was to the right of or below the superordinate one and was therefore read second. This created ambiguity since one sentence corresponded to the diagram while the other sentence corresponded to English syntax. Approximately half of the subjects chose one and half chose the other. This interpretation is corroborated by data in Table 2. In the Horizontal and Vertical formats, subjects selected the sentence that corresponded to English syntax ("B is A") over sentences that corresponded neither to English syntax nor to diagram conventions ("B has A" and "B causes A").

The results also provide information about which arrangements were chosen consistently to express each type of relationship. The data in Table 2 show marked bias favoring "A causes B" over "A has B" in the Horizontal format and the opposite in the Included format. This suggests that the Horizontal format consistently expresses causality and that the Included format consistently expresses property possession. Likewise, the Vertical and Included formats appeared to be more consistent in expressing property possession and causality than in expressing category membership.

Differences in confidence ratings were found for just five sentence pairs out of a total of 27. While it is interesting that four of these occurred in the Included format, suggesting that this spatial arrangement leads to the least tenuous interpretations, too few differences were significant overall to permit firm conclusions to be drawn. It appears that, generally, subjects who selected sentences that were not predicted to be chosen had just as good reasons for doing so as subjects who selected the sentences they were expected to choose.

EXPERIMENT 2

In instructional settings, diagrams do not contain nonsense words. Since the spatial arrangement of diagrams can so heavily bias interpretation, there is the potential for conflict between the interpretation implied in a less-than-perfectly designed diagram and the interpretation intended for the material it describes. The second experiment, therefore, examined the relationship between the appropriateness of the spatial arrangements used in diagrams and their semantic content. It was designed to determine how violating the conventions identified as effective in the first experiment would affect the comprehension of diagrams accompanying meaningful sentences.

The frequency with which sentences describing nonsense diagrams were selected was compared with selection frequency for sentences describing diagrams that contained English words. It was expected that when the diagram used the spatial arrangement shown in the first experiment to be the most appropriate for expressing a particular relationship, subjects seeing English sentences, like those seeing nonsense sentences, would consistently select the sentence that expressed that relationship. For example, if diagrams illustrating causality showed "junk food" to the left of "indigestion" and "yutcur" to the left of "nipden," then subjects would select the sentence that stated "junk food" or "yutcur" was the cause. On the other hand, if the arrangement was not the one that best expressed the relationship, English sentences would not be selected consistently while nonsense sentences would be. Thus, placing "indigestion" to the *left* of "junk food" would cause confusion by implying that indigestion causes junk food. No clear selection bias would occur—some subjects would choose the sentence that corresponded to the diagram and others would choose the sentence that made sense. Placing "nipden" to the left of "yutcur" would still be unambiguous, however, because the sentence has no interpretable semantic content; consequently, the sentence stating that "nipden" was the cause would be selected consistently. With the mean number of choices of sentences consistent with the diagrams as the dependent variable, a significant interaction between the appropriateness of the spatial arrangement and meaningfulness (English versus nonsense) would provide support for these hypotheses.

Method

Subjects

Sixty-three graduate students in education courses volunteered to take part in this experiment. Again, only subjects whose first language was English were included, and it was assumed that none would have difficulty reading the sentences. Subjects were randomly assigned to two groups, one receiving English materials ($N = 33$) and one receiving nonsense materials ($N = 30$).

Materials

The same pool of nonsense sentences was used as in the first experiment. A set of English sentences was constructed that corresponded to the nonsense sentences. For example, "Perfim causes rotlay" was matched by "Rain causes floods." The same three diagram formats were used: Horizontal, Vertical, and Included. The same three relationships were also used: category membership ("A cat is a mammal"); possession of a property ("A cup has a handle"); and causality ("Junk food causes indigestion"). Because this experiment was not concerned with differences across types of relationship, the only differences in the pairs of sentences to choose from was that in one pair the order of the concepts was reversed. This meant that pairs of sentences such as "A cat is a mammal" and "A mammal is a cat," "A cup has a handle" and "A handle has a cup," "Junk food causes indigestion" and "Indigestion causes junk food" were used, but not pairs of sentences such as "A cat is a mammal" and "A cat causes mammals." This gave a total of 9 possible permutations of diagram arrangement and type of relationship.

Subjects saw 2 "appropriate" and 2 "inappropriate" versions of each of the 9 permutations of arrangement and relationship, for a total of 36 diagrams. In "appropriate" diagrams, the arrangements were those that were concluded to be appropriate in the first experiment; that is, for example, the cause was placed to the left of, above, and outside the effect. (For the two nonsignificant cases where the Horizontal and Vertical formats expressed class inclusion, the names of the superordinate categories were placed above and to the left of the subordinate categories.) In the "inappropriate" diagrams, the positions of the names were reversed; that is, for example, the cause was placed to the right of, below, and inside the effect. Although the nonsense words did not allow subjects to determine which sequence was semantically correct, the positions of the names in the nonsense materials were also reversed. Figure 2 shows examples of the English materials. As before, a five-point scale was placed beside each item so that subjects could rate their confidence in their selections.

Procedure

The procedure was the same as that used in Experiment 1. Subjects were told to pick the sentence that, for them, offered the best interpretation of the diagram. Subjects seeing English materials were not given any additional instructions about how to respond to potentially nonsensical materials, since their freedom to choose on their own was essential to the experiment. In addition, subjects completed a brief open-ended questionnaire at the end of the experiment which asked them to state whether they found the task easy or difficult and why, and to report the strategies they had used in selecting their sentences.

Design and Analysis

The design of the study was a repeated-measures design with meaningfulness (English vs. nonsense) as a between-subjects factor and appropriateness of the spatial arrangement used in the diagram (appropriate vs. inappropriate) as a within-subjects factor. The dependent measures were the number of times each subject selected a response that was consistent with the diagram, and subjects' confidence ratings. It was expected that subjects who received English materials would choose fewer sentences that were consistent with the diagrams when the diagrams were inappropriate, while subjects who received nonsense materials would not be affected by the appro-

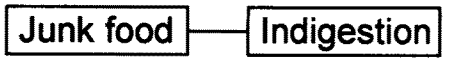
priateness of the diagrams. Therefore, support for the research hypothesis would come from an interaction between meaningfulness and appropriateness, with both groups selecting the same number of sentences that were consistent with the diagram when the diagrams were appropriate, but with the English group selecting fewer such sentences than the nonsense group when the diagrams were inappropriate.

Repeated-measures analysis of variance was also applied to the confidence ratings. Confidence ratings were expected to show an interaction similar to that found for selection frequencies.

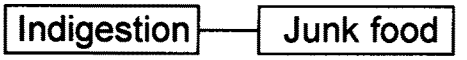
Results

The mean frequencies for the selection of sentences that were consistent with the diagrams are shown in Table 4. As predicted, subjects in both groups selected a relatively high number of sentences consistent with the diagrams when the spatial arrangements were appropriate. The number of such selections was much lower with inappropriate arrangements for the subjects who received English materials, but not for the subjects who received nonsense materials. Repeated-measures analysis of variance showed this interaction to be statistically significant, $F(1, 61) = 43.84$, $p < .001$.

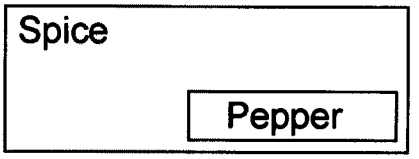
FIGURE 2 ☐ Samples of English Materials from Experiment 2. The nonsense materials were the same as in Figure 1. "Appropriate" examples are on the left; for example, the cause "junk food" is appropriately placed to the left of "indigestion." "Inappropriate" versions are on the right; for example, the cause "junk food" is on the right. In the "junk food" and "indigestion" appropriate pair, sentence "a" is consistent with the diagram while sentence "b" is inconsistent with it. In the corresponding inappropriate pair, "indigestion causes junk food" is consistent with the diagram even though it does not make sense in English. Although only six sets are shown, all possible permutations of arrangements and relationships were presented to subjects.



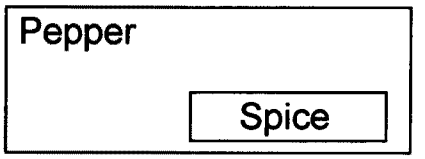
a. Junk food causes indigestion
b. Indigestion causes junk food



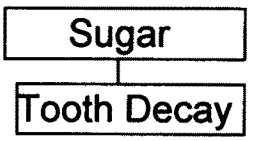
a. Junk food causes indigestion
b. Indigestion causes junk food



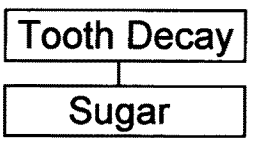
a. Pepper has a spice
b. Spice has a pepper



a. Pepper has a spice
b. Spice has a pepper



a. Sugar causes tooth decay
b. Tooth decay causes sugar



a. Sugar causes tooth decay
b. Tooth decay causes sugar

TABLE 4 ☐ Mean Frequencies of Selection of Sentences Consistent with the Diagrams in Experiment 2

	APPROPRIATENESS OF DIAGRAM					
	Appropriate		*Inappropriate*		*Total*	
GROUP	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
English	16.49	(2.18)	7.37	(3.66)	12.05	(2.74)
Nonsense	15.23	(2.21)	14.70	(2.21)	14.97	(2.21)
Total	15.89	(2.57)	11.00	(5.70)		

The main effect for materials was also significant, $F(1, 61) = 21.20$, $p < .001$, as was the within-subjects consistency effect, $F(1, 61) = 55.81$, $p < .001$. Univariate analysis of variance for inappropriate diagrams was significant $F(1, 61) = 39.01$, $p < .001$, with English subjects selecting fewer sentences corresponding to inappropriate diagrams than nonsense subjects. Univariate analysis of variance for appropriate diagrams was also significant, $F(1, 61) = 5.11$, $p < .05$, with subjects who received English materials choosing sentences consistent with diagrams more often than subjects who received nonsense materials.

Mean confidence ratings are shown in Table 5. The repeated-measures ANOVA indicated that subjects who received English materials were more confident in their responses than those who received nonsense materials, $F(1, 61) = 23.16$, $p < .001$, and that subjects' confidence in responses to appropriate diagrams was greater than their confidence in responses to inappropriate diagrams, $F(1, 61) = 19.98$, $p < .001$. An interaction between group and appropriateness was also evident, $F(1, 61) = 26.40$, $p < .001$. Differences were confined to the subjects who received English diagrams.

Responses to the questionnaire showed that 13 subjects (39%) who received English materials mentioned that inconsistency between the structure of the diagrams and the sense of English made it difficult to select a sentence. Only four (13%) of those who received nonsense materials attributed difficulty to the nonsense words. Beyond this, there was no consistency among reasons given for the difficulty of the task. Likewise, the strategies reported for interpreting the diagrams were highly idiosyncratic, for example: substituting real words for nonsense words (one subject), basing interpretations on how diagrams are used in math and computer programming (two subjects), ignoring the syntax of the sentences (two subjects), making completely random responses (one subject).

Discussion

The significant interaction between group and appropriateness supports the hypothesis that the use of inappropriate spatial arrangements has a greater impact on selecting English sentences than on selecting nonsense sentences. This finding is corroborated by the fact that some subjects who received English materials reported inconsistency between the diagrams and the sense of English as a source of

TABLE 5 ☐ Mean Confidence Ratings in Experiment 2

	APPROPRIATENESS OF DIAGRAM					
	Appropriate		*Inappropriate*		*Total*	
GROUP	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
English	4.42	(0.46)	3.87	(0.81)	4.15	(0.59)
Nonsense	3.14	(1.01)	3.18	(1.01)	3.16	(1.00)
Total	3.81	(1.00)	3.54	(0.96)		

task difficulty. It seems that the use of inappropriate arrangements led to confusion and ambiguity.

The significant difference between sentence selection frequency for appropriate diagrams, favoring the English materials, was not anticipated. This result suggests that there was even more consistency in interpretation of the diagrams when the concepts they related were known. This reinforces the importance of the relationship between the spatial arrangement and the semantic content of the diagram.

The confidence ratings were as expected. The interaction between materials and appropriateness was significant, corroborating the finding for sentence selection. Also important was the finding that confidence in interpretations of English materials was higher overall, even though reduced, when the diagrams were inappropriate. This finding suggests two things.

First, although interpretations of inappropriate English diagrams might not be consistent across a group of subjects, individuals were quite confident that their selections were correct. This result could be accounted for by a number of factors. Maybe subjects were better able to use English words than nonsense words to imagine circumstances where the sense of inappropriate diagrams might be plausible, e.g., that people with indigestion might eat junk food. Or maybe subjects decided early to select sentences exclusively on the basis of the sense of English, although this study found no evidence that this was the case. It would be interesting, therefore, to examine in a future study whether these or other factors predict subjects' selection of sentences that match the diagram or the meaning of English. Consistency within subjects would suggest that selections were made on the basis of some individual characteristic, preference, or strategy. The idiosyncrasy of responses identifying sources of task difficulty supports this notion.

Second, the lack of difference between confidence ratings found for subjects seeing nonsense diagrams in the first experiment was confirmed. Subjects who selected appropriate interpretations were no more confident than those who selected inappropriate interpretations, perhaps because they did not realize that their selections were influenced by the conventions of diagrams. This finding could also suggest that subjects who selected inappropriate interpretations—for example, those who interpreted the cause to be on the right—had just as good reasons for doing so as subjects who selected appropriate interpretations. This study did not determine what these reasons might be, however.

CONCLUSION

The results of these two experiments permit two general conclusions of relevance to graphic and instructional designers. First, the consistent and heavily biased preferences for particular interpretations of the three diagram formats illustrating the three relationships suggest that the arrangements of diagrams studied here are interpreted consistently. One explanation for this is that diagrams are read in the same way as text. Conventions have developed among English-speakers that interpret the word read first as the "subject" of the diagram. However, this explanation may not work when category membership is illustrated by placing the superordinate category above the subordinate one. Further study of this format is needed.

Second, violations of the conventional arrangement of diagrams consistently lead to ambiguity and lower confidence in interpretation, but only when the diagrams represent known concepts. An obvious implication is that designers should strictly adhere to established conventions for the spatial arrangement of diagrams. A less obvious implication is that, when the material is unfamiliar to students, as it is when they are learning it for the first time, inappropriate spatial arrangements of concepts in diagrams will probably lead to misinterpretation of the material. When the sense of the material does not dictate otherwise, students will be biased by the conventions of the diagram. When diagramming material that is unfamiliar to students, the designer should therefore be particularly careful to (1) place causes to the left of or above effects; (2) include

superordinate categories within superordinate ones rather than express category membership by horizontal and vertical arrangements; (3) place attributes to the right of, below, or inside the objects that possess them; and (4) in general, follow a left-to-right, top-to-bottom sequence that corresponds to English word order.

One limitation of this study was that it deliberately looked at the simplest possible ways of arranging concepts. The next step is to examine how the biases identified in this study might change when different arrangements are used, or when more than two concepts are shown in each diagram. Another promising area of study is individual differences in susceptibility to influence by the arrangement of diagrams or by the meaning of English when the two are at variance.

Overall, though, it is clear that the effectiveness of diagrams can be attributed to the way particular spatial arrangements are used by the designer. The designer's role in linking concepts in direct and concrete ways, described by many of those who have developed text-mapping schemes and who use diagrams as adjuncts to text, is of critical importance. □

William Winn and Cliff Solomon are at the University of Washington in Seattle.

REFERENCES

Armbruster, B. B., & Anderson, T. H. (1982). *Idea mapping: The technique and its use in the classroom, or simulating the "ups" and "downs" of reading comprehension.* (Reading Education Report #36). Urbana, IL: University of Illinois Center for the Study of Reading.

Armbruster, B. B., & Anderson, T. H. (1984). Mapping: Representing informative text graphically. In C. D. Holley & D. F. Dansereau (Eds.), *Spatial learning strategies* (pp. 189–209). New York: Academic Press.

Berkowitz, S. J. (1986). Effects of instruction in text organization on sixth-grade students' memory for expository reading. *Reading Research Quarterly, 21,* 161–177.

Breuker, J. A. (1984). A theoretical framework for spatial learning strategies. In C. D. Holley & D. F. Dansereau (Eds.), *Spatial learning strategies.* New York: Academic Press.

Geva, E. (1983). Facilitating reading comprehension through flowcharting. *Reading Research Quarterly, 18,* 384–405.

Guri-Rozenblit, S. (1988a). Impact of diagrams on recalling sequential elements in expository texts. *Reading Psychology, 9,* 121–139.

Guri-Rozenblit, S. (1988b). The interrelationships between diagrammatic representations and verbal explanations in learning from social science texts. *Instructional Science, 17,* 219–234.

Guri-Rozenblit, S. (1989). Effects of a tree diagram on students' comprehension of main ideas in an expository text with multiple themes. *Reading Research Quarterly, 24,* 236–247.

Holley, C. D., & Dansereau, D. F. (1984). The development of spatial learning strategies. In C. D. Holley & D. F. Dansereau (Eds.), *Spatial learning strategies.* New York: Academic Press.

Holliday, W. G., Brunner, L. L., & Donais, E. L. (1977). Differential cognitive and affective responses to flow diagrams in science. *Journal of Research in Science Teaching, 14,* 129–138.

Hughes, R. E. (1989). *Radial outlining: An instructional tool for teaching information processing.* Doctoral dissertation, University of Washington College of Education, Seattle.

Johnson, D. D., Pittelman, S. D., Heimlich, J. E. (1986). Semantic mapping. *Reading Teacher, 39,* 778–783.

Koran, J., & Koran, M. (1980). Interaction of learner characteristics with pictorial adjuncts in learning from science text. *Journal of Research in Science Teaching, 17,* 477–483.

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

Prater, D. L., & Terry, C. A. (1988). Effects of mapping strategies on reading comprehension and writing performance. *Reading Psychology, 9,* 101–120.

Ruddell, R. B., & Boyle, O. F. (1989). A study of cognitive mapping as a means to improve summarization and comprehension of expository text. *Reading Research and Instruction, 29,* 12–22.

Salomon, G. (1979). *Interaction of media, cognition and learning.* San Francisco: Jossey-Bass.

Schewel, R. (1989). Semantic mapping: A study skills strategy. *Academic Therapy, 24,* 439–447.

Sinatra, R. C., Stahl-Gemake, J., & Berg, D. N. (1984, October). Improving reading comprehension of disabled readers through semantic mapping. *The Reading Teacher, 37,* 22–29.

Sinatra, R., Stahl-Gemake, J., & Morgan, N. W. (1986). Using semantic mapping after reading to organize and write discourse. *Journal of Reading, 18,* 4–13.

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology, 3,* 515–557.

Winn, W. D. (1982). The role of diagrammatic rep-

resentation in learning sequences, identification and classification as a function of verbal and spatial ability. *Journal of Research in Science Teaching, 19*, 79–89.

Winn, W. D. (1983). Perceptual strategies used with flow diagrams having normal and unanticipated formats. *Perceptual and Motor Skills, 57*, 751–762.

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

Winn, W. D., Li, T-Z., & Schill, D. E. (1991). Diagrams as aids to problem solving: Their role in facilitating search and computation. *Educational Technology Research and Development, 39*, 17–30.

Wong, T. H., & Kao, H. S. R. (1991). The development of drawing principles in Chinese. In J. Wann, A. M. Wing, & N. Sovik (Eds.), *Development of graphic skills*. New York: Academic Press.