

Effects of Student-Generated Diagrams versus Student-Generated Summaries on Conceptual Understanding of Causal and Dynamic Knowledge in Plate Tectonics

Janice D. Gobert^{1,*} John J. Clement²

¹*Graduate School of Education, Harvard University, Cambridge, Massachusetts 02138 and Department of Science Studies, Western Michigan University, Kalamazoo, Michigan 49008*

²*School of Education, Furcolo Hall, University of Massachusetts, Amherst, Massachusetts 01003*

Received 2 July 1997; first revision 20 January 1998; second revision 23 February 1998; accepted 20 March 1998

Abstract: This research examines the beneficial effects of student-generated diagrams versus student-generated summaries on conceptual understanding in the domain of plate tectonics. Fifty-eight Grade 5 students read a brief expository text about plate tectonics. During their reading of the text, students were asked to either draw diagrams, produce written summaries, or simply read the text (control). Conceptual understanding was measured by the diagrams and summaries which were generated during students' reading of the text, as well as by a posttest which assessed students' understanding of both spatial/static and causal/dynamic knowledge of the domain. Results indicated that the summaries generated during the reading of the text contained more domain-related information than the diagrams which were generated during the reading of the text. However, on the posttest measures, the diagram group outperformed both the summary and text only groups in terms of understanding both the spatial/static as well as causal/dynamic aspects of the domain. Results are discussed with regard to the differential effects that generating diagrams as compared to generating summaries or simply reading has on both on-line comprehension during reading and resulting conceptual understanding of the domain. © 1999 John Wiley & Sons, Inc. *J Res Sci Teach* 36: 39–53, 1999

Diagrams are widely used in virtually all science domains at all levels of education since they are thought of as an effective way of clarifying subject matter (Lowe, 1989). In school settings as well as in research, they are primarily presented as adjunct aids to text (Dean & Enemoh, 1983; Joseph & Dwyer, 1984; Mayer, 1989) to depict fundamental spatial information which is difficult to adequately describe in text due to its linear structure. Despite the modern

*Currently a Visiting Scholar at Graduate School of Education, Harvard University

Correspondence to: J.D. Gobert

Contract grant sponsor: National Academy of Education

Contract grant sponsor: NSF

Contract grant number: MDR 9150002

emphasis on interactionist and constructivist methods of learning, diagrams are often given to students with the assumption that the presence of the diagram alone should facilitate learning.

Unfortunately, there are several problems which arise when text and diagrams are used for learning, both in school settings and in research. First, many of the diagrams used in science have merely an illustrative rather than explanatory role. Explanatory diagrams are contrasted to illustrative diagrams such that explanatory diagrams seek to explain some aspect of the subject matter, whereas illustrative diagrams serve only to depict certain characteristics (Lowe, 1991). Second, contrary to the belief of science teachers and researchers (Kindfield, 1993), students often do not know how to use diagrams effectively (Gobert, 1994; Lowe, 1989), nor do they know what features are important in diagrams (Anzai, 1991). Finally, learning from a text and diagram may increase memory load (Sweller et al., 1990) or may require additional cognitive effort to integrate the two media (Hegarty, 1992; Yee, Hunt, & Pellegrino, 1991). The result of these problems is that the intended goal, which is for students to develop deeper understanding, is not met (Lowe, 1989).

In brief, the existing research on visual displays as adjunct aids to text has demonstrated that they facilitate both recall and comprehension (Levin, Anglin, & Carney, 1987; Mayer, 1989; Mayer & Gallini, 1990), but for higher-level, conceptual understanding to occur, diagrams should have an explanatory role in presentation (Levie & Lentz, 1982; Winn, 1987). Furthermore, conceptual understanding also requires the student to be able to build a meaningful and appropriate mental representation, i.e., a mental model, of the system being taught (Lowe, 1993).

In this research, rather than present students with scientific diagrams as adjunct aids to text, we sought to foster students' conceptual understanding and mental model construction by having students generate their own diagrams while reading an expository text about plate tectonics. Specifically, in this study we compared students who generated diagrams to those who generated written summaries or simply read the text (control). Briefly, our hypothesis was that having students construct their own diagrams for learning would facilitate conceptual understanding since constructing diagrams may facilitate inference making (Schwartz, 1993). Diagramming, we thought, was also a good strategy to promote model building because of the visual nature of the domain under inquiry. Furthermore, we hypothesized that drawing diagrams would be more beneficial for mental model construction and inferencing than generating summaries because previous research has shown that diagrams both permit inferences based on perceptual cues such as spatial adjacency (Larkin & Simon, 1987) and explicitly indicate structural relationships (Brueker, 1984) which are difficult from textual representations such as those generated in the written summary task. For these reasons, we hypothesized that having students generate their own diagrams would be more beneficial to generating written summaries because diagram drawing would facilitate understanding of the spatial/static as well as causal/dynamic aspects, and the integration of this information into a mental model.

This research builds on existing research requiring students to generate their own diagrams (Schwartz, 1993). Specifically, as presented in an earlier issue of this journal, Schwartz reported two studies in which he investigated whether students could construct diagrams to structure complex novel information, and demonstrated that this was possible. Although acknowledged as important in his research, he did not, however, assess whether the process of constructing diagrams would improve comprehension or problem-solving performance. This latter issue is directly assessed in the present study.

The domain chosen for this study was plate tectonics. This is a difficult topic to learn both because of the hidden explanatory mechanisms which are outside our direct experience, and because it involves several different types of knowledge including spatial, casual, and dynamic knowledge. Specifically, conceptual understanding in this domain requires understanding the

spatial arrangement of the various layers of the earth (i.e., spatial/static information) as well as understanding the causal and dynamic movements within these layers (core as causal in heating the mantle, convection currents forming in the magma, plates moving, crust breaking/buckling, etc.). In addition to acquiring these two types of knowledge (spatial/static and causal/dynamic), several concepts need to be integrated into a complex causal chain to build a rich mental model of the system. We refer to these mental models as causal models. From these models, predictions and inferences can be made about the system's behavior: in the case of plate tectonics, explaining or depicting volcanic eruption, sea floor spreading, earthquakes, mountain formation, and continental drift.

The tasks requested of the students during the reading of the text (diagram and summary conditions) were decomposed into a learnable progression of models. That is, first we requested that students depict or describe a static model of the interior layers of the earth; next, we requested that they depict or describe the movement and processes in the various layers (i.e., a model of the dynamic processes). Finally, we requested that students depict or describe two outcomes of plate tectonics—namely, mountain formation and volcanic eruption. By ordering the tasks in this manner we sought to scaffold the development of students' understanding. In the case of the diagram condition, it was believed that ordering the tasks in this manner would promote the development of rich diagrams (external representations) and mental models (internal representations). Furthermore, it is believed that the process of constructing a series of diagrams provides a means for progressively refining one's understanding of the domain. The order in which diagrams were requested, we hypothesize, puts an emphasis on students' understanding of increasingly complex causal relationships and facilitates the construction of progressively more complete and complex models. This strategy continues in the vein of current research on progressive model construction such as Raghavan and Glaser's MARS curriculum (1995), in which smaller models (e.g., of density and force addition) provide conceptual leverage for larger macromodels (e.g., of buoyancy). In addition, White and Frederiksen's progressions from simplified to more detailed and complex causal models of electricity (1990), and White's (1993) progression of increasingly more complex models in ThinkerTools for Newtonian mechanics are examples of pedagogical tools to promote progressive model construction. These strategies also draw support from research findings on expertise which has documented experts progressing from incomplete models to improved and more complete models through processes of model construction, criticism, and revision (Clement, 1989). In our study, students were first requested to construct a static/spatial model of the layers within the earth and then instructed to construct increasingly more complex diagrams in terms of the causal and dynamic information required. We hypothesize that students are able to make use of their spatial/static model as a base to build a causal/dynamic model, which in turn facilitates the construction of models requesting the depiction of two outcomes in the world (mountain formation and volcanic eruption). This is a finer-grained progression on a shorter time scale than the studies mentioned above which also sought to develop causal models (Raghavan & Glaser, 1995; White, 1993; White & Frederiksen, 1990), but it fits the same general themes of progressive model construction as a process which takes active work and time to achieve successfully.

The studies employed a text as a learning source. The text comprehension framework underlying the studies was van Dijk and Kintsch's text comprehension model (1983). As predicted by this model, simple recall and recognition tasks are best supported by a memory for the text itself—that is, a good text base. Higher-level inference tasks are best supported by representations which reflect higher-level, more integrated representations—that is, situation models (Kintsch, 1986) or mental models (Johnson-Laird, 1983). In accordance with van Dijk and Kintsch's theory, in the present study it was assumed that students' understanding is based on

an interaction of the processing induced by either diagram drawing or summarizing with the processing of the main passage itself. Thus, differential effects on learning would be predicted in the three different learning conditions: diagram drawing, summarizing, or read only. In this study, we sought to get an in-depth look at the learning which occurs in three different learning conditions by measuring understanding during both students' reading of the text, i.e., their on-line comprehension, as well as by means of a posttest, i.e., their resulting conceptual representations or mental models.

Method

Subjects

Three classes of Grade 5 students participated, 58 students in total (34 boys and 24 girls). The students ranged in age from 10 to 12 years. Students were drawn from a small town in western Massachusetts.

Procedure

The three groups of students were each given a short expository text (approximately two pages) about plate tectonics. For each of the intervention groups, diagrams or summaries were requested at specific points in the text; four diagrams or summaries in total were requested during the reading of the text. The control group read the text only; that is, no tasks were requested while the text was being read. For the intervention groups, prompts were given in the text (prior to its respective section of text) to inform the students about the task that would be requested after each particular section of the text. An example of one of the prompts for each of the diagram and summary conditions is: "After this paragraph you will be asked to draw a picture (on the next page) of the different layers of the earth," or "After this paragraph you will be asked to describe the different layers of the earth." Then, as mentioned above, the student's diagram or summary was requested. The wording used to request drawings and summaries was highly similar, e.g., "Thinking back to what you just read, draw a picture of the different layers of the earth. Include and label all the information about these layers that you can" or "Thinking back to what you just read, describe in words the different layers of the earth. Include all the information about these layers that you can."

As previously mentioned, the diagram and summary tasks were requested in order of increasing difficulty in terms of the causal and dynamic knowledge to be integrated. First, the students were requested to either depict or describe spatial/static information only (i.e., the different layers of the earth). Second, they were asked to either depict or describe the movement and processes which occur in the layers. Third, they were asked to depict or describe what happens in the layers of the earth when mountains are formed. Finally, they were asked to depict or describe what happens in the layers of the earth when volcanoes erupt.

After students finished reading the text (and the respective tasks, as in the diagram and summary conditions), they were each given a booklet of questions containing multiple choice and short answer questions, as well as diagram tasks (both student generated and one provided in the test booklet). All questions were designed to assess either knowledge of spatial/static aspects of the domain or causal/dynamic aspects of the domain. Examples of questions assessing spatial/static knowledge are: "Where is the thinnest part of the crust?" and "If the continents were all together, would the rest of the earth be water?" Examples of questions assessing causal/dynamic knowledge are: "the movement in the crust of the earth is caused by . . . ?" and

“Rock from the floor of the Atlantic Ocean tests to be younger than rock from the middle of the North American continent because . . . ”

For this study, there are two sets of quantitative data: namely, their performance on the intermittent tasks, i.e., the diagrams and summaries that were generated at intermittent points during students' reading of the text (for those in the intervention conditions), and students' performance on the posttest (for those in all three conditions). Each will be discussed in turn.

Coding of Intermittent Diagrams and Summaries

Coding schemes were developed for each of the four intermittent tasks which the students generated during their reading of the text (diagrams or summaries). For each task, the coding scheme was used to evaluate the inclusion of main points as in the case of the summaries, or the inclusion of the main features as in the case of the diagrams. The total possible score for each of the four tasks are: Task 1, 12 points; Task 2, 13 points; Task 3, 16 points; and Task 4, 14 points.

Coding of Posttest

For the posttest data, the short answer questions, seven multiple choice questions, and diagram task provided in the test booklet were simply scored. The student-generated diagrams on the posttest were evaluated using coding schemes similar to those used for the diagrams generated during students' reading of the text. The total possible score for spatial/static knowledge on the posttest was 25; the total possible score for causal/dynamic knowledge was 45. All subsequent analyses were performed using these two types of knowledge as dependent variables.

To check for interrater reliability for the posttest, a second coder was trained on the diagram scoring technique since these were the items on the posttest which were subjectively scored. The second coder scored all of the students' posttest diagrams; a interrater reliability score of 93.8% was obtained indicating a high degree of agreement between the experimenter and the second coder.

Results

Analysis of Intermittent Data

The data were analyzed using SPSS for Macintosh (Language System Corporation, 1991). For the analysis of the intermittent data, multivariate analysis of variance is recommended for these data since it takes into account possible correlations among the four dependent measures, allowing for a clearer interpretation of the results (Stevens, 1986).

An overall multivariate analysis of variance was conducted with the four intermittent tasks entered together as variables, comparing the summaries and diagrams for informational content (note that the read-only group did not generate summaries or diagrams; thus, this group was not included in this analysis). An overall significant multivariate effect was obtained, $F = 5.718$ (Wilks), $p < .001$, indicating differences between the two groups (i.e., diagram vs. summary). Univariate tests indicated that multivariate results were primarily due to differences for the first two intermittent tasks: namely, the depiction or summary of “layers of the earth,” $F = 23.57$, $p < .001$, and that of “movement and processes in the layers of the earth,” $F = 4.87$, $p < .05$; these results are summarized in Table 1. As can be seen in Figure 1(a), the summary group's

Table 1
*Univariate analysis of variance for four intermittent tasks
 generated during reading of text*

Variable	Hypoth. MS	Error MS	<i>F</i>	<i>p</i>
Task 1 (layers of earth)	121.773	5.17	23.57	.000**
Task 2 (movement in layers)	62.146	12.76	4.87	.033*
Task 3 (mountain formation)	14.339	5.63	2.55	.118
Task 4 (volcanic eruption)	.636	5.40	.118	.773

Univariate *F* tests with (1, 40) *df*.

p* < .05 *p* < .001.

summaries contained more semantic information than did the diagram group's diagrams on all four tasks generated during the reading of the text. The differences between the summary and the diagram group were largest for the first two tasks: namely, the layers of the earth, and the movement and processes in the layers, as indicated by statistical significance on both the multivariate and the univariate tests.

A second analyses of the intermittent data was done in which the data across the four tasks were divided and pooled by the type of knowledge being assessed. That is, the first intermittent task assessed only spatial/static knowledge, whereas the second, third, and fourth tasks assessed primarily causal/dynamic knowledge. The data were grouped into their respective knowledge category, i.e., spatial/static and causal/dynamic, and analyzed a second time to get a clearer sense of students' knowledge during their reading of the text in terms of the spatial/static versus causal/dynamic variables, thus reducing the data from four intermittent measures to two dependent variables. A second multivariate analysis of variance was then conducted.

An overall significant multivariate effect was obtained, $F = 11.823$ (Wilks), $p < .001$, indicating differences between the two groups (i.e., diagram and summary) on their spatial/static and causal/dynamic knowledge during their reading of the text. Univariate tests indicated a significant difference between the groups favoring the summary group for their inclusion of spatial/static knowledge in their summaries, $F = 23.57$, $p < .001$. Although not statistically significant at the .05 level of alpha on the univariate tests, the summary group was also found to outperform

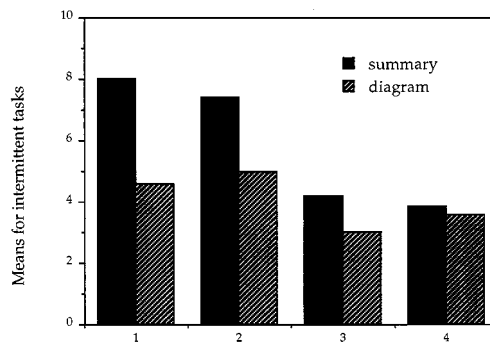


Figure 1. (a). Group comparison of summaries and diagrams generated during reading of text.

Table 2
Univariate analysis of variance for spatial/static and causal/dynamic knowledge during reading of text by group (summary or diagram)

Variable	Hypoth. MS	Error MS	<i>F</i>	<i>p</i>
Spatial/static (Task 1)	121.773	5.17	23.57	.000*
Causal/dynamic (Task 2, 3, 4 pooled)	155.428	42.41	3.67	.063 (trend)

Univariate *F* tests with (1, 40) *df*.

**p* < .001.

the diagram group on their inclusion of causal/dynamic knowledge in their summaries, $F = 3.67$, $p = .063$. These results are summarized in Table 2 and are depicted in Figure 1(b).

Analysis of Posttest Data

To compare the resulting conceptual knowledge of the students or the mental model representations as measured by the posttest, a multivariate analysis of variance was conducted with static/spatial knowledge and causal/dynamic knowledge entered together as dependent variables. Again, by entering the two dependent variables in a multivariate analysis of variance, any statistical commonality between the two measures is accounted for, allowing for a clearer interpretation of the results for each dependent variable.

An overall significant multivariate effect was obtained, $F = 2.46$ (Hotellings), $p = .049$, indicating differences for the two groups (diagram vs. summary). Univariate tests indicated that there were statistically significant differences between the groups for both static/spatial knowledge, $F = 4.388$, $p = .017$, and causal/dynamic knowledge, $F = 4.307$, $p = .018$. These data are summarized in Table 3 and depicted in Figure 2.

Differences in Spatial/Static Knowledge at Posttest

As mentioned above, the univariate test indicated that there were statistically significant differences among three groups (diagram, summary, and read-only) for the amount of spatial/sta-

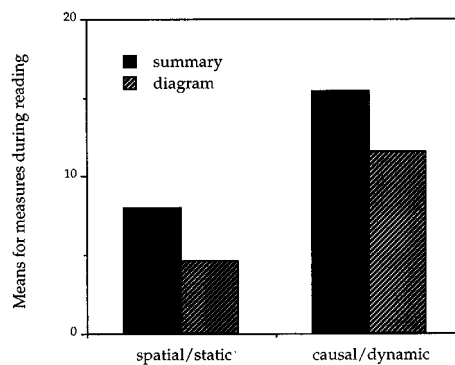


Figure 1. (b) Spatial/static and causal/dynamic knowledge reflected in summaries or diagrams generated during reading.

Table 3
Univariate analysis of variance for spatial/static and causal/dynamic knowledge at posttest, by group (read-only, summary and diagram)

Variable	Hypoth. MS	Error MS	<i>F</i>	<i>p</i>
Spatial/static	87.384	19.910	4.389	.017*
Causal/dynamic	156.584	36.351	4.308	.018*

Univariate *F* tests with (2, 56) *df*.

**p* < .05.

tic knowledge they had acquired about the domain, $F = 4.39$, $p = .017$. Follow-up contrasts were performed to compare (a) diagram group to the summary group, and (b) the summary group to the read-only group. These planned contrasts were structured in this way because it was hypothesized that the diagram group would outperform the summary group and that the summary group would outperform the read-only (control) group.

For the set contrasts assessing differences in spatial/static knowledge, a one-way analysis of variance was performed comparing (as previously stated) the diagram group to the summary group and the summary group to the read-only group (control). Although not strong enough to reach statistical significance at the .05 level of alpha, a trend in the expected direction was found; that is, the diagram group was found to be superior to the summary group in its knowledge about the spatial/static aspects of the domain, $t(56) = -1.83$, $p = .07$. For the second planned contrast, there were no significant differences found between the summary group and the read-only group, $t(56) = -1.29$, $p = .20$, in their knowledge about the spatial/static knowledge they had acquired. Table 4 summarizes the means and standard deviations. Figure 2 depicts these results.

Differences in Causal/Dynamic Knowledge at Posttest

As mentioned previously, the univariate test indicated that there were statistically significant differences among three groups (diagram, summary, and read-only) for the amount of

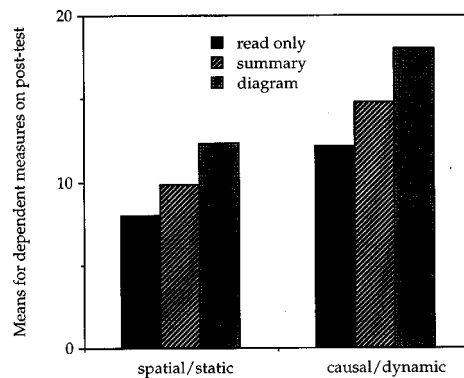


Figure 2. Group comparison of spatial/static knowledge and causal dynamic knowledge at posttest.

Table 4
Group means (M) and standard deviations (SD) for understanding of spatial/static aspects of domain

Group (Sample Size)	<i>M</i>	<i>SD</i>
Read-only (control) (<i>n</i> = 17)	7.94	4.5
Summary group (<i>n</i> = 23)	9.78	4.5
Diagram group (<i>n</i> = 19)	12.32	4.4

causal/dynamic knowledge they had acquired about the domain, $F = 4.31$, $p = .018$. Again, follow-up contrasts were performed to compare the diagram group to the summary group, and the summary group to the read-only group, reflecting the hypotheses of the study.

For the set contrasts assessing differences in causal/dynamic knowledge, a one-way analysis of variance was performed comparing (as previously stated) the diagram group to the summary group and the summary group to the read-only group (control). Although not strong enough to reach statistical significance at the .05 level of alpha, a trend in the expected direction was found; that is, the diagram group was found to be superior to the summary group in its knowledge about the causal/dynamic aspects of the domain, $t(56) = -1.72$, $p = .09$. For the second planned contrast, there were no significant differences found between the summary group and the read-only group, $t(56) = -1.38$, $p = .17$ in their knowledge about the causal/dynamic knowledge they had acquired. Table 5 shows the means and standard deviations. Figure 2 depicts these results.

Summary and Discussion of Results

What is interesting about these results is the discrepancy between the learners' understanding of the domain during their reading of the text, as measured by the intermittent diagrams and summaries, and the learners' resulting conceptual understanding or mental models of the domain, as assessed by the posttest. To review, on tasks which were generated during students' reading of the text, those who generated summaries outperformed those who generated diagrams on all four intermittent measures (two of which reached statistical significance at both the multivariate and univariate levels). When the intermittent data were pooled into spatial/static and causal/dynamic variables and reanalyzed, the summary group was found, again, to outperform the diagram on both measures. On the posttest which was designed to measure students' mental models or resulting conceptual representations of the domain, an opposite pattern was found;

Table 5
Group means (M) and standard deviations (SD) for understanding of causal/dynamic aspects of domain

Group (Sample Size)	<i>M</i>	<i>SD</i>
Read-only (control) (<i>n</i> = 17)	12.11	5.2
Summary group (<i>n</i> = 23)	14.78	6.4
Diagram group (<i>n</i> = 19)	18.00	6.3

that is, the diagram group outperformed the summary group for both the understanding of spatial/static as well as causal/dynamic aspects of the domain.

These data were interpreted as follows. The task of generating summaries supported the formation of a good text-base representation, i.e., a good representation of the text itself (van Dijk & Kintsch, 1983) from which domain-related information could be easily recalled in generating the summaries. This point is supported by the finding that the summaries contained more domain-related information than did the diagrams. However, the task of generating summaries did not support the development of higher-level, mental model representations from which rich inferences could be made, as evidenced by lower scores on the posttest than the diagram group. Hence, for the summary group, the task of generating summaries appears to have been like a rote recall task which did not promote the construction of richer mental model representations. The task of generating diagrams during the reading of the text, on the other hand, supported the formation of rich mental model representations from which greater inferencing about the domain could be done, as evidenced by better understanding elicited by the diagram group for both the spatial/static as well as the causal/dynamic aspects of the domain on the posttest. However, generating diagrams did not support the development of a verbatim-type text-base representations, i.e., memory for the text itself, from which information could be easily recalled to construct diagrammatic models.

It also could be hypothesized that for the summary group, since the media were the same (i.e., text to text), this did not promote a deeper processing of the textual material (Craik & Lockhart, 1972; Lockhart & Craik, 1990). Thus, the spatial/static knowledge of the different layers of the earth was not as well remembered, perhaps because providing summaries did not support the development of a static mental model from which the required spatial information could be easily read, or from which further inferences about the causal and dynamic processes could be made (Larkin & Simon, 1987). In the case of those who drew diagrams, it is hypothesized that they had to represent their knowledge into a diagrammatic format, and that they could not solely rely on rote memory of the text to do this. Rather, the diagramming tasks which were prompted before students read the relevant section of the text promoted a deeper, more variable processing of the text which elicited inference making and resulted in richer mental models of the domain (Craik & Lockhart, 1972; Lockhart & Craik, 1990; van Dijk & Kintsch, 1983).

Conclusions

Text-Base Representations and Mental Model Representations

In this study, we hypothesized that the task of generating diagrams while reading would promote richer mental model construction than the task of generating summaries or simply reading the text. This hypothesis was confirmed; that is, those who generated diagrams outperformed those who generated summaries in terms of their resulting conceptual representations of the domain. This is interesting in light of the fact that the summary group's summaries were found to contain more domain-related information than did for the tasks generated during students' reading of the text, the diagram group's diagrams. These findings are consistent with van Dijk and Kintsch's (1983) theory of the text comprehension, as well as studies which have shown that learner's representations of text material can be altered by changing their goals for learning (Schmalhofer & Glavanov, 1986). These findings can also be attributed to the possible advantages afforded by having the students represent their knowledge in a visual medium (as in the diagram condition) that were not afforded by the textual medium (as in the summary or read-

only conditions) (Larkin & Simon, 1987). For students who generated summaries, these written descriptions served as externalized representations of their knowledge from which they were able to make some inferences, but our data suggest that this was a more difficult process from a textually based representation than a visually based representation (as in the diagram condition).

Drawing to Learn and Children's Causal Model Construction

Our goal was to foster causal model construction in students, building upon existing research on progressive model construction and diagram construction in science (White, 1993; Raghavan & Glaser, 1995; Schwartz, 1993). We refer to our approach as “drawing to learn,” since the students learn new material by generating their own diagrams in an order of increasing difficulty in terms of the causal and dynamic information to be integrated. Drawing to learn may be effective for a variety of learners, since research has shown that even unskilled readers are able to make rich inferences from diagrams (Holmes, 1987).

In the existing literature, questions were raised whether students can be taught to produce diagrams from which inferences can be drawn (Anzai, 1991), as well as whether students will be able to draw inferences from their own diagrams once they have been constructed (Schwartz, 1993). In addition, there has been a great deal of debate about whether children are able to formulate models, as opposed to only fragmentary knowledge structures (di Sessa, 1985, 1993), and whether children can make inferences on the basis of these models (Vosniadou, 1989). Our results suggest that young students, i.e., fifth graders, can construct rich diagrams of complex spatial, causal, and dynamic systems which they can use to make inferences about the topic under inquiry. Drawing strategies such as those used here may be particularly effective in domains such as plate tectonics in which students cannot directly observe the processes occurring. In this case, students must construct internal representations of processes that involve complex spatial relationships, and depict these relationships in external representations in the form of diagrams. Results from this study suggest that the diagrammatic medium facilitates model construction processes; perhaps because memory and learning advantages are afforded by spatial explicitness of diagrams (Yates, 1978). In particular, we hypothesize that in depicting the layers of the earth in a visual medium provides perceptual cues of spatial adjacency to inferences about causal and dynamic knowledge; e.g., seeing that the heated core is spatially adjacent to the mantle made of liquid magma allows for the inference that the magma heats up, and seeing that this magma is adjacent to the plates allows for the inference that the heated magma causes plate movement (Gobert & Clement, 1994). That is, the diagram, once constructed, acts as a guide for “chains of inference” (Larkin & Simon, 1987; Schwartz, 1993). These results suggest that having students generate diagrams can improve comprehension of targeted material.

As previously mentioned, our students constructed diagrams in order of increasing complexity, from a spatial/static model first, followed by increasingly complex models in terms of the causal and dynamic information to be integrated. This order appeared to facilitate conceptual understanding and mental model construction. Although we did not manipulate the order of diagrams requested, the results obtained are compatible with current science education research in which students were also asked to generate increasingly complex causal models which has shown that simpler models provide conceptual leverage for understanding increasingly complex causal models (Raghavan & Glaser, 1995; White, 1993). Based on these data, we hypothesize that model construction via diagramming may be enhanced when students construct static models first, and then enrich these models through constructing diagrams of causal and dynamic information. The key element here appears to be designing and requesting diagramming tasks in an order that will facilitate model construction.

Implications of Further Research

An empirical question remains as to whether another type of inference task, i.e., a nondiagrammatic task, during the reading of the text would elicit similar performance as that of the diagram group. Questions of this nature have been raised, and conjectures have been made that drawing diagrams is likely to elicit similar types of inferential power as providing explanations (Chi et al., 1989, 1994). However, we argue that it is the diagrammatic medium in particular rather than inferencing in general which is supporting the types of knowledge gains made by the diagram group. In addition to promoting a deeper processing of the text, we propose that depicting their knowledge in a visual medium in particular facilitated conceptual understanding and causal model construction in this domain. A study which employs explanation-based tasks versus diagram-based ones is currently in progress to test this hypothesis empirically (Gobert, 1997). If borne out—that is, if the diagram group outperforms the explanation group—these results would support the notion of the visual media affording specific advantages over the textual media, irrespective of the inferencing that is being elicited in both conditions. Finally, further research is needed to delineate the modality-specific processes that are employed in setting up mental models on the basis of diagram-based tasks as well as textually oriented tasks such as summarization and explanation.

Implications for Instruction

One possible implication for classroom practice is to have students generate their own diagrams for learning rather than giving students diagrams to learn from. In this way, diagramming may become a general problem-solving strategy for science students. The instructional techniques used for the purposes of this study were not elaborate and therefore would be relatively simple to implement in a classroom science class. Certainly, many other useful techniques could be used in a fuller, longer-term unit. However, the instructional techniques used here were designed specifically to empirically compare the effects of diagram drawing and generating summaries on conceptual understanding.

The results of this study raise the question whether the drawing to learn strategy might be useful in other contexts such as in small-group activities or large-group discussions with the student-generated drawings used as the focus for discussion (Scott, 1987). Although not directly evaluated in this study, our results also suggest that having students construct drawings of events and processes rather than simply static pictures as illustrations may be crucial to understanding causal and dynamic processes and thus also may be crucial to model building. Laboratory exercises, for example, often emphasize drawing pictures of lab set ups or observed final states, but rarely request pictures of observable events, much less pictures of unobservable mechanisms.

The issue of depicting dynamic processes invites another topic of investigation—that is, notational strategies for representing movement and change in diagrams. Arrows were used in a natural way by the students in this study to represent movement of magma and plates. Encouraging discussions about drawings that lead to evaluations and eventual agreements about effective notations for dynamic events may have significant payoffs for teaching science (Scott, 1987; Pea, 1994).

Concluding Comments

This approach to science education examines both the process and product of science learning. It uses diagrams as important tools for reasoning and model construction as opposed to the

more conventional use of diagrams as merely illustrations of science concepts. These results, in accordance with other research on diagrams (Lowe, 1989, 1991), suggest that student-generated diagramming should be evaluated as potentially an integral part of the science curriculum. Our findings are pertinent to inform instructional and assessment practices which value visual components of understanding in science. Further efforts are needed to investigate other ways of capitalizing on students' ability to build causal models through diagramming and to encourage students to transfer the idea of mediating their learning with diagrams.

This research was supported in part by a Spencer Postdoctoral Fellowship from the National Academy of Education, awarded to the first author; and in part by a grant from the National Science Foundation (MDR 9150002), awarded to the second author.

References

- Anzai, Y. (1991). Learning and use of representations for physics expertise. In K.A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits*. Cambridge, MA: MIT Press.
- Brueker, J.A. (1984). A theoretical framework for spatial learning techniques. In C. Holley & D. Dansereau (Eds.), *Spatial learning strategies*. New York: Academic Press.
- Chi, M.T.H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13, 145–182.
- Chi, M.T.H., de Leuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18, 439–478.
- Clement, J. (1989). Learning via model construction and criticism. In Glover, G., Ronning, R., & Reynolds, C. (Eds.), *Handbook of creativity: Assessment, theory and research*. New York: Plenum
- Craik, F.I.M., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671–684.
- Dean, R.S., & Enemoh, P.A. (1983). Pictorial organization in prose learning. *Contemporary Educational Psychology*, 8, 20–27.
- di Sessa, A. (1985). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age*. Hillsdale, NJ: Erlbaum.
- di Sessa, A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10, 105–226.
- Gobert, J. (1994). Expertise in the comprehension of architectural plans: Contribution of representation and domain knowledge. Unpublished doctoral dissertation, University of Toronto, Toronto, Ontario.
- Gobert, J. (1997, August). Summarizing, explaining, and diagramming: The differential effects on text-base representations and mental models. *Proceedings of the 19th annual meeting of the Cognitive Science Society*. Stanford University, Palo Alto, CA.
- Gobert, J., & Clement, J. (1994, April). Promoting causal model construction in science through student-generated diagrams. Presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Hegarty, M. (1992). The mechanics of comprehension and the comprehension of mechanics. In K. Rayner (Ed.), *Eye movements and visual cognition*. New York: Springer-Verlag.
- Holmes, B. (1987). Children's inferences with print and pictures. *Journal of Educational Psychology*, 79, 14–18.

- Johnson-Laird, P.N. (1983). *Mental models: Toward a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Joseph, J.H., & Dwyer, F.M. (1984). The effects of prior knowledge, presentation mode, and visual realism on student achievement. *Journal of Experimental Education*, 52, 110–121.
- Kindfield, A. (1993). Biology diagrams: Tools to think with. *Journal of the Learning Sciences*, 3, 1–36.
- Kintsch, W. (1986). Learning from text. *Cognition and Instruction*, 3, 87–108.
- Language Systems Corporation. [Computer Software]. (1991). SPSS for the Macintosh, Version 4.0. SPSS Inc., Chicago, IL.
- Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65–100.
- Levie, W.H., & Lentz, R. (1982). Effects of text illustrations. *Educational Communication and Technology Journal*, 30, 195–232.
- Levin, J., Anglin, G., & Carney, R. (1987). On empirically validating the functions of pictures in prose. In D.M. Houghton & H.A. Willows (Eds.), *The psychology of illustration*. New York: Springer-Verlag.
- Lockhart, R., & Craik, F.I.M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology*, 44, 87–112.
- Lowe, R. (1989, October). Scientific diagrams: How well can students read them? In *What Research Says to the Science and Mathematics Teacher*, no. 3. Key Centre for School Science and Mathematics, Curtin University of Technology, Perth, Australia.
- Lowe, R. (1991). Expository illustrations: A new challenge for reading instruction. *Australian Journal of Reading*, 14, 215–226.
- Lowe, R. (1993). Constructing a mental representation from an abstract technical diagram. *Learning and Instruction*, 3, 157–179.
- Mayer, R. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology*, 81, 240–246.
- Mayer, R., & Bromage, B. (1980). Different recall protocols for technical texts due to advance organizers. *Journal of Educational Psychology*, 72, 209–225.
- Mayer, R., & Gallini, J. (1990). When is an illustration worth then thousand words? *Journal of Educational Psychology*, 82, 715–726.
- Pea, R.D. (1994). Seeing what we build together: Distributed multimedia learning environments for transformative communications. *The Journal of Learning Sciences*, 3, 285–299.
- Raghavan, K., & Glaser, R. (1995). Model-based analysis and reasoning in science: The MARS curriculum. *Science Education*, 79, 37–61.
- Schmalhofer, F. & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional and situational representations. *Journal of Memory and Language*, 25, 279–294.
- Schwartz, D. (1993). The construction and analogical transfer of symbolic visualizations. *Journal of Research in Science Teaching*, 30 1309–1325.
- Scott, P. (1987). The process of conceptual change in science: A case study of the development of a secondary pupil's ideas relating to matter. In *Proceedings of the Second International Seminar Misconceptions and Educational Strategies in Science and Mathematics*, Vol. 2, 404–419.
- Stevens, J. (1986). *Applied multivariate statistics for the social sciences*. Hillsdale, NJ: Erlbaum.
- Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology*, 119, 176–192.

van Dijk, T., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Vosniadou, S. (1989). Analogical reasoning as a mechanism in knowledge acquisition: A developmental perspective. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*. Port Chester, NY: Cambridge University Press.

White, B. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction*, 10, 1–100.

White, B., & Frederiksen, J. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, 24, 99–157.

Winn, W. (1987). Charts, graphs, and diagrams in educational materials. In D.M. Houghton & H.A. Willows (Eds.), *The psychology of illustration*. New York: Springer-Verlag.

Yates, F. (1978). *The art of memory*. Harmondsworth, Middlesex: Penguin.

Yee, P., Hunt, E., & Pellegrino, J. (1991). Coordinating cognitive information: Task effects and individual differences in integrating information from several sources. *Cognitive Psychology*, 23, 615–680.