Effects of discourse structure graphic organizers on EFL reading comprehension

Xiangying Jiang
West Virginia University
United States

Abstract

This study investigated the effects of a 16-week reading instruction program with discourse structure graphic organizers (DSGOs) on the development of English reading comprehension among college-level English as a Foreign Language (EFL) students. A total of 340 first and third semester students of non-English majors at a Chinese university participated in this study. A DSGO completion test and a TOEFL (Test of English as a Foreign Language) reading comprehension test were administered before, immediately after, and 7 weeks following the instructional treatment. The results showed that the DSGO instruction significantly improved discourse comprehension as measured by the DSGO completion task, and the effect was retained 7 weeks after the instructional treatment. Significant improvement was also observed in the general reading ability as measured by TOEFL reading comprehension in the immediate posttest, but the effect did not persist in the delayed posttest. These findings apply to both the first and third semester students. Pedagogical implications of the DSGO instruction are discussed.

Keywords: discourse structures, discourse structure awareness, graphic organizers, discourse structure graphic organizers, second language reading instruction

In reading research, discourse structure awareness is perceived as an important component of a reader's overall comprehension abilities (Pearson & Fielding, 1991; Trabasso & Bouchard, 2002). Discourse structures, also known as text structures, are defined as the "knowledge structures or basic rhetorical patterns in texts" (Grabe, 2003, p. 9). They are the frameworks that writers employ to convey information in an organized and coherent manner. As a result, texts are generally organized following certain basic rhetorical patterns (Hoey, 2001; Mohan, 1986; Singer, 1990; Van Dijk & Kintsch, 1983).

Discourse structures are relatively few in number. They recur regularly across texts and are often found in various combinations (Grabe, 2003; Meyer, 1985; Mohan, 1986). These structures include comparison-contrast, cause-effect, problem-solution, definition, classification, argument, description, procedure, and narrative episodes. The finite number of discourse structures and their recurrence across texts make it possible to explicitly teach these structures. When students become knowledgeable about the limited ways that texts are organized, "they will be able to

better understand the coherence and logic of the information being presented, and they will be able to locate the main ideas and distinguish them from less important information" (Grabe & Gardner, 1995, p. 78). Most of all, students will be able to transfer this knowledge across texts and content areas.

Studies on expository text comprehension have demonstrated that discourse structure awareness plays an important role in reading comprehension in the first language (L1) and the second language (L2) settings. The acquisition of discourse structures "occurs only through formal training and substantial reading experience" (Koda, 2005, p. 139). Therefore, pedagogical efforts should be made to raise students' awareness of these discourse structures. One way to translate discourse structures from texts to classroom instruction is through the use of graphic organizers (GOs) *that represent the discourse structures of the text*. Discourse structure graphic organizers (DSGOs) are instructional techniques that display discourse structures and content information visually and hierarchically (Simmons, Griffin, & Kameenui, 1988). They depict the discourse structures by representing the interrelationships among ideas and patterns of the text.

It is important to distinguish the term DSGOs used in this paper from GOs of a generic format, variably known as Venn diagrams, semantic maps, outlines, and t-bars, that do not specifically represent the discourse structures of a text. A literature review of the reading instruction using GOs has suggested that GOs can be classified into two major types: those that directly represent the discourse structures of a text (DSGOs) and those that do not represent the discourse structures of a text (Jiang & Grabe, 2007). Moreover, DSGOs provide more consistent and stronger evidence for their effectiveness than the generic type. This paper focuses exclusively on DSGOs, though the term GOs is used in the literature review to be consistent with the original sources.

Review of Literature

In L1 reading research, a great number of studies have explored the effects of GO instruction on discourse structure awareness and reading comprehension. Appendix A provides detailed information on a number of these studies on DSGOs. For detailed information on the remaining studies reviewed here, readers may refer to Appendix B in Jiang and Grabe (2007). It should be noted that in the following literature review, GO is used interchangeably with such terms as flowchart, tree-diagram, frame, or matrix. The literature review to follow highlights key research studies on the impact of DSGOs in both L1 and L2 contexts.

Geva (1983) trained first year L1 community college students to represent text structures in node-relation flowcharts. She found that learning to recognize text structure through flowcharting transferred positively to the comprehension of expository texts by less skilled readers. Armbruster, Anderson, and Meyer (1991) reported the effectiveness of a particular type of GO called a "frame" in helping the fourth and fifth grade L1 students study their social studies texts. The combined analyses of recall and recognition measures showed that students in the framing condition scored higher than students in the control condition. Guri-Rosenblit (1989) investigated the effectiveness of using a tree diagram in helping Israeli L1 college students understand the main ideas in an expository text in social sciences. The results demonstrated that

students who received the tree diagram performed significantly better in understanding the main ideas and on the recall of the relations between various elements in the text than those who received either the original or the elaborated text without a diagram. Armbruster, Anderson, and Ostertag (1987) investigated the effects of GO instruction on students with varying abilities and provided evidence that GO instruction benefits students of both stronger and weaker ability groups. The GO groups recalled about 50% more macrostructure ideas than the traditional groups. The training was effective for high, medium, and low ability students alike. Alvermann and Boothby (1986) found that length of GO treatment was an important variable in determining the amount of GO transfer. The study showed that students in the 14-day GO group comprehended and recalled significantly more information than the comparison group, whereas there was no reliable difference between the 7-day GO group and the comparison group.

In the mid 1990s, two studies by Robinson and colleagues reinforced earlier research findings. Robinson and Kiewra (1995) found that students studying GOs learned more hierarchical and coordinate relations than students who studied outlines or the text alone. They were also more successful in applying that knowledge to essay writing. Robinson and Schraw (1994) compared the computational efficiency of matrix (GO), outline, and text reading alone. They found that studying a matrix enabled college-level readers to grasp the conceptual relations in a text more efficiently when compared to studying an outline or the text alone. However, the advantage of this matrix group over the other two groups disappeared in a delayed test. The researchers interpreted this phenomenon as indicating that students in the matrix group did not make enough effort to learn the conceptual relations, ironically, due to the efficiency of the matrix representation.

More recently, Williams et al. (2005) reported that children were sensitive to text structure in expository passages as early as second grade. A program, which included the use of a matrix (GO) as an instructional technique to teach compare-contrast text structure, improved students' abilities to comprehend compare-contrast texts and enabled them to transfer their newly learned skills to uninstructed compare-contrast texts. Williams and colleagues (2007) extended the findings of Williams et al. (2005) to the content area of social studies and found that the explicit comprehension instruction, which consisted of the use of GOs for each cause-effect text structure, improved the comprehension of instructional cause-effect texts.

Although the majority of the studies have consistently demonstrated the beneficial effects of GOs in reading instruction, a small number of studies have produced controversial or inconclusive findings. Armbruster, Anderson and Meyer (1991) found that GOs had positive effects on the study of social science texts among the fifth graders, but not among the fourth graders. Hoffmann (2010) investigated the effect of using a matrix diagram on the comprehension of compare-contrast texts among a group of fifth graders. The GO group showed improvement in neither comprehension scores nor test scores on a standardized test of reading comprehension over a course of six weeks. Despite these controversies, the majority of GO studies consistently demonstrated beneficial effects of GOs in reading instruction. In a review article, Robinson (1997) reported that 14 out of 16 studies found beneficial effects for GOs compared to studying texts alone, and the effectiveness of GOs has been confirmed in a variety of settings.

The findings in L1 reading research on GO instructional approaches have important implications for L2 reading research. In the L2 context, discourse structure awareness more generally has been found to consistently facilitate the comprehension and recall of text information (Carrell, 1984; Ghaith & Harkouss, 2003; Goh, 1990; Hague & Olejnik, 1990; Lahuerta Martinez, 2002), and training in discourse structure knowledge is generally effective (Carrell, 1985; Davis, Lange and Samuels, 1988; Raymond, 1993; Tang, 1992). However, very few discourse structure training studies in L2 reading instruction have explored the use of GOs as instructional tools.

In contrast to the abundance of GO research on L1 readers, there is generally a paucity of GO research in L2 reading comprehension. To my knowledge, Tang (1992) was the only published empirical investigation on the effects of GOs on English text comprehension among ESL learners. In this study, Tang attempted to understand how a teacher-provided tree graph presenting structural knowledge of classification can help facilitate text comprehension and immediate recall in seventh-grade ESL students. She divided 45 participants into two groups. The graphic group was presented with the content of a passage in the form of a tree graph before being asked to complete that partially-completed graph and write a recall. The non-graphic group read the same passage, but as a control group, focused on studying key vocabulary, answering questions related to the material, and writing a recall. The findings showed that the graphic group made gains in the amount of information recalled in the posttest compared to the pretest, whereas the non-graphic group scored the same on both tests. In addition, most of the students in the graphic group reported that using the graphic organizer helped them comprehend the text.

Given the positive results found for the use of GOs in various L1 reading contexts, it is somewhat surprising that only one empirical L2 study on GO use has been published. At the same time, it is widely recognized that L2 students in more advanced EAP (English for Academic Purposes) contexts face significant challenges in understanding difficult texts. In particular, intermediate and advanced L2 learners deserve special attention because more and more international students are pursuing advanced education in the United States and other English speaking countries. One of the major challenges L2 learners face is trying to understand how complex L2 academic reading materials are organized, especially when the rhetorical conventions of L2 texts can differ from those of L1 texts (Grabe & Stoller, 2011). Given the strong potential of GO training for L2 academic reading development, there is a pressing need for further research on this issue.

In addition to the general lack of research on GO instruction in L2 settings, there are relatively few studies employing extended training in GO instruction in either L1 or L2 contexts. Discourse structure awareness is a knowledge base that grows incrementally from increasing exposure and explicit instruction (Koda, 2005). To better understand the potential effects of GOs on discourse comprehension, it is important that researchers extend the instructional training period. Most previous GO studies, however, involved a training period ranging from a few hours to a few weeks, during which only two to eight passages were covered. This short-term instructional treatment may explain why the effect size of GOs was found to be as low as .22 (Cohen's *d*) in a meta-analysis of 23 studies (Moore & Readence, 1984).

A different, but related, issue in GO research is the lack of evidence that GOs are designed to consistently represent a range of discourse structures. In studies designed to promote discourse

structure awareness, researchers used a great number of terms to refer to GOs, including "frame" (Armbruster et al., 1991; Armbruster et al., 1987), "tree diagram" (Guri-Rosenblit, 1989), "map" (Berkowitz, 1986), "flowchart" (Geva, 1983), and "matrix" (Robinson & Schraw, 1994). Each term represents a different way of organizing information, entailing the different levels of details involved in graphic representation. Variations in the graphic representation of discourse structures have caused unnecessary difficulties in comparing and synthesizing research findings. To avoid the "anything goes" situation noted by Griffin and Tulbert (1995, p. 86), a more unified means for GO design is needed.

Given these gaps in our knowledge of GO training on reading development, the primary goal of this study is to examine the effect of extended DSGO instruction on L2 learners' discourse comprehension and reading ability. Specifically, the study addresses the following two research questions:

1. Does a 16-week DSGO instruction program significantly improve college-level EFL students' discourse comprehension and reading ability? If there are immediate instructional effects, do these effects remain after a 7-week delay?
2. Do students' education levels (first vs. third semester) play a significant role in the effectiveness of DSGO instruction on discourse comprehension and reading ability?

Methods

Participants

The participants were 174 first semester and 166 third semester undergraduate students of non-English majors at a major university in China. Forty-nine percent of the participants were males and 51 percent were females. On average, they were 19.2 years of age and had received 9.1 years of uninterrupted English education. They were students from 14 different academic departments at the University.

The participants were from 12 intact English classes—six classes at each of the two education levels (first and third semester). Within each education level, three classes were randomly assigned to receive the DSGO instructional treatment, while the other three classes served as a comparison group and received traditional instruction without research intervention. There were 94 first semester and 85 third semester students in the DSGO experimental group, and 80 first semester and 81 third semester students in the comparison group. The 12 classes were taught by five instructors. Class assignment and instructor information can be found in Appendix B.

Materials and Instruments

DSGO construction. The textbooks used in the College English Program were from a series titled *Twenty-First Century College English* (Zhai et al., 1999). Book 1 was used for the first semester students, and Book 3 was used for the third semester students. Each book consisted of 10 units, each of which had two reading passages (Texts A and B) for classroom use. Text A was intended for intensive reading, which required a large allocation of class time for new vocabulary,

sentence structures, content, and classroom drills. Text B was intended for in-class extensive reading exercise aided by a glossary and comprehension tasks. The length of the instructional texts ranged from 700 to 1,000 words. In addition to the textbook, a companion vocabulary workbook was also used in class for extra practice on words and phrases from Texts A and B of each unit. The DSGOs were developed for Texts A and B in each unit based on the design principles proposed by Jiang and Grabe (2007). These DSGOs were then edited into partially completed GO worksheets to be used as part of the classroom activities (see Appendix C for an example).

Test of general reading ability. The reading comprehension section of three unpublished TOEFL forms was used to assess participants' general reading ability. These tests were used by Educational Testing Service (ETS) on July 11, August 2, and October 25, 1997. The TOEFL reading test was used to examine whether the DSGO instructional treatment had any effect on the general academic reading ability. The general reading ability measure is considered a more distant transfer effect of the DSGO instruction compared to the performance on the DSGO completion task. Students in each class were assigned to the three test forms by means of a pseudo randomization procedure. This randomization procedure ensured that each student was assigned to a different test form in the pretest, posttest, and delayed posttests, respectively, and that an approximately equal number of students took each form during each test.

Each test form had a total of 50 multiple-choice items based on five reading passages. The answers were scored as either correct or incorrect. The total achievable score was 50. The internal consistency reliability coefficients (Cronbach's alpha) for Forms A, B, and C based on students' performance in the pretest were found to be 0.83, 0.78, and 0.80, respectively.

Test of discourse comprehension. A DSGO completion test in three parallel forms (Forms A, B, and C) was developed for this study to examine whether the effect of the DSGO instruction could be transferred to other similar instructional texts. Care was taken to ensure that the reading passages in different forms were similar in length, text organization, and readability (see Table 1). The construction of the partially completed DSGOs followed the same procedures used in preparing the partially completed DSGO worksheets for classroom activities. As in the TOEFL reading test, students in each class were assigned to the three test forms during pretest, posttest, and delayed posttests by means of a pseudo randomization procedure.

Table 1. *Passages for the discourse comprehension test*

	Passage in Form A	Passage in Form B	Passage in Form C
Title	What's wrong with our weather	Climate: A powerful force	Close to home
Number of words	758	756	754
Flesch-Kincaid Grade Levels	10.4	9.8	10.4
Text structures	Cause-effect Problem-solution Description Compare-contrast	Cause-effect Problem-solution Process-sequence Compare-contrast	Cause-effect Problem-solution Description Compare-contrast

The partially completed DSGOs for each reading passage included 20 blanks for the students to fill in after they finished reading the passage (see Appendix C for an example). Detailed scoring rubrics were developed for each form prior to scoring. To ensure inter-rater reliability, about 25% of the pretest papers in each form were scored by two raters. The inter-rater reliability coefficients (Cronbach's alpha) were above 0.99 for all three forms. As a result, one rater scored all the remaining test papers. The reliability coefficients (Cronbach's alpha) of the DSGO completion test based on students' performance in the pretest was found to be 0.88, 0.84, and 0.82 for Forms A, B, and C, respectively.

Procedures

Pre-instruction teacher and student training. The week before the semester started, the researcher met with the other two instructors of the DSGO experimental group and provided them with two 90-minute training sessions. One session focused on knowledge of discourse structure while the other focused on training in DSGO tasks. The goal of the first session was to help the instructors obtain a clear understanding of common discourse structures and also learn how to identify relationships among ideas in texts. The second session provided further practice in discerning important discourse structures in parts of texts, filling in blanks in partially completed DSGOs, and constructing DSGOs for simple texts.

For students to be prepared for future discussions on discourse structures and DSGOs, the experimental classes were given two 45-minute pre-instruction training sessions the week after the pretest. The training materials were a simplified and shortened version of the materials used in the two instructor training sessions. To provide a fair and accurate assessment of the performance in the posttest and delayed posttest for all students, the training sessions were also given to the students in the comparison classes early in the semester. The purpose of offering training to the students in the comparison group was to ensure that they had gained enough knowledge about the DSGO tasks before being required to perform these tasks in the posttest and delayed posttest. In other words, the training was aimed to strengthen the face validity of the DSGO completion test and to help minimize a possible novelty/familiarity effect on task performance.

Instruction. All classes met twice a week and each class meeting consisted of two 45-minute

periods with a 10-minute in-between break. The program curriculum required that instructors finish each textbook unit in eight 45-minute periods or four class meetings in two weeks. The current study involved 16 weeks of instruction covering eight textbook units. Generally speaking, instruction on Text A and its practice exercises took up to six 45-minute class periods. Text B usually took one 45-minute class period, and the exercises in the companion vocabulary workbook took another one period of class time.

The DSGO instruction was built into the regular curriculum for the DSGO experimental classes. The DSGO lesson for each text took one 45-minute class period. During a typical DSGO lesson, the instructor started off by giving a brief introduction to the overall structure of the text. Then the students were asked to fill in partially completed DSGOs designed for the text, either individually, in pairs, or in groups of three. Finally, the instructor went over the answers and conducted some post-DSGO activities such as asking students to summarize the main ideas of the text and answer comprehension questions. The comparison group did not have access to the DSGO instructional materials. The comparison classes were taught in the way that their instructors have always been teaching. Classroom observations confirmed that the instructors of the comparison group neither lectured on discourse structures nor used DSGOs to aid reading comprehension.

Except for the DSGO practice, the experimental and comparison groups shared most of the traditional classroom activities in everyday teaching, including: (a) pre-text vocabulary learning, (b) intensive lexico-grammatical analysis of the text, (c) practice in vocabulary, word building, and sentence structures, and (d) exercises in cloze, translation, and structured writing. The DSGO experimental group spent 75 % of the class time on traditional classroom activities and 25% on DSGO activities, while the comparison group spent 100 % of the class time on traditional classroom activities. A reasonable concern was that the time spent on DSGO instruction by the experimental group might negatively affect their learning of other language skills taught in the traditional curriculum. To address this concern, students' grades on four program-wide unit achievement tests over the 16-week period were collected and analyzed. The findings are reported in the Results section.

Classroom observation. All classes, except those taught by the researcher herself, were observed during the instruction of a complete textbook unit (four class meetings over two consecutive weeks). The class observation showed that no topics related to the discourse structure knowledge were discussed in the comparison classes. The lessons in these classes focused on traditional classroom practices, such as vocabulary learning, lexico-grammatical analysis of the text, and other classroom activities.

The class observations in the DSGO experimental group were focused on the fidelity of the instructional treatment. Attention was directed to the degree to which the teachers properly implemented the key elements of the DSGO lessons, including the amount of time spent on the DSGO tasks, the appropriateness of discourse structure analysis by the instructor, and the level of student participation in classroom activities. The observations showed that instructors in the experimental group followed the DSGO lesson plans faithfully, and the classes proceeded as planned using the DSGO instructional approach.

Test administration. To assess the effect of the DSGO instruction, each student received a total of three tests: a pretest, a posttest, and a delayed posttest. The pretest was administered during regular class time in the first week of the semester. All participants had 50 minutes to complete the TOEFL reading test. The test papers were collected before the DSGO completion test was administered. For the DSGO completion test, students had approximately 10 minutes to read the passage and 20 minutes to work on the DSGO completion task. The posttest was administered in the last week following the conclusion of the class instruction. The delayed posttest was administered seven weeks after the instruction. The same test procedures used in the pretest were followed in the posttest and delayed posttest.

Data Analysis

As mentioned in the previous section, three parallel forms of the DSGO completion test were designed and used in the study. A close examination of the pretest scores suggested that the forms varied significantly in difficulty. Although the test forms were randomized with approximately equal numbers of students taking each form at each test time, it is preferable that the test forms be equated. Since the sample size for each form was relatively small and the difference in difficulty between the forms was not constant along the score scale, the method of linear equating (Kolen & Brennan, 1995) was adopted to equate the forms. These equated scores were used in subsequent data analysis, in replacement of the original DSGO completion scores.

Results

Effects of Graphic Organizer Instruction

Descriptive statistics (mean, standard deviation and sample size) on the DSGO completion test and the TOEFL reading comprehension test are presented in Table 2. To establish a baseline for subsequent statistical analyses, a Bonferroni *t* was employed to compare possible difference in the DSGO completion pretest and TOEFL reading comprehension pretest between the DSGO experimental group and the comparison group at $\alpha = .05/2$ level. At the time of pretesting, the two groups were significantly different from each other on the DSGO completion test, $t = -2.52$, $p = .012$ (two-tailed). On the other hand, the two groups were not significantly different from each other on the TOEFL reading comprehension test, $t = -0.62$, $p = .54$ (two-tailed).

After the baseline comparisons, a multivariate analysis of covariance (MANCOVA) was performed on the present data with instructional groups (DSGO experimental vs. comparison) and education levels (first semester vs. third semester) as the between-subject variables and time of the test (posttest and delayed posttest) and test types (DSGO completion and TOEFL reading) as the within-subject variables. This is a doubly multivariate repeated measures design in which each participant was measured twice (posttest and delayed posttest) on two variables (DSGO completion and TOEFL reading). Since the experimental group and comparison group were significantly different in the DSGO completion pretest, these pretest scores were used as a covariate for the DSGO completion posttest and delayed posttest. To satisfy the requirements of a MANCOVA, the TOEFL reading comprehension pretest scores were used as a covariate for the TOEFL reading comprehension posttest and delayed posttest.

Table 2. *Descriptive statistics for GO completion test and TOEFL reading comprehension*

Instructional groups	Education levels	*n*	Pretest		Posttest		Delayed posttest	
			M	*SD*	*M*	*SD*	*M*	*SD*
GO Completion Test								
Experimental	First-semester	94	10.56	5.17	14.73	3.66	15.66	4.23
	Third-semester	85	12.44	4.95	15.69	2.74	16.69	2.97
	Total	179	11.56	5.09	15.18	3.29	16.15	3.71
Comparison	First-semester	80	12.59	4.60	12.36	5.03	14.19	4.36
	Third-semester	81	13.26	4.77	13.39	3.61	14.41	4.70
	Total	161	12.93	4.68	12.88	4.38	14.30	4.52
TOEFL Reading Comprehension								
Experimental	First-semester	94	17.38	6.38	24.20	6.57	22.60	6.56
	Third-semester	85	19.82	6.51	29.65	6.49	26.87	6.10
	Total	179	18.53	6.54	26.78	7.06	24.61	6.68
Comparison	First-semester	80	18.31	6.81	21.14	6.09	21.79	7.03
	Third-semester	81	19.60	7.05	27.83	7.03	28.41	6.87
	Total	161	18.97	6.94	24.55	7.37	25.16	7.68

The MANCOVA showed a significant three-way interaction among test types (DSGO completion and TOEFL reading), time of the test (posttest and delayed posttest) and instructional groups (DSGO experimental vs. comparison) at the 0.5 level, $F(1, 331) = 6.25$, $p = .013$. In addition, there was a significant three-way interaction among test types, instructional groups, and education levels (first semester vs. third semester) at the .05 level, $F(1,331) = 4.57$, $p = .033$. There was no significant four-way interaction among test types, time of the test, instructional groups, and education levels, $F(1, 331) = 1.18$, $p = .278$.

To further examine the test types × time of the test × instructional groups interactive effect, a repeated measures ANCOVA was performed on the DSGO completion test scores, and a repeated measures ANOVA was conducted on the TOEFL reading scores, respectively. In both procedures, the within-subject variable was time of the test and the between-subject variable was instructional groups. In the ANCOVA procedure, the DSGO pretest scores were used as a covariate.

For the DSGO completion test, the repeated measures ANCOVA showed that the main effect of instructional groups was significant, $F(1, 339) = 56.93$, $p < .001$, partial $\eta^2 = .144$. There was neither a significant within-subject effect (time of the test) nor a significant interactive effect (time of the test × instructional groups).

For the TOEFL reading comprehension test, the repeated measures ANOVA showed that there was a significant two-way interaction between time of the test and instructional groups, $F(1, 327) = 10.84$, $p = .001$. To further examine this significant two-way interactive effect, two ANOVA procedures were performed on the TOEFL reading scores in the posttest and delayed posttest, respectively. During the posttest, the DSGO experimental group significantly outperformed the comparison group, $F(1, 364) = 12.74$, $p < .001$, partial $\eta^2 = .034$. However, during the delayed

posttest, no significant difference was found between the two instructional groups, F (1, 344) = 0.45, p > .05.

To examine the significant three-way interaction among test types, instructional groups, and education levels, repeated-measures ANCOVA was performed on the DSGO completion test scores, and a repeated measure ANOVA was conducted on the TOEFL test scores, respectively. In both analyses, the within-subject variable was time of the test and the between-subject variables were instructional groups and education levels. In the repeated measures ANCOVA, the DSGO pretest scores were used as a covariate. Both analyses were focused on the interactive effects between the instructional groups and education levels, because a non-significant interaction between these two variables would indicate that the DSGO instructional benefits were found irrespective of the education levels. Indeed, the repeated-measures ANCOVA on the DSGO completion test scores showed that there was no significant interaction between instructional groups and education levels, F (1, 339) = 0.01, p > .05. Likewise, the repeated measures ANOVA on the TOEFL scores showed that there was no significant interaction between instructional groups and education levels, F (1, 335) = 2.12, p > .05.

These findings were sufficient to answer the two research questions of this study. The first question of this study is whether the 16-week DSGO instruction program significantly improves college-level EFL students' discourse comprehension. The results showed that the DSGO instructional program had a significant effect that was reflected in both the DSGO completion test and TOEFL reading test administered immediately after 16 weeks of instructional treatment. However, after a seven-week delay, the DSGO instructional effect was reflected in the DSGO completion test alone.

The second question is whether students' education levels (first vs. third semester) played a significant role in the effectiveness of DSGO instruction on discourse comprehension and reading ability. Since there was no significant interactive effect between instructional groups and education levels on either the DSGO completion test scores or the TOEFL reading test scores, the results were interpreted as indicating that education levels did not significantly influence the effectiveness of the DSGO instruction. In other words, the DSGO instructional effects apply equally well to both the first and third semester students.

Acquisition of Other Language Skills

One legitimate concern about the research study was whether or not the DSGO experimental group would be at a disadvantage in relation to the comparison group in learning other language skills, because this group was allocated less class time to learn the skills emphasized in the traditional classroom. To address this concern, the DSGO experimental group and the comparison group were compared for their performances on four unit achievement tests taken over the 16-week period. These unit achievement tests, developed by the College English Program, were given to all students in the program. Although the format of these tests were not standardized, all four tests consisted of 75% objective items, including a multiple-choice cloze test and other multiple-choice items in listening, grammar and structure, and reading comprehension (and 25% sentence translation or writing). The scores on the objective items of the unit achievement tests are used in the following analyses.

Table 3 shows the descriptive statistics of the scores in Tests 1 through 4 for the DSGO experimental group and comparison group, respectively. A total of 317 participants were included in this part of the analysis. Other participants were excluded due to missing cases.

Table 3. *Descriptive statistics for scores on the unit achievement tests*

	Instructional groups	*n*	*M*	*SD*	*Min.*	*Max.*
Test 1	Experimental	171	56.24	8.73	29.50	75.00
	Comparison	146	59.03	8.47	24.50	73.50
Test 2	Experimental	171	56.05	7.76	30.00	70.50
	Comparison	146	57.40	7.69	36.50	73.50
Test 3	Experimental	171	52.46	9.83	27.00	72.50
	Comparison	146	52.95	10.12	32.00	74.50
Test 4	Experimental	171	55.73	6.98	33.50	69.00
	Comparison	146	53.20	7.25	28.00	65.50

Repeated-measures ANOVA was conducted to examine the effect of instructional groups on traditional language skills acquisition as measured by the objective items in the unit achievement tests. The within-subject variable was the unit achievement tests 1-4, and the between-subject variable was the instructional groups. The results showed a significant interaction between unit achievement tests and instructional groups, $F (3, 313) = 9.78$, $p < .001$. A post hoc analysis with Bonferroni t compared the DSGO experimental group and the comparison group on each of the unit tests at $\alpha = .05/4 = .0125$ level. In Test 1, the comparison group did significantly better than the DSGO experimental group, $t (315) = -2.88$, $p = .004$ (two-tailed). However, there was no significant difference between the two groups on Test 2, $t (315) = -1.55$, $p = .121$ (two-tailed) or Test 3, $t (315) = -0.44$, $p = .661$ (two-tailed). Results from Test 4 showed a reversed trend; the DSGO experimental group performed significantly better than the comparison group, $t (315) = 3.16$, $p = .002$ (two-tailed). Although the comparison group performed significantly better than the DSGO experimental group on Test 1, they were nonetheless outperformed by the DSGO group on Test 4. Contrary to the belief that less class time devoted to traditional classroom activities would lead to reduced ability of the DSGO students in learning the traditional language skills, the results of this study showed that the DSGO instruction may even be beneficial to the acquisition of these skills.

Instructor Effect

Another concern about the current findings was that instructors might be a confounding factor on class performance. It is not clear whether the classes that received the DSGO instruction were assigned better instructors than the comparison classes (see Appendix B for instructor assignment to classes). To examine this issue, a MANOVA was performed on the DSGO completion and TOEFL reading posttests to examine four orthogonal contrasts: Class 1 (Instructor 1) vs. Classes 2 & 3 (Instructor 2); Class 4 (Instructor 2) vs. Classes 5 & 6 (Instructor 3); Class 7 (Instructor 1) vs. Classes 8 & 9 (Instructor 4); Class 10 (Instructor 4) vs. Classes 11 & 12 (Instructor 5). The multivariate F test showed that the effects of contrasting instructors were not significant, $F (8, 706) = 0.51$, $p = .849$. The results were interpreted as indicating that there

was no significant instructor effect on class performance.

Discussion

Effects of DSGO Instruction

This study examined the effects of DSGO instruction on reading comprehension among non-English major college EFL students. Analysis of the measure of DSGO completion showed a significant instructional effect. On the DSGO completion tasks during both the posttest and the delayed posttest, the instructional effect was found to be .144 on the partial η^2, or equivalently .82 on Cohen's *d*. This effect size would be considered large (Cohen, 1988). As mentioned earlier, Moore and Readence (1984) concluded, based on a meta-analysis of 23 studies, that the effect of GO instruction is generally low. Most previous GO studies had a training period ranging from a few hours to few weeks. In contrast, the present study showed that the instructional effect as measured by DSGO completion test can be improved by prolonging the training period.

This study confirmed the positive effects of GO training on reading comprehension found in most prior L1 research (Alvermann & Boothby, 1986; Armbruster et al., 1991; Armbruster et al., 1987; Balajthy & Weisberg, 1990; Geva, 1983; Guri-Rosenblit, 1989; Spiegel & Barufaldi, 1994; Robinson & Schraw, 1994; Williams et al., 2005). Moreover, this study has also reaffirmed the findings of Tang (1992) that ESL learners who received training in GOs improved their reading comprehension significantly.

In addition, the effect of DSGO instruction remained persistent on the measure of DSGO completion even after a seven-week delay. Carrell (1985) found that the effect of text structure training remained persistent in the delayed posttest three weeks after instruction. The present study provides evidence that the effect can be retained as long as seven weeks.

The relatively strong outcome of the DSGO treatment, as reflected in the DSGO completion test, should not be simply construed as resulting from a familiarity factor, because the DSGO group was tested on what they had been practicing. By offering the comparison group two 45-minute training on discourse structure knowledge and DSGOs, we are more confident in attributing the superior performance in the DSGO group to their improved ability in discourse comprehension. Our belief has been further strengthened by two important findings of this study. First, the DSGO group significantly outperformed the traditional instruction group in the DSGO completion posttest even after a seven-week delay. Since the transient familiarity effect should have been somewhat neutralized after this long period of delay, it is reasonable to conclude that the superior performance of the DSGO group was more likely due to a more permanent effect of improved text structure knowledge. Second, stronger evidence for the DSGO treatment effect was provided by the superior performance of the DSGO treatment group over the traditional instruction group in the immediate TEOFL reading posttest. Since a familiarity effect is not possible under this test condition, it is reasonable to conclude that the superior performance of the DSGO group was due to transfer of the DSGO treatment effect to general reading ability. Although some level of the familiarity factor cannot be entirely ruled out, the ability to recognize

discourse structures and fill in partially completed DSGOs should be perceived as a good skill for reading comprehension itself (Grabe, 2009; Trabasso & Bouchard, 2002). If a student does better at a skill widely recognized as capable of improving reading abilities in other settings, we should perceive their improving performance as indicating the development of a skill critical for improving reading comprehension, rather than just better performance on an isolated task.

In addition to using the DSGO completion test scores as a measure of DSGO instructional effect, the current study examined the more distant transfer effect of the DSGO instruction on the learners' general reading ability as measured by the TOEFL reading comprehension. Analysis of the TOEFL reading comprehension scores showed that the DSGO experimental group did significantly better than the comparison group immediately after instruction, but the difference was not retained after a seven-week delay. The partial η^2 for the instructional effect was found to be .034 during the posttest, which is equivalent to .38 on Cohen's d. This effect size would be considered small-to-medium (Cohen, 1988).

One possible reason for this level of effect size is that the reading passages used to measure general reading ability lack strong discourse structures. The traditional paper-based TOEFL reading comprehension test included five reading passages, each of which was 300-350 words in length. Generally speaking, these relatively short passages were not able to contain salient discourse structures and, as such, were insensitive to the beneficial effects of discourse structure knowledge (Chapelle, Enright, & Jamieson, 2008). This problem has been rectified in the new TOEFL iBT reading test that contains longer and more complex texts (650-700 words) with multiple discourse structures. In the future, it would be interesting to investigate whether the effect size of the DSGO instruction would increase if the new TOEFL iBT reading comprehension test were used to assess the general reading ability.

Another possible reason for this level of effect size is the fact that the general reading ability test reflects a distant transfer effect of the DSGO instruction. Previous research on discourse structure awareness and DSGO instruction often employed free recall or tasks closely related to those practiced as measures of reading comprehension. It is reasonable to assume that the treatment effect under these circumstances was more readily transferred and captured. Although the effect size was in the small-to-medium range, this result has provided evidence that the DSGO instruction is able, even if only in a relatively small way, to positively influence general reading ability, which is the ultimate goal of reading instruction.

The positive effect of DSGO instruction on general reading ability disappeared after seven weeks. It is possible that the 16-week instruction period was not long enough to make this process highly internalized and automatic. Consequently, students reverted back to their old approach to discourse processing after a seven-week delay. Further research is needed to investigate the optimal length and intensity of DSGO training required to make this process highly internalized and automatic. In addition, the new TOEFL iBT reading comprehension test should be used in replacement of the paper-based TOEFL reading comprehension test which seems to lack sensitivity to the beneficial effect of discourse structure knowledge.

Finally, the study found that education levels of the EFL students played no important role in determining the effects of DSGO instruction on either the DSGO completion test or the TOEFL

reading test. In other words, the positive effects of the DSGO instruction occurred irrespective of education levels. However, since this study employed a fixed factor design for education levels, the results should not be generalized to other education levels. The participants in this study represent a homogenous group of learners with relatively similar levels of English proficiency. Ideally, this study should have selected education levels with greater differences between each other (e.g., first- vs. seventh-semester) so that the interactions between GO instruction and education levels could be examined in greater clarity. However, since English is only offered to non-English undergraduate majors for up to two years at most Chinese universities, this option was not possible in this study. To determine whether the DSGO instruction is effective for learners of all proficiency levels, further research will be needed to include participants with a wider range of language proficiency.

Limitations of the Study

This study has two major limitations. First, this study did not randomly assign the instructors to different classes due to administrative reasons. One common way to circumvent the instructor effect is to assign the same teacher to both control and treatment conditions. However, this practice does not guarantee the internal validity of the findings, because the instructor could be biased toward the treatment class. In reality, it is often not practical to exert a perfect control over the instructor effect in a large-scale experiment involving several hundred participants. The fact that no significant instructor effect was found clearly indicated that the present experimental design had worked well for this study.

The other limitation of the study was that the delayed posttest took place seven weeks after the instructional treatment, which seemed to be unnecessarily conservative. An alternative approach would have the delayed posttest with a three or four week delay, which would have made the present data more comparable to earlier studies (e.g., Carrell, 1985). In fact, Nation and Webb (2011) noted that delayed posttests are most meaningful if implemented in a 3-4 week range, and they are much less likely to be meaningful beyond a four-week delay. Future training studies are needed to pursue this issue.

Despite the limitations, this study was able to implement a large-scale DSGO program covering 16 instructional passages over a course of 16 weeks. DSGO treatment studies with comparable length and intensity are rare in the literature. This study has provided evidence that an extended training period is able to increase the effect size of the DSGO instruction.

Pedagogical Implications

The current study has a number of implications for materials development and classroom teaching. First, due to the number of commonly used discourse structures, it is possible to consistently represent these structures in instructional texts with a set of easily designed DSGOs. This study successfully integrated DSGO instruction into an existing university curriculum by constructing DSGOs to reflect the discourse structures of each passage in the textbook. It has shown that principles and basic designs of DSGOs (Jiang & Grabe, 2007) can be applied to a wide range of instructional texts. The DSGO instruction can be embedded within a standard reading curriculum.

Second, the principles and basic DSGO designs can be taught to teachers through a brief training period. This study provided two 90-minute pre-instruction training sessions, during which instructors were taught how to recognize discourse structures and how to integrate the DSGOs in reading instruction. These instructors were able to absorb the discourse structure knowledge and provide effective instruction to their students in a relatively short period of time if given appropriate DSGO diagrams to work with.

Third, active involvement by the students is critical to learning the DSGOs effectively. To achieve active involvement, classroom activities must be task-oriented and engaging. In this study, students were asked to complete the partially completed DSGO tasks themselves. By working on the partially completed DSGO tasks, the students were compelled to engage at a deeper level of information processing.

Finally, the traditional instructional approach, which emphasizes vocabulary and sentence level comprehension, can be demotivating. Although the students in the comparison group were better equipped with basic language skills as shown in the first unit achievement test, they were outperformed by the DSGO experimental group by the time of the fourth test. Adding DSGO tasks to the curriculum was not only able to improve students' knowledge of discourse structures, but also helped to facilitate the acquisition of other linguistic skills. Thus, the incorporation of the DSGO tasks into the reading curriculum should be viewed as an effective use of class time.

Acknowledgments

References

Alvermann, D., & Boothby, P. (1986). Children's transfer of graphic organizer instruction. *Reading Psychology*, *7*, 87–100.
Armbruster, B., Anderson, T., & Meyer, J. (1991). Improving content area reading using instructional graphics. *Reading Research Quarterly*, *26*, 393–416.
Armbruster, B., Anderson, T., & Ostertag, J. (1987). Does text structure/summarization instruction facilitate learning from expository text? *Reading Research Quarterly*, *22*, 331–346.
Balajthy, E., & Weisberg, R. (1990). Transfer effects of prior knowledge and use of graphic organizers on college developmental readers' summarization and comprehension of expository text. In J. Zutell & S. McCormick (Eds.), *Literacy theory and research: Analysis from multiple paradigms: Thirty-ninth yearbook of the National Reading*

Conference (pp. 339–345). Chicago, IL: National Reading Conference.

Berkowitz, S. (1986). Effects of instruction in text organization on sixth-grade students' memory for expository reading. *Reading Research Quarterly*, *21*, 161–178.

Carrell, P. (1984). The effects of a rhetorical organization on ESL readers. *TESOL Quarterly*, *18*, 441–469.

Carrell, P. (1985). Facilitating ESL reading by teaching text structure. *TESOL Quarterly*, *19*, 727–752.

Chapelle, C., Enright, M., & Jamieson, J. (Eds.) (2008). *Building a validity argument for TOEFL.* New York, NY: Routledge/Taylor and Francis Group.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Davis, J. N., Lange, D. L., & Samuels, S. J. (1988). Effects of text structure instruction on foreign language readers' recall of a scientific journal article. *Journal of Reading Behavior*, *20*, 203–214.

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

Ghaith, G. M., & Harkouss, S. A. (2003). Role of text structure awareness in the recall of expository discourse. *Foreign Language Annals*, *36*, 86–96.

Goh, S. T. (1990). The effects of rhetorical organization in expository prose on ESL readers in Singapore. *RELC Journal*, *21*, 1–13.

Grabe, W. (2003). Using discourse representations for reading development. In M. Swanson & K. Hill (Eds.), *Proceeding of the 2002 JALT Conference* (pp. 9–17). Tokyo, Japan: JALT Publications.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York, NY: Cambridge University Press.

Grabe, W., & Gardner, D. (1995). Discourse analysis, coherence, and reading instruction. *Lenguas Modernas*, *22*, 69–88.

Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading* (2nd ed.). New York, NY: Pearson Education.

Griffin, C., & Tulbert, B. (1995). The effect of graphic organizers on students' comprehension and recall of expository texts: A review of the research and implications for practice. *Reading and Writing Quarterly*, *11*, 73–89.

Guri-Rosenblit, 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.

Hague, S. A., & Olejnik, S. (1990). What L2 readers remember: Is it related to their awareness of text structure? In J. Zutell & S. McCormick (Eds.), *Literacy theory and research: Analyses from multiple paradigms: Thirty-ninth yearbook of the National Reading Conference* (pp. 409–417). Chicago, IL: National Reading Conference.

Hoey, M. (2001). *Textual interaction: An introduction to written discourse analysis*. New York, NY: Routledge.

Hoffmann, K. F. (2010). *The impact of graphic organizer and meta-cognitive monitoring instruction on expository science text comprehension in fifth grade students*. (Unpublished doctoral dissertation). North Carolina State University, Raleigh.

Jiang, X., & Grabe, W. (2007). Graphic organizers in reading instruction: Research findings and issues. *Reading in a Foreign Language*, *19*, 34–55.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. New York,

NY: Cambridge University Press.

Kolen, M., & Brennan, R. (1995). *Test equating: Methods and practices*. New York: Springer.

Lahuerta Martinez, A. C. (2002). Empirical examination of EFL readers' use of rhetorical information. *English for Specific Purposes*, *21*, 81–98.

Meyer, B. (1985). Prose analysis: Purpose, procedures and problems. In B. Britton & J. Black (Eds.), *Understanding expository text* (pp. 269–285). Hillsdale, NJ: Erlbaum.

Mohan, B. A. (1986). *Language and content*. Reading, MA: Addison-Wesley.

Moore, D. W., & Readence, J. E. (1984). A quantitative and qualitative review of graphic organizer research. *Journal of Educational Research*, *78*, 11–17.

Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle.

Pearson, P. D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (pp. 815–860). New York, NY: Longman.

Raymond, P. M. (1993). The effects of structure strategy training on the recall of expository prose for university students reading French as a second language. *The Modern Language Journal*, *77*, 445–458.

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

Robinson, D. H., & Kiewra, K. A. (1995). Visual argument: Graphic organizers are superior to outlines in improving learning from text. *Journal of Educational Psychology*, *87*, 455–467.

Robinson, D. H. (1997). Graphic organizers as aids to text learning. *Reading Research and Instruction*, *37*, 85–105.

Simmons, D. C., Griffin, C. C., & Kameenui, E. J. (1988). Effects of teacher-constructed pre- and post-graphic organizer instruction on sixth-grade science students' comprehension and recall. *Journal of Educational Research*, *82*, 15–21.

Singer, M. (1990). *Psychology of language: An introduction to sentence and discourse processes*. Hillsdale, NJ: Lawrence Erlbaum.

Spiegel, G. F., & Barufaldi, J. P. (1994). The effects of a combination of text structure awareness and graphic postorganizers on recall and retention of science knowledge. *Journal of Research in Science Teaching*, *31*, 913–932.

Tang, G. (1992). The effect of graphic representation of knowledge structures on ESL reading comprehension. *Studies in Second Language Acquisition*, *14*, 177–195.

Trabasso, T., & Bouchard, E. (2002). Teaching readers how to comprehend text strategically. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 176–200). New York, NY: The Guilford Press.

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

Williams, J. P., Hall, K. M., Lauer, K. D., Stafford, K. B., DeSisto, L. A., & deCani, J. S. (2005). Expository text comprehension in the primary grade classroom. *Journal of Educational Psychology*, *97*, 538–550.

Williams, J. P., Nubla-Kung, A. M., Pollini, S., Stafford, K. B., Carcia, A., & Snyder, A. E. (2007). Teaching cause-effect text structure through social studies content to at-risk second graders. *Journal of Learning Disabilities*, *40*, 111–120.

Zhai, X., Zheng, S., & Zhang, Z. (Eds.), (1999). *Twenty-first century college English: Reading*

and writing, *Book 1* and *Book 3*. Shanghai, China: Fudan University Press; High Education Press.

Appendix A
Studies in the effect of Discourse Structure Graphic Organizers

Studies	Participants	Text	Treatment/Comparison	Outcome	Conclusion
Williams et al. (2005)	128 Second-graders, L1	Nine compare-contrast passages	1. Text structure program (TS) focused on the structure of the text, using such techniques as GOs (matrix). 2. Content-only program (CO) focused on general information and interesting facts, rather than the structure of the text. 3. No instruction (NI)	TS group outperformed CO and NI groups on recall of clue words, locating clue words, matrix sentence generation (oral and written), immediate recall (oral), near transfer (oral) far transfer (oral), & content measures.	1. TS group demonstrated superior comprehension of the compare-contrast structure without any loss of content compared with CO and NI groups; 2. TS group demonstrated transfer to an uninstructed compare-contrast structure.
Williams et al. (2007)	243 Second-graders, L1	Ten cause-effect passages	Same as above	TS group outperformed CO and NI groups on locating clue words, underlining clauses, effect questions in explicit teaching (one cause—one effect), and in near and far transfer (one cause—multiple effects). There was no difference between TS and CO groups on content measures.	TS had greater comprehension of cause-effect text without a loss in the amount of content acquired. TS program also showed transfer effect on some comprehension measures.
Robinson & Kiewra (1995)	Experiment A: 110 college students enrolled in an undergraduate psychology course, L1	A text of 6,500 words from a introductory psychology book	1. Text only (T) 2. Text plus outline (O) 3. Text plus GO (GO)	GO group outperformed O and T groups in coordinate relations, and outperformed outline group in contrasting premises.	Given enough time, students in the GO group were more successful in learning hierarchical relations and coordinate relations, and applying newly learned knowledge to integrated writing.
	Experiment B: 42 college students enrolled in an undergraduate psychology course, L1	Same as above	Same as above except that subjects studied the material for a hour and reviewed the material for 15 minutes after a day delay	GO group outperformed O or T groups on hierarchical relations, coordinate relations, and application of newly learned knowledge.	

Robinson & Schraw (1994)	Experiment A: 48 college students, L1	A text describing the social groupings of six fish, three-minute text display time	1. Matrix 2. Outline 3. Text Only	The main effect of display was significant, indicating that the matrix group outperformed the Text Only group on comparison and pattern items.	Matrix is computationally more efficient than outline or text. The advantage of the matrix group disappeared after a 25-minute testing delay, ironically, due to the efficiency of the matrix display.
	Experiment B: A different group of 45 college students, L1	Same as in Experiment A except the text display time was reduced to a minute	Same as Experiment A	As in Experiment A, there was a significant effect of display even under reduced display time condition.	
	A different group of 45 college students, L1	Same as in Experiments A & B except that the display time was increased to five minutes a 25-minute testing delay was added		The main effect of display created a non-significant result, indicating the matrix group did not perform better than the other two groups.	
Hoffmann (2010)	162 fifth-graders, L1	Eight compare-contrast passages	1. GO plus metacognitive monitoring (MM) 2. GO 3. MM 4. Comparison Group (CG)	Students in the GO plus MM and MM groups showed improved comprehension scores, whereas the GO group showed no improvement in comprehension scores.	Comprehension scores increased only in conditions where students received training in metacognitive monitoring.

Appendix B

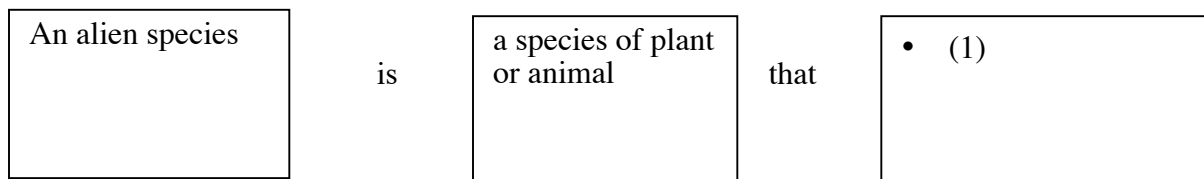*Class Assignment and Instructor Information*

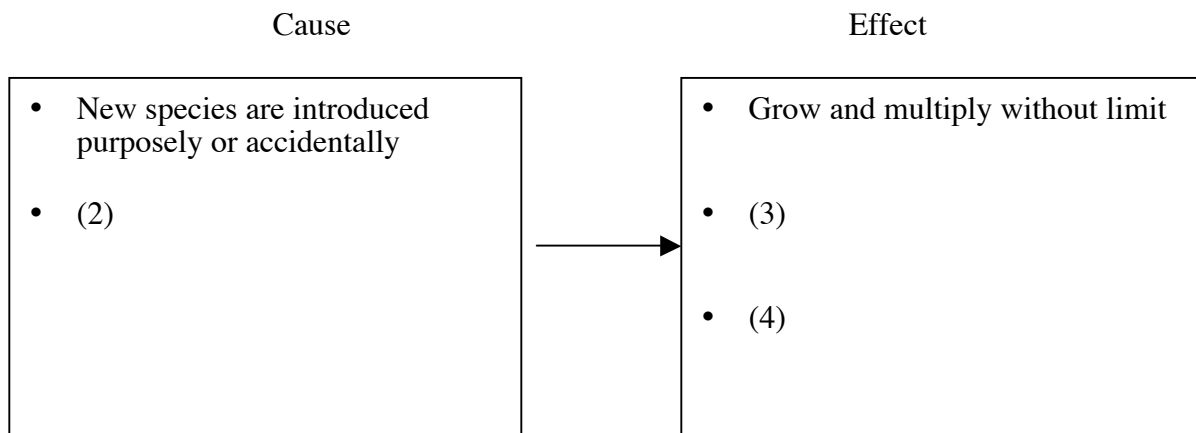| Level	Class	Condition	Instructor	Gender	Years of Teaching	Highest degree
First-semester	1	Experimental	Instructor 1 (Researcher)	Female	7	MA
	2	Experimental	Instructor 2	Female	11	BA
	3	Experimental	Instructor 2			
	4	Comparison	Instructor 2			
	5	Comparison	Instructor 3	Male	35	BA
	6	Comparison	Instructor 3			
Third-semester	7	Experimental	Instructor 1 (Researcher)	Female	7	MA
	8	Experimental	Instructor 4	Female	27	BA
	9	Experimental	Instructor 4			
	10	Comparison	Instructor 4			
	11	Comparison	Instructor 5	Male	20	MA
	12	Comparison	Instructor 5			

Appendix C

GO Completion Worksheet: An Example
Directions: Fill in the numbered blanks in the graphic organizers based on the text your read.

1. What is an alien species? (Paragraph 1)

An alien species	is	a species of plant or animal	that	• (1)

2. The introduction of alien species and its consequences (Paragraph 1)

Cause   Effect

Cause	Effect
• New species are introduced purposely or accidentally • (2)	• Grow and multiply without limit • (3) • (4)

About the Author

Xiangying Jiang is an Assistant Professor of TESOL at West Virginia University. Her research focuses on second language reading development, including comprehension strategies, discourse structure, oral reading fluency, vocabulary, and the influence of first language literacy on second language literacy development. Email: Xiangying.Jiang@mail.wvu.edu