



## Scaffolding EFL Reading with Text-Structure Visual Displays: Idea Maps and Idea Matrices

以文本架構的視覺呈現輔助英文閱讀:比較圖示法與圖表法之成效

Hsi-chin Chu<sup>1</sup>

Received: 4 January 2020 / Revised: 4 January 2020 / Accepted: 21 February 2020 /

Published online: 18 May 2020

© National Taiwan Normal University 2020

### Abstract

Awareness of text structure has been perceived as crucial for effective text processing and comprehension, especially for EFL readers in academic contexts. This study investigated the instructional effect of two types of text-structure visual display—idea maps and idea matrices—in two phases of implementation. In the first phase, freshman non-English majors in two experimental groups went through respective training on the completion of idea maps and idea matrices in support of textbook reading, while the control group did not have such access. Pretest and posttest measures indicated a significant effect of idea map tasks on a general reading test. For text retention in recall, no effect was found for both groups. Nevertheless, descriptive statistics revealed that both experimental groups retained more higher level ideas and less lower level ideas. For inferences generated in recall, idea matrix completion produced significantly fewer incorrect inferences but more correct inferences, as reflected in the descriptive statistics. In the second phase, a swapping of treatment was done between the two experimental groups. Both groups reported a higher preference for the idea map task with reasons regarding text processing, task demand, and design features. Thus, the higher level of relational processing demanded by the two-dimensional matrix may dilute the benefit of idea localization.

---

✉ Hsi-chin Chu  
hcchu@ntnu.edu.tw

<sup>1</sup> National Taiwan Normal University, 162 He-ping E. Rd., Sec. 1, Taipei 10610, Taiwan

## 摘要

對以英文為外語的學習者而言,閱讀學術性的文章時能覺知文本架構有助於閱讀理解。此項能力在學術閱讀的重要性自不待言。本研究透過兩個階段的教學,以兩種文本架構的視學呈現-示(idea map)與圖表(idea matrix)-輔助閱讀以了解並比較其成效。在第一個階段,兩組大一學生分成兩個實驗組分別作圖示填空及圖表填空兩種練習來支援閱讀,第三組為對照組。前測與後測則施以閱讀能力測驗及讀後回憶書寫。研究結果顯示圖表練習增進閱讀能力,但是兩種練習對文本記憶並無影響,雖然敘述統計顯示兩組實驗組經過練習後高階的文意記得較多,低階的文意記得較少。同時圖表練習也顯著地減少了錯誤的推論,增加了正確的推論。在第二個階段,經過兩組實驗組互換訓練方式後,問卷顯示兩組均較喜歡圖示練習,圖表練習較不受青睞主要原因為呈現兩個面向的文意需要較費力的連結,因此抵消了將重點聚焦的好處。

**Keywords** Text structure · Idea map · Idea matrix · Recall · Inference

關鍵詞 文意架構 · 圖表 · 圖示 · 記憶書寫 · 推論

## Introduction

Awareness of text structure has been perceived as crucial for effective text processing and comprehension, especially for EFL readers in academic contexts. Freshman non-English majors who do not have extensive exposure to texts characterizing English writing conventions might not be able to perceive the target text in its due complexity, which might lead to a somewhat ill-formed mental representation of text and thus less text retention. Synthesis of second language reading research also points to the importance of text structure and discourse organization in instruction [9]. The recognition of textual logic via the construction of text structure may form the basis for reader response, be it intra-personal response writing or interpersonal post-reading discussion. As reiterated by Kintsch [16], the construction of a textbase can precede the integration of text and knowledge, mediating the move from textbase representation to knowledge transformation.

To facilitate textbase construction, graphic display of idea structure in maps or matrices is often provided by the textbook writers and classroom teachers to scaffold readers' anchoring of key idea chunks. The map characterizes a core box or nodes that encloses the main idea which links to layers of branch ideas in hierarchical relations, spreading out in all directions [23]. The matrix displays ideas in rows and columns in a table that contains the topics and their categories, showing a cross-topic relations [14]. Since different designs for visual display of idea structure may command different processing loads and thus comprehension outcomes [6, 28], it is therefore important to understand the effects of training in using the two types of visual display. Previous studies investigating the instructional effects of text-structure visual display in the EFL context are dearth; a comparison between idea maps and idea matrices is scant. Given the popularity of supplying idea maps in EFL textbooks and the feasibility in generating idea matrices with word processors by teachers, it is also of practical interest to understand the effectiveness of these two types of text-structure visual display.

This study thus compared the effects of idea maps and idea matrices on students' reading proficiency, text retention, inference generation, and task preference. Two experimental groups and one control group were involved in this study in two phases:

the first phase investigating the treatment effects on three measures of reading comprehension; the second, after exchanging treatments, on the preferred approach and the reasons for preference.

## Literature Review

### Idea Maps and Idea Matrices: Design, Processing, and Mental Representations

The literature on text-structure visual display has yet to establish a unanimous use of terminology. The design features of the two visual displays investigated in this study, idea maps and idea matrices, generally correspond to those ascribed by previous researchers to types of maps and matrices. The idea map comprises a central node with main ideas denoted and links to subordinate nodes in radiating directions, which resembles the concept map or knowledge map in Nesbit and Adesope [23]. The difference is that, in this study, rather than having concepts, commonly comprising one or two keywords as written units, nodes encloses ideas that are extracted and may be condensed from the text, hence the idea map. The idea map thus may be more suitable for supporting lengthy text read by university learners and is similar to Eppler's [7] mind map which addresses semantic, rather than concept, relation.

A matrix is a type of graphic organizer mainly used in supporting learning from expository text [18]. It is a table of two dimensions, organizing ideas vertically by topic and horizontally by category, allowing the comparison of topics based on categories [14, 22]. It is similar to Crooks and Cheon's [4] skeletal template for graphic organizers, consisting of a matrix with columns for main topics and rows for repeatable subtopics. Unlike the idea map, which can cover all the ideas in one map for a lengthy text, two or three tables/matrices may be required to display varied types of textual logic (e.g., problems–causes–solutions) within a text.

To demonstrate the features that distinguish idea matrices and idea maps, a sample of each on the textbook passage, "No More Dead Space," is presented in the appendices. The idea matrices (Appendix 2) in two tables exemplifies the two thinking patterns respectively: contrast and listing of factors. The major topics are displayed in the first row (*past and now*) and the categories showing a secondary-level logic (*window display and way people walk*) are shown in the left column; in this way, the concepts are displayed in two dimensions. By contrast, idea maps focus on the visual display of hierarchical concept relations that are linked by lines, without indicators of secondary logical relation, as demonstrated in Appendix 1. The words extracted from the text for the nodes/slots are largely the same.

The visual supports, idea maps and idea matrices, help reduce the extraneous cognitive load because they share two principles of learning for multimedia design [21], *coherence principle* and *spatial contiguity principle*. The *coherence* principle has it that learning may be enhanced when materials competing for resources in working memory and consequently posing extraneous cognitive load, in this case, details, are excluded (p. 89). The two types of visual display may direct readers' attention to the processing of germane information, eliminating the processing of unimportant textual elements.

Another principle, *spatial contiguity principle* [20], underlines the function of more relating in text processing for idea matrices than for idea maps. *Spatial contiguity principle* postulates that ideas/words displayed in tandem may be integrated with ease for their proximity enables their simultaneous processing in the working memory. Both the idea maps and the idea matrices localize, i.e., place side by side, the top-level ideas in the design. The topical localization is shown in the idea map in the first circle of radiating nodes; in the idea matrix, in the top row in the table. Beyond the topical localization, the idea matrix demands additional cross-topic categorical localization, which could be repeated in several rounds.

Beyond topical relating, readers of different tasks may engage in different routes of text processing: the idea map users may chain ideas nearby within topics in a superordinate-to-subordinate hierarchy while the idea matrix users exercise cross-topic relational processing [6] for categories that are spatially segregated in the text, yielding the sum of the information in a square [18]—a more compact, smaller two-dimensional space displaying a higher degree of spatial contiguity than the diffusional idea map.

Not only readers' text processing but also the resultant mental representation could vary between the two tasks. As elucidated by Schnotz and Bannert [28] and Danielson and Sinatra [6], different graphics may show different relation patterns and yield varied mental models because the design prescribes how we organize the ideas into our long-term memory [29]. Similar to Robinson and Schraw [26] in their comparison of outline and matrix (p. 404), the idea map, like outline, addresses intra-topic relations while the idea matrix provides inter-topic relations via cross-topic categorization. The relation for the former may require more automatic local-bridging inferences, while that for the latter the more effortful global-coherence inferences [17]. Thus, constructing the idea map may be less cognitive-demanding than the idea matrix, resulting in a mental model less robust than that mediated by the idea matrix, a trade-off between processing and product.

Nonetheless, in gauging the effectiveness of the graphic text adjunct for EFL learners, one element has to be taken into account—language proficiency. Given the limited working memory capacity, Cognitive Load Theory [31], on one hand, speaks to us that localizing important ideas, as in the repeated rounds exercised for the idea matrix task, may reduce the extraneous text-processing load and enhance learning. On the other, the design of the idea matrix may entail an intrinsic cognitive load due to a connection of highly interactive ideas ([29], p. 681). On top of these calculations, language processing, which is not yet automatic for L2 readers, may tap processing capacity. Therefore, it is unknown whether this extra processing load would suppress EFL readers' capacity in searching for cross-topic relations for the idea matrix tasks and thus affect the essential processing of basic information, or whether rows of localized ideas displayed in the idea matrix would ease the processing, allowing readers more capacity for the generative processing for inferential understanding. Hence, a comparison of the training effects of idea matrix use with those of idea map completion would help elucidate the issues.

### Previous Studies on the Effects of Text-Structure Visual Displays

A plethora of research on the effect of studying or constructing maps/graphic organizers have been conducted, allowing meta-analysis studies, notably L1, to synthesize

the research. One such study by Nesbit and Adesope [23] on concept maps and knowledge maps attested significant but varied effects on knowledge retention. Schroeder, Nesbit, Anguiano and Adesope's [30] meta-analysis of concept maps also showed positive overall effects on learning. In the EFL context, Kansizoğlu [13] meta-analyzed types of visual display and revealed an overall effect of visual displays on listening, reading, writing, grammar, and vocabulary/concept learning. The measures he analyzed were language skills and language components, which differ from those commonly inspected in L1: text or knowledge retention. The skills examined in this L2 meta-analysis thus included reading skill. Yet, there was no mention of how reading was measured.

Unlike studies on the effect of map tasks, scant research was done on the effects of matrix tasks alone, though several L1 studies compared the effects of matrix with other display types, mostly text alone and outlining. These comparisons found the superiority of matrix task in speeding up search for relationship [27], in recall of subtopic information [15], and in the identification of global and local relations as well as in the favorable learner perceptions [14].

Kiewra et al. [15], after obtaining the findings that matrices promoted recall of subtopics but not overall recall, contemplated that recall may not genuinely capture the advantage of relational processing of outlines and matrices. Yet, Robinson and Kiewra [25] employed multiple measures, including recall and essay writing, to compare the effects of providing multiple outlines and matrices for a chapter-length text reading. The matrix task was found to promote the hierarchical and coordinate relations in the production measures of recall and essay writing; however, the facts not presented in the matrices were not retained as much as in the control condition. It was conjectured that the attention to the facts may be diverted to the relational processing driven by the matrix, a cost-benefit effect. The lengthy texts used in this study may echo the need of EFL readers to deal with the English coursebooks, which are laden with multiple thinking patterns requiring multiple matrices to display, instead of a cohered idea map.

In the EFL context, Tseng [33] compared two versions of concept maps, comprehensive vs. thematic representation, used as text adjunct to the history texts read by 10th graders in Taiwan. The former encompasses all textual ideas in temporal and spatial relationships, the latter only the gist in causation relationship. The finding indicated that the visual display representing the text content may impose on readers' mental representation, as revealed in the better retention yielded via the comprehensive maps and the enhanced reasoning mediated by the thematic maps. Another EFL study investigating the effect of concept maps [32] found that concept maps scaffolded comprehension assessed by comprehension test and by propositional recall primed by the fill-in-the-blank task.

Long-term intervention studies on text adjunct of visual display's in EFL instructional milieu are sporadic. For one, Jiang's [11] EFL university students, in a 16-week program, filled a partially completed DSGO (discourse structure graphic organizers), before discussion. Immediate effects were found on general reading ability measured by TEOFL test and both immediate and delayed effects on DSGO completion tasks. Specifically, the task of DSGO partial completion helped to control the task challenge and allowed for learner agency. However, using the target treatment tool, DSGO, to measure the comprehension outcome in pre- and posttests may fail to address the transfer effect to the independent reading without supports. Another long-term

treatment using matrices as an instructional tool was conducted by Maxim [19] in a one-semester elementary-level university German course. Throughout a semester, students read a novel scaffolded by idea matrices by episodes. Although insignificant effects on recall were drawn, the sole use of matrices in the instruction placed no disadvantages on the experimental group in terms of language proficiency and recall of texts.

Thus far, research has yet to investigate the use of idea matrices in scaffolding EFL reading in the instructional context, not to mention to compare the training effects of idea maps and idea matrices. Several implications are drawn from the review. First, the apprehension of text structure instigated by one-shot provision of graphs or maps might be temporary and transient. To internalize the structuring process and transfer it to other text readings, a long-term, repeated activation may be necessary. Second, since the text structure in the texts read by university students is never so prototypical that a simple graph can grasp, flexible and multiple visual displays tailored to the contents of each reading text should be designed for ability transfer. Third, the commonly reader-generated tasks such as recall should be adopted to measure reading outcomes in a fine-tuned manner, such that ideas in recall can be teased apart into levels of textual importance. Fourth, in addition to measuring the retention of essential text elements, the generative dimension of comprehension as shown in the inferences generated in recall can be measured as well so as to capture the impact of relational processing. Therefore, this study aims to compare the effects of two reading supports, idea maps and idea matrices, on EFL university students' reading comprehension in terms of general reading, retention, and inference. Students' preferences of display type and reasons for the preference are investigated as well. Four research questions are posed.

## Research Questions

1. Do tasks of idea map completion and idea matrix completion affect EFL students' general reading ability? Is there a difference between the two?
2. Do tasks of idea map completion and idea matrix completion affect EFL students' text retention? Is there a difference between the two?
3. Do tasks of idea map completion and idea matrix completion affect EFL students' inference generation? Is there a difference between the two?
4. Which type of visual display of text structure did students prefer to use? and why?

## Methods

### Participants and Design

This study employs a three-group pre- and posttest quasi-experimental design. Three sections of freshman non-English majors of *intermediate-high* level for a course, English I and II, participated in this study in two semesters. Two sections were assigned to two experimental conditions, one working on idea maps, hence map group (30 students), another on idea matrices (27 students), hence matrix group. A third section was designated as the control group (34 students). GEPT-Reading pretest scores showed no difference in reading proficiency among the three groups at the outset of the study ( $F(2, 88) = 0.33, p > 0.05, \eta^2 = 0.01$ ).

## Materials and Treatment

Two stages of treatment were involved. The first stage extended from the first semester to the first quarter of the second semester, during which five lesson units of the textbook *Q: skills for success 4*, 2nd edition [5], were covered. The second stage then followed until the end of the second semester, during which the two experimental groups swapped the treatment approach for the remaining three lesson units. All three groups followed the same schedule covering the same lesson units, except that the control group did not have access to the visual display of text structure.

Each lesson unit includes two reading passages of around 1000 words on the same subject in social sciences, with one adopted for close reading and another for general understanding. The passages are argumentative, generally in a combination of problem-solution, cause-effect, and comparison-contrast structures. An idea map and several idea matrices were designed for each reading passage, with partial node/slots filled (see Appendices 1 and 2 for samples). Unlike self-generated tasks, partially-completed maps/matrices, designed by teachers, can serve as a cognitive scaffold [4] which reduces the processing demand, saves the processing time, and avoids disinteresting the learners [8].

To facilitate the processing from linear text reading to the construction of hierarchical text structure, two types of coaching materials were developed for both experimental groups: the material-oriented signaled texts and the learner-oriented prompts [24]. For the signaled texts, five devices were used to mark the texts: (1) an text-extraneous *bracket* addressing thinking pattern (e.g., [cause-effect-solution]) placed at the beginning of each segment; (2) *slashes* segmenting the global and local text chunks; (3) *enclosure* of logical cues such as factors, thanks to, past, or today; (4) main points *boldfaced*; and (5) sub-points *underlined*. For the learner-oriented prompt, a set of questions were provided to direct students' relational processing of global textual chunks. These materials were designed in accordance with Mayer's [21] *signaling* principle for noticing, selecting, extracting, and organizing, and *segmentation* principle for text management.

A typical cycle of a unit lesson (in two two-period sessions) included (1) self-directed map/matrix work on assigned passage at home, (2) in-class instructed close-reading of the first passage, (3) checking/discussing on the answers to the idea map/matrix task, (4) extension activities, and (5) pair/group-work of map/matrix task on the second passage.

## Instrumentations and Data Collection

To assess the effect of treatment in the first stage, pre- and posttest of GEPT-reading and written recall were administered. A split-block design was employed using two sets of GEPT tests and two recall passages, Dolphin (592 words at 9.4 Flesch-Kincaid grade-level) and Animal Test (573 words at 9.8 grade-level), with a comparable argumentative idea structure. Students read the passage for 12 min, and then with the passage removed, wrote either in Chinese or English or a combination of the two in 20 min, as recalls in either or both L1 and L2 were equally valid. [12]. Then, students proceeded with GEPT-reading test - intermediate high, in 50 min. To understand students' preference and reasons for preference of the two types of tasks, at the

completion of the second phase treatment, students in the experimental groups responded to two probing questions.

## Data Analyses

**Analysis of GEPT Score** ANCOVA analysis, with group (3 levels) as an independent variable, pretest score as a covariate, and posttest score as a dependent variable, was performed. In the case of substantial effect size for group, i.e., practical significance, pairwise contrasts were examined.

**Coding, Calculation, and Analyses of Recall** For recall, the weighted pausal system as suggested by Bernhardt [3] was used and a set of procedure was followed to establish scoring templates, scoring procedure, and inter-coder reliabilities. To create scoring templates, two native speakers marked the pauses when reading aloud to build idea unit systems, yielding an agreement rate of 79% for “Dolphin” and 85% for “Animal Testing.” The discrepancy was then resolved by the researcher and two Chinese-speaking coders, by taking the syntactic features of written Chinese into consideration. A total of 112 units for “Dolphin” and 104 units for “Animal Testing” were produced. Next, the two trained coders weighted the idea units by 4 levels according to the textual importance, such that a quarter of the total units, 28 and 26 respectively, were assigned to each level for the two articles. The inter-coder reliabilities were 0.81 for “Dolphin” and 0.83 for “Animal Testing.” Again the differences in weighting were resolved via discussion among the two coders and the researcher.

The two coders then checked students’ written recall respectively against the weighted idea unit templates, resulting in an inter-coder reliability of 0.78 for level 1 units, 0.84 for level 2 units, 0.88 for level 3 units, 0.81 for level 4 units, and 0.90 for unit total. Then, each recalled unit was weighted before summed up for the total recall score. The sums were then divided by the total full scores (280 for Dolphin and 260 for Animal Test) yielding the *percentage of weighted recalled units*, the score for recall total. For the score of each level unit, the raw number of recalled units was divided by its total (28 and 26 respectively) as the *percentage of recalled units*.

**Coding of Inferences** The residual in recall that did not have a match with the idea unit systems—ideas extraneous of text contents—was coded as inference into two types: correct and incorrect. The former are those in line with the textual ideas; the latter, those generated due to misinterpretation, misreading, or insufficient linguistic knowledge. The unit of inference was defined as utterance that expresses a complete thought [34], or a self-contained argument. The inter-coder reliability for correct inference was 0.69, and for incorrect inference, 0.83. The inference scores were yielded by averaging.

The same statistical procedure for GEPT scores was performed on the five measures of text retention, total and levels 1 to 4, and three measures of inferencing, total, correct, and incorrect. Altogether eight rounds of ANCOVAs were performed, followed by an inspection on the paired contrasts if substantial effect size, i.e., practical significance, was yielded.



**Analyses of Questionnaire Data** Chi-square was performed on the preference for idea maps/idea matrices. Content analysis with constant comparisons was done on the reasons given to identify key concepts and form categories.

## Results

### Results on Reading Proficiency

Preliminary check on GEPT-reading score indicated that the assumptions of ANCOVA analyses, reliability and homogeneity of regression slopes, were met. As shown in Table 1, after adjusting for the variance from pretest mean scores ( $\eta p^2 = 0.44$ ), no main effect of group ( $F(2, 87) = 2.33, p > 0.05$ ) was found. Despite insignificant, the effect size ( $\eta p^2 = 0.05$ ) was small to medium. A check on the range of the 95% confidence intervals among the three post hoc mean contrasts showed that the mean difference ( $M_{\text{diff}} = 6.68, SE = 3.12$ ; Table 2) between the map group ( $M = 87.93^a, SE = 2.27$ ) and the control group ( $M = 81.26^a, SE = 2.13$ ) did not contain zero (95% CI = [0.48, 12.87],  $p < 0.05, d = 0.54$ ; Table 2), hence a significant difference. The task of idea map did enhance students' performance in GEPT-reading test. Therefore, the response to RQ1 is that the task of idea maps increased students' general reading ability while that of idea matrix did not. And there was no difference between the two training tasks in enhancing general reading comprehension.

### Results on Text Retention

Preliminary checks on reliability and homogeneity of effect slopes ensured no violation of ANCOVA assumptions for the five rounds of analysis: the recall of total and recall of four levels of idea units. For posttest total recall (Table 1), with the pretest total scores ( $\eta p^2 = 0.14$ ) adjusted for, no effect of group was yielded ( $F(2, 87) = 0.24, p = 0.78$ ) and the effect size was small ( $\eta p^2 = 0.01$ ). Post hoc pairwise comparisons were therefore not performed.

**Table 1** ANCOVA results for group on three sets of measures

	Pretest $\eta p^2$	$F(2, 87)$	$p$	$\eta p^2$
GEPT-reading	0.44	2.33	0.10	0.05
Retention-level 1	0.25	2.36	0.10	0.05
Retention-level 2	0.17	2.39	0.10	0.05
Retention-level 3	0.07	0.33	0.77	0.01
Retention-level 4	0.00	1.02	0.36	0.02
Retention-total	0.14	0.24	0.78	0.01
Correct inference	0.04	1.40	0.25	0.03
Incorrect inference	0.03	2.91	0.06	0.06
Total inference	0.01	0.69	0.50	0.02

**Table 2** Pairwise comparisons on the mean differences for three sets of measures

	Map-Control	Matrix-Control	Map-Matrix
GEPT-reading	6.68* (3.12) [0.48, 12.87] $d=0.54$	3.91 (3.20) [-2.44, 10.27] $d=0.32$	2.76 (3.29) [-3.77, 9.30] $d=0.22$
Retention-level 1	-1.14 (1.94) [-5.00, 2.72] $d=0.15$	-4.25* (2.00) [-8.22, -0.29] $d=0.55$	3.12 (2.06) [-0.97, 7.20] $d=0.40$
Retention-level 2	-4.74 (2.75) [-10.20, 0.72] $d=0.44$	-5.43(2.76) [-10.91, 0.06] $d=0.50$	0.69 (2.90) [-5.08, 6.46] $d=0.06$
Retention-level 3	2.11 (3.30) [-4.44, 8.67] $d=0.16$	-0.50 (3.40) [-7.25, 6.25] $d=0.04$	2.61 (3.49) [-4.32, 9.54] $d=0.20$
Retention-level 4	5.23 (3.67) [-2.06, 12.53] $d=0.36$	2.08 (3.75) [-5.37, 9.53] $d=0.14$	3.15 (3.88) [-4.57, 10.87] $d=0.21$
Retention-total	1.16 (2.61) [-4.03, 6.34] $d=0.11$	-0.75 (2.68) [-6.07, 4.57] $d=0.07$	1.91 (2.77) [-3.59, 7.41] $d=0.18$
Inference-correct	-0.62 (0.45) [-1.51, 0.28] $d=0.36$	0.11 (0.43) [-0.75, 0.96] $d=0.06$	-0.73 (0.47) [-1.65, 0.19] $d=0.42$
Inference-incorrect	-0.03 (0.44) [-0.92, 0.85] $d=0.01$	-1.01*(.46) [-1.92, -0.10] $d=0.57$	0.97* (0.48) [0.02, 1.92] $d=0.55$
Inference-total	-0.61 (0.69) [-1.98, 0.76] $d=0.22$	-0.73 (0.68) [-2.07, 0.62] $d=0.27$	0.12 (0.74) [-1.35, 1.58] $d=0.04$

$N=30$  for map group;  $N=27$  for matrix group;  $N=34$  for control group

\*  $p < 0.05$ . Standard error in parentheses

95% CI in brackets

$d$  = Cohen's  $d$

Breaking down the total score into the component scores of level 1 to level 4, the results vary. For the posttest recall of level 1 idea units, with the pretest scores accounted for ( $\eta p^2 = 0.25$ ; Table 1), group did not show a difference ( $F(2, 87) = 2.36$ ,  $p > 0.05$ ), albeit with a small to medium effect size ( $\eta p^2 = 0.05$ ; see Table 1). A check on the 95% confidence intervals among the three post hoc mean contrasts (see Table 2) showed a significant contrast ( $M_{\text{diff}} = -4.25$ ,  $SE = 2.00$ ) between the matrix group ( $M = 8.51^a$ ,  $SE = 1.49$ ; Table 3) and the control group ( $M = 12.76^a$ ,  $SE = 1.33$ ; Table 3) with the range of confidence interval not including zero (95% CI [-8.22, -0.29],  $p < 0.05$ ,  $d = 0.55$ ; Table 2). Therefore, training in idea matrix construction decreased the retention of level 1 ideas while idea map production did not. Meanwhile, there was no difference between the two tasks.

For the posttest recall of level 2 idea units, as revealed in Table 1, with pretest scores ( $\eta p^2 = 0.17$ ) adjusted, no effect of group was found ( $F(2, 87) = 2.39$ ,  $p > 0.05$ ), despite

**Table 3** Adjusted posttest scores for total and for levels of idea units recalled across groups

	Map group ( $N = 30$ )	Matrix group ( $N = 27$ )	Control group ( $N = 34$ )
Level 1	11.62 <sup>a</sup> (1.41)	8.51 <sup>a</sup> (1.49)	12.75 <sup>a</sup> (1.33)
Level 2	15.89 <sup>a</sup> (2.00)	15.20 <sup>a</sup> (2.07)	20.63 <sup>a</sup> (1.85)
Level 3	24.44 <sup>a</sup> (2.40)	21.83 <sup>a</sup> (2.53)	22.33 <sup>a</sup> (2.26)
Level 4	30.66 <sup>a</sup> (2.67)	27.51 <sup>a</sup> (2.80)	25.43 <sup>a</sup> (2.50)
Total	23.58 <sup>a</sup> (1.90)	21.67 <sup>a</sup> (2.00)	22.42 <sup>a</sup> (1.78)

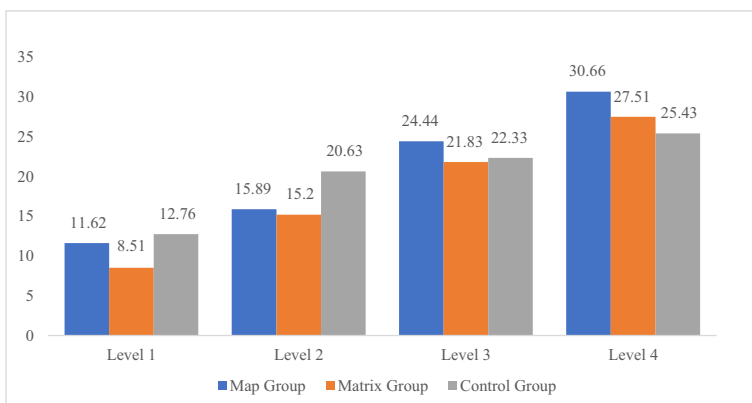
<sup>a</sup> Adjusted for pretest scores. *SE*: standard error

of a small to medium effect size ( $\eta p^2 = 0.05$ ). The range of 95% CI among the three post hoc contrasts was therefore inspected and no significant contrast was found, for all three intervals included zero (see Table 2). As such, training students to produce a visual display, of either type, did not make a difference in the retention of level 2 ideas.

For the posttest recall of level 3 idea units, as indicated in Table 1, with the impact of pretest scores ( $\eta p^2 = 0.07$ ) covaried, group did not make a difference ( $F(2, 87) = 0.33$ ,  $p = 0.72$ ,  $\eta p^2 = 0.01$ ). Pairwise comparison was thus not ensued. Table 1 also shows that, for level 4 idea unit, no relation was found between pretest scores and posttest scores, indicating that with pretest scores adjusted for or not, the effect of group would not be affected. In addition, no significant difference was found among the three groups on posttest scores ( $F(2, 87) = 1.02$ ,  $p = 0.36$ ,  $\eta p^2 = 0.02$ ; Table 1). Hence, no paired group contrasts were examined.

Notwithstanding the insignificance yielded in most of the between-group comparisons in recall measures, descriptive statistics showed trends varying between the experimental groups and the control group. As revealed in Fig. 1, unlike the recall of level 1 and level 2 idea units, for which the control group produced higher adjusted posttest means than the two experimental groups, for level 3 unit, map group generated a higher adjusted means than control group and, for level 4, both experimental groups' adjusted posttest means were higher than control group.

The answer to RQ2 is therefore summarized as follows. For retention, training in idea-matrix completion decreased the retention of level 1 ideas, but made no difference



**Fig. 1** Adjusted posttest means for four levels of units recalled across groups

in the retention of level 2, level 3, and level 4 ideas, whereas training in idea map completion did not make a difference in the retention of ideas at all four levels. In addition, no difference was found between the two instructional tasks on the retention of ideas at any level. Nonetheless, there was a tendency for the experimental groups to generate a smaller percentage of lower level idea units and a greater percentage of higher level idea units, the idea map group especially, than the control group.

## Results on Inference Generated in Written Recall

Preliminary check on homogeneity of effect slopes for the analysis on total, correct, and incorrect inferences indicated that the assumptions of ANCOVA were met. For the posttest production of total inferences, after pretest means ( $\eta p^2 = 0.01$ ) were covaried, no group effect was found ( $F(2, 87) = 0.69, p < 0.05, \eta p^2 = 0.02$ ; Table 1). Post hoc contrasts were thus not examined.

For correct inference, with pretest variance ( $\eta p^2 = 0.04$ ) accounted for, group did not show a difference on the adjusted posttest means ( $F(2, 87) = 1.40, p = 0.25$ ) with a small effect size ( $\eta p^2 = 0.03$ ; Table 1). Again, no follow-up contrasts were checked.

For the posttest generation of incorrect inferences, after adjusting for pretest means ( $\eta p^2 = 0.03$ ), group did not reveal a difference ( $F(2, 87) = 2.91, p > 0.05$ ), yet with a small to medium effect size ( $\eta p^2 = 0.06$ ; Table 1), endorsing a check on the three pairs of group contrast. The range of 95% CI interval shown in Table 2 revealed a mean difference of  $-1.01$  ( $SE = 0.46$ ) between the matrix group ( $M = 2.08^a, SE = 0.34$ ; Table 4) and the control group ( $M = 3.09^a, SE = 0.30$ ; Table 4), with the 95% CI interval not containing zero ( $[-1.92, -0.10], p < .05, d = 0.57$ ) (Table 2). Similarly, the contrast ( $M_{\text{diff}} = -0.97, SE = 0.48$ ) between the matrix group and the map group ( $M = 3.06^a, SE = 0.33$ ; Table 4) also indicated a 95% CI range ( $[-5.73, -0.41], p < 0.05, d = 0.55$ ) not including zero as well (Table 2). Therefore, the matrix group, after treatment, produced less incorrect inferences than the control group and the map group.

Descriptive statistics, as shown in Fig. 2, revealed a reverse pattern among the three groups between the production of correct inferences and incorrect inferences. While the matrix group produced a significantly lower adjusted mean frequency of incorrect inferences ( $M = 2.08^a, SE = 0.34$ ) than the other two groups, they produced the highest adjusted mean frequency of correct inferences ( $M = 3.58^a, SE = 0.32$ ) among the three groups, albeit insignificant. By contrast, the map group produced the least correct inferences but much more incorrect inferences than the matrix group.

Hence, in response to RQ3, the idea matrix tasks yielded significantly less incorrect inferences. Idea matrix tasks also generated significantly fewer incorrect inferences than idea map tasks. For correct inference, no effect was found with the two types of

**Table 4** Adjusted posttest means for group on types of inference in written recall across group

Type of inference	Map group ( $N = 30$ )	Matrix group ( $N = 27$ )	Control group ( $N = 34$ )
Correct	2.85 <sup>a</sup> (0.32)	3.58 <sup>a</sup> (0.32)	3.47 <sup>a</sup> (0.30)
Incorrect	3.06 <sup>a</sup> (0.33)	2.08 <sup>a</sup> (0.34)	3.09 <sup>a</sup> (0.30)
Total	5.88 <sup>a</sup> (0.50)	5.77 <sup>a</sup> (0.51)	6.49 <sup>a</sup> (0.45)

<sup>a</sup> Adjusted for pretest scores. *SD* standard deviation; *SE* standard error

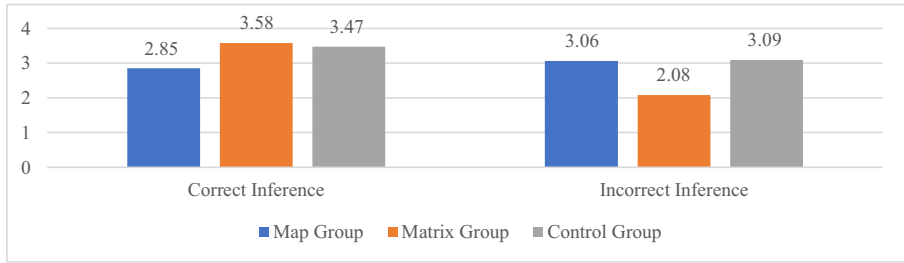


Fig. 2 Posttest adjusted means for two types of inference across groups

task, respectively, nor a comparative effect between the two. Nevertheless, descriptive statistics showed a reversed effect between the two tasks on correct and incorrect inferences.

### Results on Students’ Preference

To answer RQ4, a chi-square test for independence, with a 2 by 2 design, was performed for the first questionnaire item tapping students’ preference for the two tasks, and the result indicated no significant association between group and preference for task, ( $\chi^2(1, N = 57) = 0.03^a, p = 0.86$ ) with a small effect size (Cramer’s  $V = 0.06$ ). A majority of students (62.1% of map group and 67.9% of the matrix group) opted for idea maps. Despite more exposure,

Table 5 Reasons for the preference of types of text-structure visual display

	Idea map	Idea matrix
Reader process	<ul style="list-style-type: none"> <li>*Easier to locate words in texts.</li> <li>*Seeing ideas from big to small.</li> <li>*Clarify relations in levels.</li> <li>*Depict the construction and reflect memory in the brain in a way more like the outcome of thinking.</li> </ul>	<ul style="list-style-type: none"> <li>*Able to sort ideas and see logic.</li> <li>*See the structure at a deeper level.</li> </ul>
Design feature	<ul style="list-style-type: none"> <li>*More visualizing; less textual.</li> <li>*Lines and arrows directing the thread of thought thus providing more cues to fill in.</li> <li>*Grasp [present] the gist/summarize the text in one sighting.</li> </ul>	<ul style="list-style-type: none"> <li>*Ideas displayed in a system</li> <li>* Ideas displayed in a sequence, from top to down in a table.</li> <li>*Ideas displayed following a line of thought.</li> <li>*More information contained.</li> <li>*Showing arguments.</li> <li>*Ideas in table easier to relate in texts.</li> <li>*Tables are easier to produce.</li> </ul>
Task demand	<ul style="list-style-type: none"> <li>*More freedom to construct, with ideas in bubbles going in all directions.</li> <li>*Less time consumed, with fewer words (just keywords) to complete.</li> </ul>	
Affective response	<ul style="list-style-type: none"> <li>*More fun.</li> <li>*More user-friendly.</li> </ul>	
Perceived utility	<ul style="list-style-type: none"> <li>*Good for getting keywords.</li> <li>* Good for general understanding.</li> </ul>	<ul style="list-style-type: none"> <li>*Good to use for long chapters.</li> <li>*Good for summary writing.</li> </ul>

the matrix group did not prefer the task of idea matrix, with their preference rate being even lower than the map group. For questionnaire item 2, Table 5 summarizes the result of analysis on students' reasons for their preference of the two tasks, in five categories: reader process, design feature, task demand, affective response, and perceived utility.

In giving reasons for preference, contrastive reasons for a dislike of the other approach were also given but only on the idea matrix task, with four on the effortful text processing: (1) *the need to condense words into phrases*, (2) *the need to reduce sentences into gist*, (3) *the need to search for words far away to connect*, and thus (4) *the requirement of thinking to connect*; one on the idea matrix design: *tables being constraining on thinking*; another one on task demand: *requiring more effort and time to organize ideas*, and therefore negative affective response: *boring*.

## Discussion

The finding that the map group yielded a superior performance in the GEPT-reading test reflected the advantage of idea map completion in promoting EFL university students' general reading proficiency, a recognition phase of reading. The effect may resonate with the positive finding confirmed in the L2 meta-analysis by Kansizoğlu [13] on the effects of graphic organizers on language and reading skills, support Tajeddin and Tabatabaei [32] on the effect of concept map on general reading comprehension, and corroborate Jiang [11] on the intervention effects of DSGO competition on EFL students' reading comprehension. The positioning of nodes and the lines to link nodes in the visual display may enhance the map group's essential processing for meaning [20] and thus the construction of a mental model closer to the textbase [16]. The fact that the experimental students across the board favor more of idea maps may reflect the ease in the task operation, as evidenced in students' account of reasons for preference such as *easy cross-reference between text (words) and map*, *line of thoughts from general to specific and thus corresponding to textual development*, *more visual and less textual elements*, *more freedom and less constraining*, *summarizing with one sighting*, and thus the affective response: *more fun and user-friendly* (Table 5).

The map group's excel in general reading comprehension, nevertheless, did not extend to the productive phase of retention—which was well attested in L1 meta-analyses on the effect of maps/graphic organizers (e.g., [23]), a case with the matrix group as well. Yet, the reverse trends in the recall of higher and lower level ideas between the experimental groups and the control group indicated that both text adjuncts boosted the retention of higher level ideas and suppressed that of lower level ideas. It could be that the benefit of the visual displays on retention obtained by L1 learners was subdued by the additional language processing by L2 learners. The L2-trained users of idea maps/matrices might have to allocate their capacity, compressed by language processing, to relating higher level ideas extracted, leaving little capacity for the processing of lower level ideas. Much like Robinson and Kiewra [25] who found visual display posed effect on the recall of subtopics but not on overall recall, the

quantity of higher level ideas enhanced may be canceled out by the suppressed lower level ideas.

For idea matrix tasks, no advantage was found in general reading proficiency; rather, the benefit was on the elimination of incorrect inference generated in memory for texts. The effects of idea matrix text supplement on text building evidenced in L1 studies were not present in this study. L1 studies on matrices showed the effect on hierarchical and coordinate relations in recall and essay writing [25] and on the recall of subtopics [15], while the present studies did not show an effect on general reading nor text recall. The L1–L2 discrepancy in results may be attributed to several factors. First, most of the matrix studies were one-shot factorial investigating the effect of target task rather than intervention studies on transferred task, the effect of which is more difficult to yield because the pretest measure has been accounted for. Secondly, the short passages in the GEPT reading test may not require extensive relational processing that was commended by longer passages with multiple paragraphs, a processing better captured by idea matrix task. Thirdly, the task of idea matrix, with ideas related in two dimensions, could be too demanding for the EFL learners to construct even with the support of text signaling and structural prompt ([4], as of GO notetaking) and with the tables partially filled. Although these supports were implemented during treatment, it is likely that, even with such scaffolding, the task was still at a frustrating level for the present readers, not to mention when no such supports were present at pre- and posttest. This is resonated by a larger proportion of students who did not favor idea matrix work, as revealed in their reasons: it is *a table that is constraining, one needs to rephrase words, integrated sentences, search for and connect words in remote text parts, engage in thinking, and therefore more effort and time is consumed for organizing ideas, hence a boring task.*

Notwithstanding the less preference for idea matrix task, there was a growth in the generative phase of text comprehension evidenced by the significantly less incorrect and the more correct, shown in descriptive statistic, inferences generated, a phenomenon reflecting a more accurate mental model built for deep comprehension [10]. The two-dimensional matrices may induce heavier use of relational reasoning ([1]; Danielson and Sinatra 2016), via text reduction and integration, than the hierarchical idea maps. More global-level relating may summon more inferences to establish global coherence, leading to a situation model [16] more in line with the textual logic, thus less inaccurate inferences generated. Although cognitively more demanding, a relatively small proportion of students were able to perceive the strength of matrix design as *systematic display of ideas, in a top-to-down fashion, thus easy to refer in the text, reflecting a line of thought, containing more information in a smaller space, hence enabling learners to sort ideas and see logical structure at a deeper level.* Moreover, *tables are easy to produce.*

In a nutshell, the varied findings between the two treatment groups on three comprehension measures may be interpreted as an interplay of Mayer's [20] three levels of text processing: the extraneous processing of language, the essential processing for basic understanding, and the generative processing for text-knowledge integration. The idea map that is more in line with the text model should have evoked an essential processing and resulted in a textbase construction. However, in this study, only the recognition phase of comprehension captured such effect. The production phase of recall failed to attain an effect, despite a weighting on higher level idea shown in retention. It is likely the extraneous language processing hampered the due effect. Moreover, the idea matrix requiring multiple

reasoning processing could have instigated a generative processing for inferences. Though significantly fewer inaccurate inferences were produced by the matrix group, a significantly higher production of accurate inferences was not attained. Again, the extraneous language processing may circumvent such effect to surface. The dissociation in effects among the three measures within groups also points to the likelihood that essential processing may function at a different level from generative processing.

## Implications

The present findings may shed light on material production and classroom instruction. EFL textbook writers and reading teachers may adopt text-structure visual displays, idea maps or idea matrices, in accordance with students' language proficiency and learning goals. For freshmen non-majors, who are yet to perceive the benefits of idea matrix work—apprehending textual logics so as to connect large textual chunks in lengthy academic texts—idea map construction may be the initial step toward general understanding. Yet, the layers of facts/ideas in idea maps eventually have to be framed in the secondary thinking logics so as to elucidate the complex thoughts displayed in the arguments of academic texts. Idea matrix tasks could follow after students' language proficiency attains maturity and academic demands surface. As put by one student, “idea maps are good for general understanding; idea matrices, for the summary of lengthy texts.”

Given the importance of scaffolding text-structure construction in EFL context, more research should be conducted not only on the effects of varied design of visual display but also on the effects of coaching materials, text signals and learner prompts, so as to understand their roles in readers' noticing, selecting, extracting, organizing, and integrating of textual ideas during the text-structure construction process. Future research should also implement measurements of comprehension, production phase especially, to assess levels of text understanding. With more measurements employed, dimensions of comprehension outcome could be examined to profile the effects of supporting text-structure visual display on EFL learners. Moreover, in this study, partially completed idea maps and idea matrices were designed for students to complete. It is unknown whether teacher/publisher-contrived or students-generated map/matrix would yield different results. Hence, how text-structure visual display is exercised in the classroom constitutes another issue for future studies.

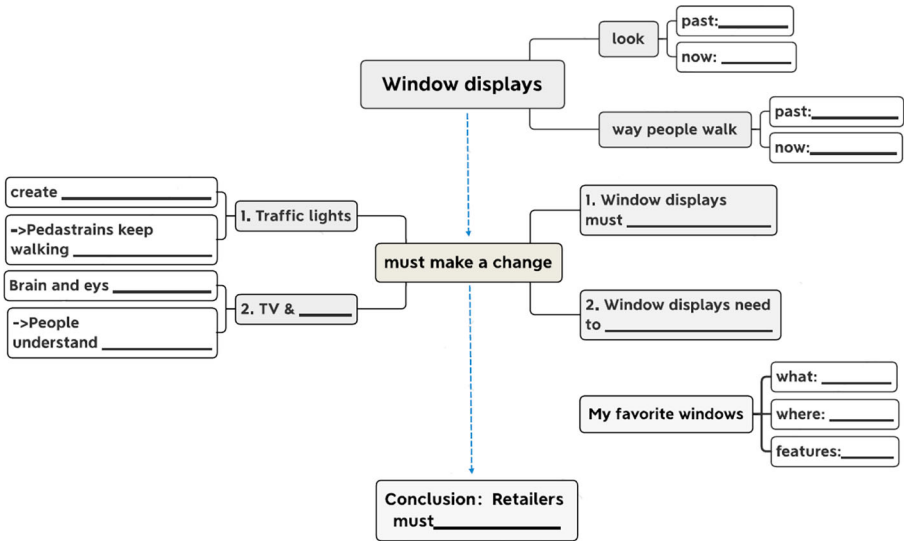
**Funding Information** The research reported here was supported by MOE through grant 107F015202 under Teaching Practice Research Program.

## Compliance with Ethical Standards

**Conflict of Interest** The author declares that she has no conflict of interest.



### Appendix 1. A sample idea map



### Appendix 2. Sample idea matrices

#### I. Contrast

	past	now
Window display	big	
Way people walk		

#### II. Cause-Effect-Solution

	Cause 1: traffic lights	Cause 2: TV & _____
Effect 1	Creates _____	Brain and eye _____
Effect 2	Keep walking _____	Understand _____
Solutions: Store windows must/need to	_____	_____ _____, creative & answer questions
My favorite windows	1. What/Where: _____ 2. Features: _____	

## References

- Alexander, P. A., Dumas, D., Grossnickle, E. M., & List, A. (2016). Measuring relational reasoning. *The Journal of Experimental Education*, 84(1), 119–151.
- Chu, H. C. (2015). Illuminating thinking through post-reading online discussion for EFL learners: an exploratory study. *English Teaching and Learning*, 39(3), 1–32.
- Bernhardt, E. B. (2001). *Understanding advanced second-language reading*. New York: Routledge.
- Crooks, S. M., & Cheon, J. (2013). Strategies for note-taking and computer-based graphic organizers. In G. Schraw, M. T. McCrudden, & D. Robinson (Eds.), *Learning through Visual Displays* (pp. 187–222). IAP Information Age Publishing.
- Daise, D., & Norloff, C. (Eds.) (2015). *Q: Skills for Success, Level 4: Reading and Writing*. (2nd edition). N. Y.: Oxford University Press.
- Danielson, R. W., & Sinatra, G. M. (2017). A relational reasoning approach to text-graphic processing. *Educational Psychology Review*, 29, 55–72.
- Eppler, M. J. (2006). A comparison between concept maps, mindmaps, conceptual diagrams, and visual metaphors as complementary tools for knowledge construction and sharing. *Information Visualization*, 5, 202–210.
- Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, 28, 717–741.
- Grabe, W. (2004). Research on teaching reading. *Annual Review of Applied Linguistics*, 24, 44–69.
- Graesser, A. C., Swamer, S. S., Baggett, W. B., & Sell, M. A. (1996). New models of deep comprehension. In B. Britton & A. Graesser (Eds.), *Models of understanding text* (pp. 1–32). Mahwah: L. Erlbaum.
- Jiang, X. (2012). Effects of discourse structure graphic organizers on EFL reading comprehension. *Reading in a Foreign Language*, 24(1), 84–105.
- Johnson, R. E. (1970). Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Behavior*, 9, 12–20.
- Kansizoğlu, H. B. (2017). The effect of graphic organizers on language teaching and learning areas: A meta-analysis study. *Education and Science*, 42(191), 129–164.
- Kauffman, D. F., & Kiewra, K. A. (2010). What makes a matrix so effective? An empirical test of the relative benefits of signaling, extraction, and localization. *Instructional Science*, 38, 679–705.
- Kiewra, K. A., Mayer, R. E., Dubois, N. F., Christenson, M., Kim, S. I., & Risch, N. (1997). Effects of advance organizers and repeated presentations on students' learning. *The Journal of Experimental Education*, 65(2), 147–159.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition* (pp. 93–100). Cambridge University Press. *Instructional Science*, 38, 679–705.
- Long, D., Seely, M., Oppy, B., & Golding, J. (1996). The role of inferential processing in reading ability. In B. Britton & A. Graesser (Eds.), *Models of understanding text* (pp. 189–214). Mahwah: L. Erlbaum.
- Manoli, P., & Papadopoulou, M. (2012). Graphic organizers as reading strategy: Research findings and issues. *Creative Education*, 3(3), 348–356.
- Maxim III, H. (2002). A study into the feasibility and effects of reading extended authentic discourse in the beginning German language classroom. *The Modern Language Journal*, 86, 20–35.
- Mayer, R. E. (2009). *Multi-media learning* (2nd ed.). New York: Cambridge University Press.
- Mayer, R. E. (2013). Foster learning with visual displays. In G. Schraw, M. T. McCrudden, & D. Robinson (Eds.), *Learning through visual displays* (pp. 47–74). IAP Information Age Publishing.
- McCrudden, M. T., & Rapp, D. N. (2017). How visual displays affect cognitive processing. *Educational Psychology Review*, 29, 623–639.
- Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, 76(3), 413–448.
- Renkl, A., & Scheiter, K. (2017). Studying visual displays: How to instructionally support learning. *Educational Psychology Review*, 29, 599–621.
- 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(3), 455–467.
- Robinson, J. S., & 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., & Skinner, C. H. (1996). Why do graphic organizers facilitate search processes: Fewer words or computational efficiency? *Contemporary Educational Psychology*, 21, 166–180.

28. Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Language and Instruction, 13*, 141–156.
29. Schroeder, N. L., & Cencki, A. T. (2018). Spatial contiguity and spatial split-attention effects in multimedia learning environments: A meta-analysis. *Educational Psychology Review, 30*, 1–23.
30. Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review, 30*, 431–455.
31. Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*, 123–138.
32. Tajeddin, Z., & Tabatabaei, S. (2016). Concept mapping as a reading strategy: Does it scaffold comprehension and recall? *The Reading Matrix, 16*(1), 194–208.
33. Tseng, J. (2010). Designs of concept maps and their impacts on readers' performance in memory and reasoning while reading. *Journal of Research in Reading, 33*(2), 128–147.
34. Werth, P. (1999). *Text worlds: Representing conceptual space in discourse*. Essex: Pearson Education Limited.