

**TAIWANESE SENIOR HIGH SCHOOL STUDENTS' CONCEPTIONS
OF TEXTUAL AND VISUAL ELEMENTS IN ENGLISH TEXTBOOKS:
AN EXPLORATORY STUDY**

Wen-Chiao Yu
Peichin Chang

ABSTRACT

Conceptual beliefs learners hold may influence learning strategies and outcomes of learning. Little research, however, has investigated English as a Foreign Language (EFL) learners' views of multimodality presented in the most accessible learning materials, their English textbooks. A total of sixteen 11th graders participated, half from the language gifted (LG) class and half from the regular social science (RSS) class. They were interviewed to investigate conceptions and how proficiency level regulates conceptions. The phenomenographic method was applied to iteratively analyze the students' utterances. Five qualitatively different categories of conceptions emerged, including, from the lowest in rank, *Read for textual information only*, *Read with engaging visual aids*, *Read bi-modally for facts*, *Read bi-modally to comprehend*, and *Read bi-modally and critically*. The results revealed that bi-modal reading has the potential to enhance comprehension as indicated by the greatest number of conceptions, *Read bi-modally to comprehend*, and motivate learning, as in *Read with engaging visual aids*, ranked the second. Next, investigating the tie of proficiency to conceptions suggests that while the LG group was attuned bi-modally in advanced cognitive processing, the RSS group was more divided, in that some heavily relied on text while others were comfortable using both text and images to read critically.

Key Words: conceptions, multimodality, reading comprehension

INTRODUCTION

Multimodality refers to an approach through which different semiotic resources co-deploy and co-contextualize a text (Kress, 2000). It conceptualizes a way to communicate via the interplay of more than one

mode of semiotic resources to project meaning (Kress, 2000; Kress & van Leeuwen, 2001). Kress (2000) suggested that meanings were made through the use of different modes and that communication between people could be optimized only when various semiotic resources were used concurrently to achieve meaning. Jewitt (2008) explicated that different modal affordances composed meaning in their own ways. Words are sequential and temporal while images are non-sequential and spatial. Meanings are therefore more likely and effectively to get across when words and images work in tandem to promote understanding of texts and provide an alternative reading path (Jewitt, 2008; Kress & van Leeuwen, 1996; Unsworth, 2001).

Nowadays, it is imperative to expand what it means to be literate (Kress, 2004; The New London Group, 1996). There has been a continuous emphasis regarding how language learning materials should be developed to enhance EFL learning for learners with different learning styles (Gardner, 2006; Tomlinson & Masuhara, 2011). EFL learners may enhance their reading ability by capitalizing on their visual intelligence and tap into their strength in the visual domain to develop other intelligences (Arnold & Fonseca, 2004; Gardner, 2006). Visual elements may serve as comprehensible inputs which assist language learners to activate background knowledge, retrieve information and make meanings (Arnold & Fonseca, 2004; Pan & Pan, 2009). Leveraging visual images has become part of the competence indicators listed in the K-12 English national curriculum guidelines in Taiwan (Ministry of Education, 2003). As much as multimodal resources can benefit language learners, when multimodal resources are not properly deployed, multimodal resources might conflict with learners' existing knowledge (Ajayi, 2012). Examining textbooks for Chinese college students, Liu and Qu (2014) therefore proposed that to optimize comprehension, visual-verbal coherence is critical, and urged EFL textbook compilers to properly arrange bi-modal resources in their textbooks. Tsai (2016) also found that a broader way to conceive text-image relation should be in place so that EFL students may be better supported to move beyond factual understanding to critical interpretation of the lessons.

Despite burgeoning research investigating the effect of multimodal resources on vocabulary learning and reading comprehension (Carpenter & Olsen, 2012; Jones, 2004; Kost, et al., 1999; Liu, 2004; Liu & Qu, 2014; Rowe, Silverman & Mullan, 2013; Shalmani & Sabet, 2010; Yanguas, 2009) and on other aspects of language learning or literacy development

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

(e.g. Choi & Yi, 2016; Connors, 2013; Early & Marshall, 2008), little is known related to how EFL learners conceptualize multimodal texts, which the current study undertakes, given that conception may regulate their learning and is critical to learning outcome (Baxter Magolda, 2004). This study therefore investigated how learner conception benefits their learning specifically in recognizing and interpreting visual and textual elements when reading.

The current study sought to investigate EFL senior high school students' conceptions of the text-image relation in English subject textbooks, a major source of their English learning. Learners' conceptions play a critical role in either facilitating or impeding learning (e.g. Baxter Magolda, 2004; Schommer, 1990). The definitions of conceptions vary among scholars: beliefs that one holds (Benson & Lor, 1999); conceptions as a philosophy or epistemological belief concerning the nature and justification of knowledge (Hofer & Pintrich, 2002); conceptions as developmental progression that reflect an individual's ability to regulate their learning (Baxter Magolda, 2004); conceptions as preexisting knowledge and a structure that predisposes learners' thinking (Torff & Sternberg, 2001); and conceptions as a part of underlying metacognitive awareness (Hofer, 2001). Researchers mostly agreed that conceptions are a set of intuitive beliefs that are constructed based on prior experiences and influence learners in a variety of ways, which also guides the current study.

This study was guided by the following research questions:

1. What are the EFL senior high school students' conceptions of textual and visual elements in 11th grade English textbooks?
2. How are the senior high school students' conceptions regulated by their English proficiency levels?

LITERATURE REVIEW

A great number of studies have emphasized the importance of multimodality in language learning (e.g., Kress, 2000; Liu & Qu, 2014; Royce, 2002). Multimodal texts were found to activate and supplement prior knowledge, help organize reading inputs and enhance reading or listening comprehension (Bezemer & Kress, 2016; Kost, et al., 1999; Levie & Lentz, 1982; Royce, 2002). Multimodal language learning also has great potential to engage learners' attention (Choi & Yi, 2016). In the

following review, first, the benefits of multimodal language learning on vocabulary learning and reading comprehension are discussed. Next, studies investigating other aspects of language learning informed by a multimodal approach are presented. While not specifically investigating reading comprehension, these studies shed light on how multimodal resources supported comprehension via audio and the construction of meaning after reading. The third section reports some attempts in textbook analysis related to the textual and visual elements in textbooks. Finally, the promotion of motivation through multimodal language learning is discussed, to supplement the discussion of the cognitive domain of language learning.

Effects of Multimodal Resources on Vocabulary Learning and Reading Comprehension

Multimodal resources were generally found to benefit vocabulary learning. Specifically, combined modalities and visuals were found to enhance vocabulary learning better than textual mode or without any modal aids. Kost, et al., (1999) investigated incidental vocabulary growth in reading aided by textual and/or pictorial glosses. They found that a combination of text and pictures helped the learners to perform better in developing their vocabulary. It was revealed that participants performed better, retained longer and recalled more when more than one mode was employed. Jones (2004) investigated the use of annotations in the recognition and recall of vocabulary while listening in learning a second language. It was found that annotations, both verbally and visually, helped with recognizing vocabulary. Additionally, the participants recalled better if the testing mode matched the kind of annotations received. In all, having access to annotations allowed the participants to perform better than without. The researcher concluded that students should be supported by any mode of information that suits their needs and preferences in listening. Carpenter and Olsen (2012) found that participants learned new vocabulary in a foreign language better when pictures of corresponding words were provided instead of direct translations of the words, on condition that the learners were not overconfident in the power of pictures to help them learn. Rowe, Silverman and Mullan (2013) further investigated gestures in addition to pictures in aiding preschoolers' word learning. They reported that while more advanced learners may learn without the nonverbal aids, learners of low English-language skills

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

benefited from either gestures or pictures. The researchers concluded that the effect of verbal and nonverbal supports varies based on individual differences.

Regarding the benefits of multimodalities on reading, it was generally found that visuals combined with textual aids enhanced reading better than textual aids alone. Similar to the vocabulary learning indicated above, a variety of modes combined better enhanced reading comprehension than a single mode. Liu (2004) investigated the support of bi-modal resources and employed comic strips to assist both high and low proficient L2 learners to read. It was found the lower proficient learners outperformed the higher proficient group (who received the high-level text only) in reading high-level texts aided by comic strips. However, providing the higher proficient students with a comic strip while reading a high-level text did not enhance their recall. In addition, comic strips did not help either higher or lower proficient students in comprehending simpler texts. Specifically, if comic strips do not mirror the text's complexity, they tend to interfere with comprehension. Pan and Pan (2009), however, found that a group of EFL low-proficiency college students not only benefited from pictures in comprehending texts, but the presence of pictures actually supported comprehension of both simple and difficult texts, different from Liu's (2004) findings. Yanguas (2009) investigated the effects of different multimedia glosses on text comprehension and vocabulary learning. It was found that the multimedia gloss groups outperformed the control group in noticing and recognizing more targeted words; however, in producing the targeted words, no significant difference was found in all groups. In reading comprehension, the combination gloss group (text + picture) outperformed all other groups. Similarly, Shalmani and Sabet (2010) compared the effects of vocabulary annotations in textual, visual and a combination of both modes on EFL college students' reading comprehension done by using electronic courseware. They found that compared to the groups aided with a single modal annotation, the visual-textual annotation group performed better in reading tasks. Their finding revealed that the students could decode the texts more easily because of the bi-modal annotations.

Effects of Multimodal Resources on Other Aspects of Language Learning

Different aspects of language learning have also been explored when applying multimodal resources. Guichon and McLornan (2008)

investigated L2 undergraduate students' listening comprehension of BBC audios. The students were exposed to different modes: (1) sound only, (2) image and sound, (3) image, sound and L1 subtitles, and (4) image, sound and L2 subtitles. It was found that the students performed better when exposed to several modalities and L2 subtitles were more helpful than L1 subtitles, which aligns with Kost, et al., (1999) in investigating vocabulary learning. Both Early and Marshall (2008) and Connors (2013) explored how a multimodal approach may assist adolescent and young learners in interpreting and appreciating literary texts. It was found that the students, regardless of age, were able to engage in deeper reading of the literary works or actively draw on both text and images in graphic novels to construct literary meanings.

Visual-verbal Coherence and Reading Comprehension

While learners actively conceptualize what is to be learned, their conceptions may also be affected by the way visual and textual resources are construed. Therefore, stronger text-image complementarity in the texts, if available, may also enhance reading comprehension (Liu & Qu, 2014). The infusion of visual and verbal modes in second language learning textbooks has been an active area of research; however, few have investigated how the two modes work in tandem to deliver meaning. Giaschi (2000) carried out an image analysis in ESL/EFL textbooks and found that a number of messages that were relayed through the images did not seem to be relayed through the written text. ESL/EFL teachers, therefore, should note the distinct affordance transmitted by different semiotic vehicles. Examining textbooks for Chinese college students, Liu and Qu (2014) analyzed the visual and verbal semiotic modes in two representative EFL textbook series for Chinese college students and found that both textbooks demonstrated visual-verbal coherence and that the different bi-modal deployment exhibited in the textbooks may be motivated by the proficiency level of the target learners. They proposed that to optimize comprehension, visual-verbal coherence is critical, and urged EFL textbook compilers to properly arrange bi-modal resources in the textbooks. Investigating whether and how text and image cohere and complement each other in 8th-grade EFL textbooks, Tsai (2016) found that a more literal connection is present in the textbooks and the connection is also quite random. In some lessons, bi-modal complementarity is well achieved, but not in others. The researcher urged that a broader way to

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

conceive text-image relation should be in place so that EFL students may be better supported to move beyond factual understanding to critical interpretation of the lessons.

Multimodal Language Learning and Motivation

From affective perspectives, multimodal pedagogy helped to motivate students, especially those reluctant or struggling learners, to read. Multimodal teaching enables not only cognitive but also affective engagement. Choi and Yi (2016) found that using multimodal materials allowed learners to gain deeper understanding of content knowledge, effectively express what they have learned and more importantly, to obtain a sense of achievement and self-esteem. Garcia, Garcia and Hernandez (2011) investigated four advanced EFL learners' beliefs towards multimodal texts used in foreign language learning and found that texts as such were perceived as promoting critical thinking, motivating, and beneficial in improving communicative skills.

In summary, past studies generally found the benefits of multimodality in guiding learners to comprehend, recall, and retain reading materials. They were also better motivated presented with multimodal materials and able to actively construct meaning after reading. Different modes exhibit different potentials in facilitating reading, and visual-verbal coherence, if achieved, may better enhance reading comprehension. Despite the findings, there is a gap in research which probes learners' conceptions, given that conception may regulate their learning and is critical to learning outcome (Baxter Magolda, 2004). This study therefore examined how language learners conceptualized the interplay of both verbal and visual resources to deliver meanings to fill the gap in the current research in multimodal language learning.

METHOD

Phenomenographic Method

Learners' conceptions play a critical role in either facilitating or impeding learning (e.g. Baxter Magolda, 2004; Schommer, 1990). To generate conceptions, the phenomenographic method has often been operationalized to understand qualitatively different ways of understanding and various aspects of what people experience and how

people conceptualize things (Marton, 1986). The phenomenographic approach combines interview, protocol, and discourse analysis and is used to “identify qualitatively different, hierarchically related conceptions of learning” (Tsai, 2009, p. 1094) among a “small, purposive sample of subjects” (Marton & Booth, 1997, p. 111). Learners’ conceptions are often classified into qualitatively different categories, which are then ranked into a hierarchy of conceptions. In the hierarchy, more advanced conceptions may resemble those of the experts while those lower in the hierarchy may be less relevant and precocious (Marton, 1986).

Participants

A total of sixteen EFL eleventh graders at a senior high school in central Taiwan were recruited to participate. On average, all of the participants in the study had about ten years of English learning experience. To investigate the role proficiency plays, half of the participants were recruited from the language gifted (LG) class while the other half of the participants were from a regular social science (RSS) class. The participants from the LG class presumably had a higher aptitude for language and language proficiency¹. In addition, prior to and during their participation, they also received more training in all four aspects of the English language than the other students. For example, in Grade 10 and 11, they had “English listening and speaking”, “Oral communication in English” and “British and American literature and drama” classes. From the monthly exam in October, 2016 to that in March, 2017, the students in the LG class had an average score of 65, which is considerably higher than those in the RSS class, with an average of 50. The monthly exams were carefully written to consistently evaluate all 11th grade students’ performances and included vocabulary, grammar, cloze and reading comprehension sections, and were not limited to lessons in the textbooks. All of the participants were assigned a number for identification purposes.

¹ To be qualified for the LG class, the students needed to (1) obtain A+ or above on English in the Comprehensive Assessment program (CAP), (2) win in the top three places in nationally or city-based English speech contests, (3) acquire the certificate of GEPT Intermediate Level or pass the first stage of GEPT High-intermediate Level, or (4) rank high in admission by recommendation and screening exam which considers Chinese and English scores in CAP, Chinese subject test, English subject test and an oral test.

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

Interviews

The interview involved two parts. Part one contains general questions regarding the participants' English learning experiences and background. Part two, also the main part of the investigation, solicited the participants' views related to their English reading habits and experiences (see Appendix A). The interview questions for assessing conceptions of textual and visual elements in English textbooks were modified from Tsai (2004, 2009) and Yang and Tsai (2010). Prior to the interviews, potential interview questions were examined and piloted with five students². The questions were then refined for formal investigation. On average, each interview lasted thirty to forty minutes. All of the interviews were conducted in Mandarin and audio-recorded. The recorded contents were transcribed for further analysis.

In the interview, the students drew on Book 4, LungTeng edition for senior high school, in use when the interview took place. Among these lessons, some of them were referred to more often than others. For instance, many referred to Lesson 4 while few mentioned Lessons 6 and 7. Appendix B shows some snapshots of the lessons along with the utterances referring to the snapshots.

Analysis

To understand conceptions, the researchers first extracted relevant utterances in the verbatim transcripts related to conceptions and highlighted the keywords that exemplified the participants' views. In segmenting, relevant utterances were labeled as ideas. An idea, the basic unit for analysis operationalized in the current study, is counted when it is unique and non-repetitive. An idea could be a phrase, a sentence, or the main idea of several sentences. Each idea showed a unique way of viewing the phenomenon being studied. The following instance demonstrates what qualifies as an idea:

Teacher: How do you make sure you have learned something from a lesson?

² The questions were carefully designed and piloted to ensure that a broader range of conceptions could be elicited, which include conceptions related to both multimodal awareness and absence of such awareness. Our assumption was that both the presence and absence of the conceptions related to multimodality are valid and should be elicited.

Participant: Whenever I pass the test.

Teacher: Which lesson impresses you most?

Participant: This one, Twilight. Because of the movie. Also, *the pictures in the lessons are interesting*. Idea 1: Getting interested

The keywords in the ideas were then assigned one label that represents a conception of text-image relation. Next, similar conceptions were sorted and grouped to form qualitatively different categories. All of the research data were sorted through rounds of discussions before any attempt to delineate categories of conceptions was made. Multiple iterative rounds of grouping and regrouping were performed to make sure each category adequately represents the depth and breadth of the participants' views. Finally, the stabilized categories of conceptions were further arranged in hierarchical order (Tsai, 2004). The number of conceptions was then calculated and rendered as a percentage. As a whole, the data yielded 57 conceptions.

The process was validated by a second coder, the second author of this study. The two coders independently segmented and coded one fifth of the total transcripts based on the agreed criteria. Inter-coder reliability was counted and the agreement between two coders is 86%³, indicating acceptable reliability. Disagreement and differences of coding and grouping were discussed and reconciled through further discussion. The first author then completed the rest of the coding.

RESULTS

The first research question investigated a group of EFL high school students' conceptions of textual and visual elements in English textbooks. Through phenomenographic analysis, five different categories of conceptions emerged and were ranked in hierarchical order, including Read for textual information only, Read with engaging visual aids, Read bi-modally for facts, Read bi-modally to comprehend, and Read bi-modally and critically. The first category involves little awareness related to multimodality, thus occupying the lowest rank in the hierarchy. The second category, while involving bi-modal awareness, was ranked low

³ Reliability was calculated using the formula: $2M/(N1+N2)$, (Holsti, 1969) and M represents the number of agreed coding by the two coders while N1 and N2 refer to the numbers of coding done by the two coders respectively.

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

because most of the responses are quite shallow, in which the students simply expressed, for example, whether they like something or whether something is interesting. By contrast, the third, fourth and fifth categories involve different and more heightened levels of multimodal awareness. Table 1 presents the five categories, along with the subcategories.

Table 1

Frequencies of Conceptions of Textual and Visual Elements in the Eleventh Grade English Textbooks

Conceptions	Definitions	n	Frequency (%)
1. Read for textual information only	Processing textual mode to know and memorize content	5	8.7
(1) Textual understanding		4	7
(2) Words memorization		1	1.7
2. Read with engaging visual aids	Affective responses to the lessons	13	22.8
(1) Being attracted by the visuals		7	12.3
(2) Getting interested in the visuals		5	8.8
(3) Feeling familiar with the visuals		1	1.7
3. Read bi-modally for facts	Using both textual and visual modes to retrieve factual knowledge	7	12.3
(1) Remembering content		3	5.3
(2) Recalling content		3	5.3
(3) Retaining content		1	1.7
4. Read bi-modally to comprehend	Using textual and visual modes to enhance comprehension	21	36.9
(1) Understanding the meaning		11	19.3
(2) Guessing the meaning		7	12.3
(3) Anticipating content		3	5.3
5. Read bi-modally and critically	Analyzing and evaluating of the bi-modal connection	11	19.3
(1) Making connections between textual and visual modes		7	12.3
(2) Evaluating the use of textual and visual modes		3	5.3
(3) Organizing content		1	1.7
Total		57	100

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

Table 1 shows that Read bi-modally to comprehend, a conception ranked relatively higher in the hierarchy, has the highest percentage (36.9%), followed by Read with engaging visual aids (22.8%) which involves reaction in the affective domain. Around one fifth of the students (19.3%) conceptualized textual and visual connection as enabling advanced cognitive activity as in Read bi-modally and critically, ranked third in frequency. The lowest in the hierarchy, Read for textual information only, also accounts for the lowest in percentage (8.7%). Combined, the two more advanced conceptions, conceptions 4 and 5, make up more than 50% of the total conceptions. Examples are provided below to delineate each conception. A sequential identifying number is given to represent each student. For example, “#1” refers to “the first participant”.

(1) Read for textual information only

Read for textual information only concerns how the participants read with a text-oriented mindset. This category accounts for the smallest percentage among all (8.7%). The subcategories are ‘Textual understanding’ and ‘Words memorization’. Generally, the participants reported that they relied only on textual mode when reading. For instance,

#1: I can’t see why the pictures are here. They do not help me read and understand. (‘Textual understanding’)

#14: I usually read the article first before noticing the pictures . . . Sometimes I ignore the pictures because I enjoy looking at the pictures only after I fully digest the article. (‘Textual understanding’)

(2) Read with engaging visual aids

Read with engaging visual aids is concerned with whether the participants were aware of and drawn to the visuals in the lessons. Thirteen out of the 16 participants held this view, which constitutes 22.8% of the total conceptions. This category involves ‘Being attracted by the visuals’, ‘Getting interested in the visuals’ and ‘Feeling familiar with the visuals’. ‘Being attracted by the visuals’ refers to the view that the visuals help the participants to focus their attention on the lessons and that the visuals were eye-catching. ‘Getting interested in the visuals’ refers to the fact that the participants became interested in the lessons because of the visuals.

‘Feeling familiar with the visuals’ refers to the familiarity the participants had upon looking at the visuals. For example,

#5: Without pictures in the lesson, I will not even try to read because the lesson will not be interesting. (‘Getting interested in the visuals’)

#7: The pictures of cartoon characters are familiar . . . and are close to our lives. (‘Feeling familiar with the visuals’)

#12: While I look at the pictures, I am attracted by them. And I want to continue learning the rest of the lesson. (‘Being attracted by the visuals’)

(3) Read bi-modally for facts

Read bi-modally for facts emphasized the use of both textual and visual modes to retrieve facts from the lessons. The facts are specific information and factual knowledge. The subcategories include ‘Remembering content’, ‘Recalling content’ and ‘Retaining content’. For instance,

#7: Pictures help me memorize the settings, vocabulary or the content. (‘Remembering content’)

#12: You can look at the pictures to recall what the lesson is about before taking tests. (‘Recalling content’)

#7: I need pictures to leave deeper impression about the lesson. (‘Retaining content’)

(4) Read bi-modally to comprehend

In Read bi-modally to comprehend, the participants employed both textual and visual modes to enhance reading comprehension. This category, accounting for 36.9% of the total conceptions, is highest in frequency and concerns a higher cognitive level, involving ‘Understanding the meaning’, ‘Guessing the meaning’ and ‘Anticipating content’. This conception differs from Read bi-modally for facts in that Read bi-modally for facts emphasizes the use of both textual and visual modes to retrieve facts in the lessons. The facts referred to here are often specific information and factual knowledge such as settings and

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

vocabulary⁴. However, Read bi-modally to comprehend requires more cognitive effort to engage in a higher level of thinking as opposed to merely retrieve or look for facts. For instance,

#3: Pictures guide you to read this article but they become supportive after reading. ('Understanding the meaning')

#10: I usually look at the pictures before I read...The pictures may have something to do with the title. ('Anticipating content')

#15: The pictures are used to take a guess of the content before reading and are used to assist me to read. ('Guessing the meaning')

(5) *Read bi-modally and critically*

Read bi-modally and critically concerns analyzing and evaluating bi-modal connections. This category, accounting for 19.3% of the total conceptions, features critical thinking and is ranked highest in the hierarchy. The subcategories involve 'Making connections between textual and visual modes', 'Evaluating the use of textual and visual modes', and 'Organizing content'. For example,

#3: I tend to keep reading to see if there are any pictures that can match the descriptions in the article. Some pictures were matched with the content of the article. Others had little connection. And still others were quite different from what I previously guessed. ('Making connections between textual and visual modes')

#2: If I see the pictures, I will think that the lesson is about travel. Actually, these pictures have something to do with the lesson. But the lesson is

⁴ Our classification also resonates with Bloom's taxonomy which identifies six forms of thinking, in which "Remembering" occupies the lowest rank and "Understanding" is a rank above. The subcategories of *Read bi-modally for facts*, 'Remembering content', 'Recalling content' and 'Retaining content', also resonate with the keywords related to "Remembering" in Bloom's taxonomy (e.g. recite, recall, and repeat, etc.). By contrast, the subcategories of *Read bi-modally to comprehend* involving 'Understanding the meanings', 'Guessing the meanings' and 'Anticipating content', resonate with a different set of keywords characterizing "Understanding" in the taxonomy (e.g. explain, comprehend, and predict, etc.) (Clark, 1999).

mainly about jobs. ('Evaluating the use of textual and visual modes')

#7: When there are lots of main ideas, I can organize the content by using pictures. ('Organizing content')

The second research question investigates how English proficiency regulates students' conceptions of textual and visual elements. Of the 16 participants, half of them were recruited from the LG class and the other half were from the RSS class. Overall, because the RSS group generated more conceptions than the LG group, the group outnumbers the LG group in all categories except one, also the highest in the hierarchy, Read bi-modally and critically. It was found that both groups exhibited a similar distribution of conceptions with the most frequent being Read bi-modally to comprehend and the least, Read for textual information only. The LG group demonstrated fewer conceptions in the two conceptions in the lower hierarchy, Read for textual information only and Read bi-modally for facts. In contrast, the RSS group is more divided in the distribution. Table 2 gives the respective distributions.

The two groups showed greater gaps in two categories. First, considerably more RSS students relied solely on textual message to comprehend than the LG students (Read for textual information only: 7.02% vs 1.75%). And when reading bi-modally, the number of RSS who read for facts also doubled that of LG students (Read bi-modally for facts: 8.77% vs 3.51%). However, the number of the two most advanced conceptions combined is very close between the groups, suggesting that while generally the LG group was attuned bi-modally in advanced processing (i.e., read to comprehend and read critically), the RSS group was more divided; some relied heavily on text only while others were comfortable using both text and image to read critically.

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

Table 2

Frequency of Conceptions in the RSS and LG Groups

Conceptions	RSS	LG	Total
1. Read for textual information only	4 (7.02%)	1 (1.75%)	5 (8.77%)
(1) Textual understanding	3 (5.26%)	1 (1.75%)	4 (7.02%)
(2) Words memorization	1 (1.75%)	0	1 (1.75%)
2. Read with engaging visual aids	7 (12.28%)	6 (10.53%)	13 (22.81%)
(1) Being attracted by the visuals	3 (5.26%)	4 (7.02%)	7 (12.28%)
(2) Getting interested in the visuals	4 (7.02%)	1 (1.75%)	5 (8.77%)
(3) Feeling familiar with the visuals	0	1 (1.75%)	1 (1.75%)
3. Read bi-modally for facts	5 (8.77%)	2 (3.51%)	7 (12.28%)
(1) Remembering content	2 (3.51%)	1 (1.75%)	3 (5.26%)
(2) Recalling content	3 (5.26%)	0	3 (5.26%)
(3) Retaining content	0	1 (1.75%)	1 (1.75%)
4. Read bi-modally to comprehend	12 (21.05%)	9 (15.79%)	21 (36.84%)
(1) Understanding the meaning	6 (10.53%)	5 (8.77%)	11 (19.30%)
(2) Guessing the meaning	4 (7.02%)	3 (5.26%)	7 (12.28%)
(3) Anticipating content	2 (3.51%)	1 (1.75%)	3 (5.26%)
5. Read bi-modally and critically	5 (8.77%)	6 (10.53%)	11 (19.30%)
(1) Making connections between textual and visual modes	3 (5.26%)	4 (7.02%)	7 (12.28%)
(2) Evaluating the use of textual and visual modes	2 (3.51%)	1 (1.75%)	3 (5.26%)
(3) Organizing content	0	1 (1.75%)	1 (1.75%)
Total	33 (57.89%)	24 (42.11%)	57 (100.00%)

Figure 1 further presents the data in line chart.

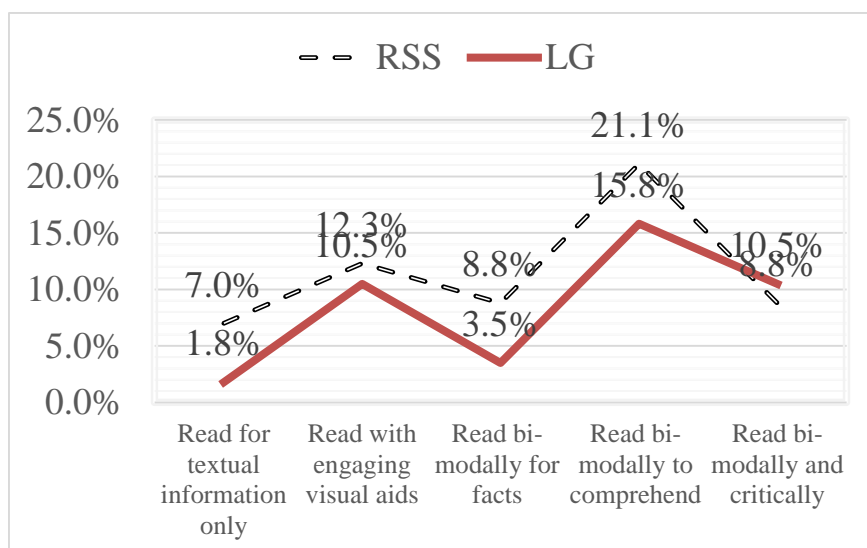


Figure 1. Percentage of RSS and LG groups' conceptions

In summary, more of the conceptions the students contributed are associated with the more advanced cognitive strategies, which involve attempts to comprehend and critically evaluate the lessons while reading. In addition, their conceptions related to the positive affective effect of bi-modal resources are also frequent. In contrast, fewer students relied solely on texts to comprehend. This is interesting and contrary to the knowledge that text still dominates the learning scene. Regarding how proficiency level may regulate conceptions, while the distributions are similar between the groups, the LG group is more concentrated in the two more advanced conceptions, Read bi-modally to comprehend and Read bi-modally and critically while the RSS group is more divided between the lower level conception and the more advanced conception.

DISCUSSION

The present study investigated a group of 16 Taiwanese senior high school students' conceptions of textual and visual elements in English textbooks and the tie between the students' conceptions and their proficiency levels. The results revealed five major conceptions, ranked

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

hierarchically, Read bi-modally and critically, Read bi-modally to comprehend, Read bi-modally for facts, Read with engaging visual aids, and Read for textual information only, among which Read bi-modally to comprehend and Read with engaging visual aids are the two most frequent conceptions. This indicates that bi-modal resources have potential in engaging learners both cognitively and affectively.

Textual and Visual Elements and Reading Comprehension

The results of the current study align with the studies which investigated the outcome of students' learning supported by multimodal resources, ranging from their potential to stimulate recall and retention of vocabulary, enhance comprehension, meet individual's learning needs and motivate learning (Choi & Yi, 2016; Garcia, et al., 2011; Jones, 2004; Kost, et al., 1999; Liu, 2004; Pan & Pan, 2009; Shalmani & Sabet, 2010; Yanguas, 2009). This suggests that the co-existence of text and image may not only facilitate language learning both cognitively and affectively, but learners' conceptions of such bi-modal resources may also contribute to their learning. The study found that cognitively, more of the students' conceptions concerned more advanced cognitive processing, as in taking advantage of the bi-modal connection to comprehend content and in evaluating the different affordances of both modes. This corresponds to many studies which indicated the affordances of multimodal language learning in enhancing comprehension (Liu, 2004; Pan & Pan, 2009; Shalmani & Sabet, 2010; Yanguas, 2009) and with Garcia, et al. (2011), who found that EFL students perceived multimodal texts for foreign language learning as promoting critical and reflective thinking.

However, while a great deal of past research indicated the benefits of multimodality in vocabulary learning in terms of how learners better recall and retain words (e.g., Carpenter & Olsen, 2012; Jones, 2004; Kost, et al., 1999), the conceptions associated with this aspect of multimodal language learning are few, taking up only 8.7% of the total conceptions and also occupies the lowest rank in the hierarchy. Therefore, given that the majority of the conceptions involve more advanced cognitive abilities related to comprehension and critical evaluation and that learners' conceptions play a critical role in facilitating learning (e.g. Baxter Magolda, 2004; Schommer, 1990), there may be a need for teachers to facilitate and support learning associated with text comprehension and critical reading.

Additionally, a great proportion of the conceptions involved the positive affective effect images bring, which enhances the students' motivation in reading, also indicated in past research (Choi & Yi, 2016; Garcia, et al., 2011). However, it should be noted that the conceptions involved mostly quite superficial affective reactions. In designing learning materials or instruction, caution should be exercised against using visual aids superficially as mere attractive or eye-catching cues. Given that some students pointed out that when different modes simply occur in texts and do not work in concert to project a coherent meaning, they would be confused and uninspired, it is suggested that visuals should be discussed in a way that complements the texts in class so that the different modes may better support reading comprehension (Liu & Qu, 2014).

Slight influence of proficiency on learners' views of textual and visual elements

Generally, the LG group exhibits the tendency of having more advanced conceptions and fewer lower level conceptions while the RSS group is more divided between advanced and lower level conceptions. Specifically, more RSS students conceptualized reading as relying solely on the textual mode or as attempts to retrieve facts only, the two less sophisticated conceptions on the hierarchy. This may warrant further research to tease out whether these conceptions pertain to multimodal learning or language learning in general. Regardless, despite this tendency, it should be noted that the difference in distribution of the conceptions is slight. Liu (2004) indicated that visuals may assist learners of both low and high proficiency levels if the visuals appropriately support the text. While visuals did not help the high proficiency students in recalling difficult text, visuals did better support the low proficiency students. This echoes our finding that more RSS students took advantage of visuals to help them retrieve facts, which may subsequently support more advanced processing such as comprehending and evaluating later. However, Pan and Pan (2009) found that visuals supported low proficiency learners to comprehend both simple and difficult texts, which also resonates with the current study which found that RSS students utilized visuals to engage in both lower level processing such as retrieving facts and higher level processing such as comprehending and even critically evaluating texts.

Conceptions in Hierarchical Order

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

The present research revealed that some learners displayed a mix of advanced and unsophisticated conceptions. In other words, a student could hold a more advanced conception and a less developed conception at the same time. This indicated that the conceptions emerging in the current study might not have strict qualitative differences, compared to previous conception research. Given that the present study explored textual and visual elements in one major source, the language learning textbooks, the research scope may be comparatively narrower than exploring conceptions of learning (Saljo, 1979), conceptions of science learning (Tsai, 2004) and conceptions of language learning (Benson & Lor, 1999; Yang & Tsai, 2016), which covered a broader and more general scope. Instead, the current study is more focused, which therefore yields the conceptions which are hierarchically ordered based on how the students approached reading predominantly in cognitive terms. Thus, it is inferred that the present study might not yield conceptions of qualitative differences, which explains why a student can exhibit sophisticated and less mature conceptions. Again, the present research is exploratory and offers only preliminary evidence on students' conceptions of multimodal learning resources. The research scope can be expanded as discussed above and can also include materials other than textbooks to be more comprehensive.

In summary, this study revealed that EFL 11th graders' conceptions of multimodal reading involved more advanced cognitive processing, i.e., comprehending and critical reading, on one hand and affective domain on the other. This aligns with many other studies which investigated the learning outcomes of multimodal reading (Liu, 2004; Pan & Pan, 2009; Shalmani & Sabet, 2010; Yanguas, 2009). In addition, the finding that the conceptions involved more advanced cognitive processing, regardless of proficiency levels, may inform reading instruction to prioritize text comprehension and critical reading beyond vocabulary memorization.

PEDAGOGICAL IMPLICATIONS, LIMITATIONS, AND CONCLUSION

Some pedagogical implications can be drawn from the present study. The current study revealed that the students tended to emphasize how text and image assisted them to comprehend texts. Given this, stronger text-image complementarity in the texts, if available, may better enhance reading comprehension, which is also supported by Liu and Qu (2014). It is also vital for EFL learners to understand how to read bi-modal texts by

tapping into the potential of textual and visual relations to enhance reading comprehension. EFL instructors should draw students' attention to how the two modes work in tandem to stimulate critical thinking.

The study was an attempt to describe a phenomenon to gain insights into the EFL students' conceptions of textual and visual elements in English textbooks for reading comprehension. In the future, more research can be done to examine the same phenomenon on a larger scale and with different aims to provide a more comprehensive picture of how English learners interact with multimodal learning materials. In addition, the students' prior learning experiences, if explored, may help us trace the source of their conceptions of reading multimodally. Moreover, to extend the research scope and respond to the prevalence of multimodal communications in the virtual domain, technology-assisted language learning materials can also be incorporated in future investigations. However exploratory, this study has offered some insight into EFL learners' views of bi-modal resources in printed materials, and future research along this line is warranted.

To conclude, the present study aimed to investigate how EFL senior high students viewed textual and visual elements in textbooks and how language proficiency regulated their conceptions. The students generally held that the bi-modal resources helped to promote reading comprehension, a rather advanced cognitive behavior as opposed to reading focusing on the textual mode only. The present study also found that the tie between English proficiency and the students' conceptions was moderate. In EFL reading, teachers may tap into the variety of conceptions the students held and offer more support to enhance their learning. Awareness of how visuals and texts work together may promote reading comprehension and multimodal literacy.

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

REFERENCES

- Ajayi, L. (2012). How teachers deploy multimodal textbooks to enhance English language learning. *TESOL Journal*, 6(1), 16-35.
- Arnold, J., & Fonseca, M. C. (2004). Multiple intelligence theory and foreign language learning: A brain-based perspective. *International Journal of English Studies*, 4(1), 119-136.
- Baxter Magolda, M. B. (2004). Evolution of a constructivist conceptualization of epistemological reflection. *Educational Psychologist*, 39(1), 31-42.
- Benson, P., & Lor, W. (1999). Conceptions of language and language learning. *System*, 27(4), 459-472.
- Bezemer, J. J., & Kress, G. R. (2016). *Multimodality, learning and communication: A social semiotic frame*. London and New York: Routledge.
- Carpenter, S. K., & Olson, K. M. (2012). Are pictures good for learning new vocabulary in a foreign language? Only if you think they are not. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(1), 92-101.
- Choi, J., & Yi, Y. (2016). Teachers' integration of multimodality into classroom practices for English language learners. *TESOL Journal*, 7(2), 304-327.
- Clark, D.R. (1999). Bloom's Taxonomy of Learning Domains. Retrieved from <http://www.nwlink.com/~donclark/hrd/bloom.html>
- Connors, S. P. (2013). Weaving multimodal meaning in a graphic novel reading group. *Visual Communication*, 12(1), 27-53.
- Early, M., & Marshall, S. (2008). Adolescent ESL students' interpretation and appreciation of literary texts: A case study of multimodality. *The Canadian Modern Language Review*, 64(3), 377-397.
- Gardner, H. (2006). *Multiple intelligences: New horizons*. New York: Basic Books.
- García, D., García, J., & Hernández, Y. (2011). Students' beliefs: Multimodal texts as pedagogical tools in foreign language learning. *PAPELES*, 3(5), 21-35.
- Giaschi, P. (2000). Gender positioning in education: A critical image analysis of ESL texts. *TESL Canada Journal*, 18(1), 32-46.
- Guichon, N., & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for CALL resource design. *System*, 36(1), 85-93.
- Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review*, 13(4), 353-383.
- Hofer, B. K., & Pintrich, P. R. (2002). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, NJ: Lawrence Erlbaum.
- Holsti, O. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley.
- Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, 32(1), 241-267.
- Jones, L. (2004). Testing L2 vocabulary recognition and recall using pictorial and written test items. *Language Learning and Technology*, 8(3), 122-143.
- Kost, C. R., Foss, P., & Lenzini, J. J. (1999). Textual and pictorial glosses: Effectiveness

- on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals*, 32(1), 89-97.
- Kress, G. (2000). Multimodality: Challenges to thinking about language. *TESOL Quarterly*, 34(2), 337-340.
- Kress, G. (2004). Reading images: Multimodality, representation and new media. *Information Design Journal*, 12(2), 110-119.
- Kress, G. R., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. London: Routledge.
- Kress, G., R., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Edward Arnold.
- Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *ECTJ*, 30(4), 195-232.
- Liu, J. (2004). Effects of comic strips on L2 learners' reading comprehension. *TESOL Quarterly*, 38(2), 225-243.
- Liu, X., & Qu, D. (2014). Exploring the multimodality of EFL textbooks for Chinese college students: A comparative study. *RELJ Journal*, 45(2), 135-150.
- Marton, F. (1986). Phenomenography—A research approach to investigating different understandings of reality. *Journal of Thought*, 21(3), 28-49.
- Marton, F., & Booth, S. (1997). *Learning and awareness*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Ministry of Education. (2003). *General guidelines of Grades 1-9 curriculum for elementary and junior high school education*. Taipei: Ministry of Education.
- Pan, Y. C., & Pan, Y. C. (2009). The effects of pictures on the reading comprehension of low-proficiency Taiwanese English foreign language college students: An action research study. *VNU Journal of Science, Foreign Languages*, 25(3), 186-198.
- Rowe, M. L., Silverman, R. D., & Mullan, B. E. (2013). The role of pictures and gestures as nonverbal aids in preschoolers' word learning in a novel language. *Contemporary Educational Psychology*, 38(2), 109-117.
- Royce, T. (2002). Multimodality in the TESOL classroom: Exploring visual-verbal synergy. *TESOL Quarterly*, 36(2), 191-205.
- Säljö, R. (1979). Learning about learning. *Higher Education*, 8(4), 443-451.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82(3), 498-504.
- Shalmani, H. B., & Sabet, M. K. (2010). Pictorial, textual, and picto-textual glosses in E-reading: A comparative study. *English Language Teaching*, 3(4), 195-203.
- The New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66(1), 60-93.
- Tomlinson, B., & Masuhara, H. (Eds.). (2011). *Research for materials development in language learning: Evidence for best practice*. London: Bloomsbury Publishing.
- Torff, B., & Sternberg, R. J. (2001). *Understanding and teaching the intuitive mind: Student and teacher learning*. Mahwah, NJ: Lawrence Erlbaum.
- Tsai, C. C. (2004). Conceptions of learning science among high school students in Taiwan: A phenomenographic analysis. *International Journal of Science Education*, 26(14),

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

1733-1750.

- Tsai, C. C. (2009). Conceptions of learning versus conceptions of web-based learning: The differences revealed by college students. *Computers & Education*, 53(4), 1092–1103. doi:10.1016/j.compedu.2009.05.019.
- Tsai, H. J. (2016). *Multimodal discourse analyses of junior high school English textbooks in Taiwan*. (Unpublished master's thesis). National Taiwan Normal University, Taiwan.
- Unsworth, L. (2001). *Teaching multiliteracies across the curriculum*. Buckingham: Open University Press.
- Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning & Technology*, 13(2), 48-67.
- Yang, Y. F., & Tsai, C. C. (2010). Conceptions of and approaches to learning through online peer assessment. *Learning and Instruction*, 20(1), 72-83.

Wen-Chiao Yu & Peichin Chang

ACKNOWLEDGMENT

We would like to thank the anonymous reviewers for their constructive comments to help improve the earlier versions of this paper.

CORRESPONDENCE

*Wen-chiao Yu, Department of English, National Taiwan Normal University, Taiwan
Email address: joe18amycom@gmail.com*

*Peichin Chang, Department of English, National Taiwan Normal University, Taiwan
Email address: peichinchang@ntnu.edu.tw*

PUBLISHING RECORD

Manuscript received: January 17, 2019; Revision received: July 29, 2019; Manuscript accepted: August 8, 2019

APPENDIX

Appendix A. Interview questions

Part I. General questions regarding the participants' English learning experiences and background

1. How long have you learned English?
2. How many years of English reading experience do you have?

Part II. Specific questions about English reading process

3. How do you describe the way you comprehend the lessons in English textbooks?
4. How do you start reading a lesson in English textbooks? How do you use the English textbooks at hand to learn English?
5. When you read a lesson in English textbooks, what do you pay attention to?
6. How do you describe the lessons that impress you the most in the English textbooks?
7. What elements in English textbooks raise your reading motivation? Why?
8. What should an effective lesson in English textbooks look like?
9. How do you know when you have learned something in a lesson in English textbooks?
10. Why are visual elements used in English textbooks? How and when should the visual elements be deployed in English textbooks?
11. What role do visual designs play in reading a lesson in English textbooks? What are the functions of incorporating visual images into English textbooks?
12. Based on your English reading experiences, how do you describe the relations between images and words in a lesson?

Appendix B. Snapshots of visuals in Book 4, LungTeng edition

SNAPSHOTS	CORRESPONDING UTTERNACES
Lesson 4 	T: 課本和考卷的排版有什麼不同？ S: 課本有時候會有圖，大致上可以猜個意思，但是考卷全部都是白紙黑字，看不下去，或是我看過一大段，但都不知道他是什麼東西。 T: 對，像是這邊有兩張圖，有幫助你什麼嗎？ S: 這個吧 T: 電影封面 S: 我一看就知道在講哪一個故事，再加上這個圖，後面這三個看起來就是很陰險的樣子，應該和他們不合，是壞人。 T: 自己讀文章是什麼樣的過程？ S: 我覺得讀文章很累. . .比較不習慣的文字我會覺得很累，只要出現 1, 2 個不會的我就會覺得不舒服。 T: 課本有圖片會好一些？. . .可以用幾個例子來告訴我嗎？ S: 因為第一件事情，我會先看圖片多不多，因為可以先猜，打開課本我就會先看圖片，想是第一課這個，這個，和這個就是在講非洲取水困難嘛. . .還有他們在排隊，或許是和環境有關係，還有這個 T: 環保標章 S: 嗯，環保標章，手機，或許可以猜一下. . . [學生敘述了一下左方圖的內容] T: 我注意到你都是指著圖篇來講故事，圖片重要嗎？
Lesson 1 	
Lesson 4	

CONCEPTIONS OF TEXTUAL AND VISUAL ELEMENTS

(continued)

S: 喚起記憶



Lesson 2



▲ Taiwan



▲ Argentina

Lesson 2



▲ Argentina

T: 你會先預習課文，視覺方面呢？你會看一下版面或是圖表嗎？

S: 還好，我剛是找單字久找很久了

T: . . . 有任何印象深刻的課文嗎？

S: (第一課)在講名字，很貼近，有文化對比

T: 你對文化差別有興趣，你怎麼開始讀一篇課文？

S: 我會先看標題，畫不會字詞，然後開始查

T: . . . 這些圖在視覺上有輔助或幫助嗎？

S: 還好，比較容易疏忽掉

S: 圖片還有什麼功能？嗯. . . 就像電影. . . 嗯，就像生物課本一樣，動脈那些，心臟血怎麼流，比較能理解，搭配圖一起看比較好，看文字的話只能自己想像

T: 從圖片中，假設是這隻牛，你可以知道什麼訊息？

S: 一開始介紹這個地方，或是這些動物和牠們的習性，看到文字不是很理解，可能就會想事不是有和這個圖有關係的，這樣更能理解