Visual Literacy in Secondary EFL Education in China: A Semiotic Perspective

Haizhen Wang
Soochow University, China

Abstract
Human societies use a variety of semiotic modes for meaning construction. The ubiquity of images in today's world, alone or in combination with other modes, makes competence with images or visual literacy a prerequisite of competence in life. Now identified as an essential literacy skill for 21st century learners in the digital age, visual literacy has been introduced into contemporary English curricula in a number of countries around the world. In China, the skill of viewing is added as one important component of language competence in *National English Curriculum Standards for High School* issued in 2017, thereby making the cultivation of visual literacy vital to secondary EFL education. This paper aims to present and discuss the current situation of visual literacy in secondary EFL education in China. It first provides an analysis of the ways of visual-verbal interplay for meaning construction in textbooks, and finds an uneven distribution of image-text relations in the current EFL textbooks, with a majority of images playing a supportive or illustrative role. Then it explicates the bi-dimensional three-level construct of visual literacy that is specified in the EFL curriculum, with which the subskills of visual literacy assessed in the latest large-scale EFL tests are compared, with a view to finding out the degree of correspondence between the subskills measured in the tests and the subskills expected in the curriculum. The result shows that only one dimension of reading visual texts is assessed, and the other dimension of writing visual texts is not evidenced. Finally suggestions are made for varied visual inputs, an expanded construct of visual literacy, and improvement of the washback effect of tests.

Keywords: viewing, visual literacy, EFL education, semiotic

1. Introduction

All that man can know is mediated by signs. Signs can be verbal or non-verbal, and can be drawn from auditory, visual or tactile stimuli. Semioticians who study sign systems would agree that only a small proportion of human communication is verbal and non-verbal communication is abundant in our daily life. Language is but one of the communicative resources through which meaning is (re)made, distributed, and

interpreted (Early et al., 2015; Jewitt, 2008). The pervasiveness of visual signs now in our surrounding world is abundantly obvious thanks to the wide-spread use of the Internet and the ubiquity of new information and communication technologies. As a result, visual communication becomes a recognized division within the study of semiotics and visual literacy becomes a part of education within a semiotic framework (Bopry, 1994). Visual literacy has been identified as an essential literacy skill for 21st century learners in the digital age (Burkhardt et al., 2003). Our postmodernism "is best imagined and understood visually, just as the nineteenth century was classically represented in the newspaper and the novel" (Mirzoeff, 1998, p. 5). The shift from the print culture to the visual culture influences enormously our perception of self, our attitudes, beliefs, values, and our ways of meaning-making and communication. While the current young generation born in the information era, or an era of the so-called "bain d' images" (Avgerinou, 2009, p. 28), are very familiar with the expressions of visual culture, they do not "naturally possess sophisticated visual literacy skills, just as continually listening to an iPod does not teach a person to critically analyze or create music" (Felten, 2008, p. 60). Students should be taught or trained with knowledge and skills of visual literacy. It is high time we changed the reliance on verbal literacy as the sole basis for teaching and testing in schools.

Visual literacy has been included as a key part of literacy education in contemporary English curricula in a number of countries including Australia, Sweden, Denmark, Norway, Finland, Singapore, Canada, and New Zealand (Callow, 2020). In China, it was not until the year of 2017 when the skill of viewing was added to language competence, in parallel with the skills of listening, reading, speaking and writing, in *National English Curriculum Standards for High School* (hereafter the new Curriculum, Ministry of Education of the People's Republic of China, 2020).

The facilitating role of visual input has been widely discussed in the fields of L1 literacy education (Baker, 2001; Hassett & Schieble, 2007; Lewis, 2001; O'Neil, 2011) and second language acquisition (SLA). For example, existing studies on the effect of multimodal input on Chinese college students' EFL learning reveal that the visual input mode enhances short term memory (Zhang, 2010), plays a better role than the aural input mode in listening comprehension (Xiao et al., 2018) and in incidental vocabulary acquisition (Gu & Zang, 2011), and that the text-image-aural input is more effective for incidental vocabulary acquisition than the mono-mode of text input or aural input (Lian & Huang, 2010), and can effectively reduce learners' communication apprehension in vocabulary learning (Zheng & Xu, 2020).

While multiple input modes are available as educational resources, the combination of visual (images) and verbal (language or texts) modes is found the most common multimodal text form used in classroom contexts (Callow, 2020). The verbal and visual modes utilize the meaning-making features peculiar to their respective semiotic systems so as to enable learners a better understanding. However, the image-text relations are complex due to the fact that images afford different ways of shaping knowledge rather than functioning simply as an illustrative feature for the written text (Kress, 1997). Different relations between images and texts therefore extend the ways in which meaning can be constructed. How do images and texts work together to construct meaning? Scholars have discussed at length visual-verbal intersections of meaning construction in L1 literacy education (Callow, 2020; O'Neil, 2011). For example, in Callow's (2020) study of picture books, three visual and verbal

intersections are found: the meanings of each mode may be equivalent to each other, augment each other or at times offer divergent or contradictory meanings. However, EFL learning differs from L1 acquisition in that English in the EFL learning context is not an official language nor used for daily communication, and that the foreign language exposure is much limited to the classroom setting. While picture books and graphic novels are widely used as learning materials for K-12 visual literacy education in America (Hassett & Schieble, 2007; Hassett & Curwood, 2009; O'Neil, 2011) and Australia (Callow, 2008, 2020), textbooks are the primary source of input for EFL education in China. To develop visual literacy, Chinese EFL learners are expected to read in the textbooks images that interact with texts in different ways for meaning construction before they themselves are able to use or compose images in different ways for meaning construction. Since we are not clear about the ways in which images interact with texts for meaning construction in the current EFL textbooks in China, an analysis of image-text relations is therefore necessary. Yet few studies do this. The limited number of unpublished MA theses on this topic studies image-text relations in the Chinese EFL textbooks of either junior high school (Huang, 2020; Liu, 2020) or senior high school (Bao, 2020). None provides a comprehensive analysis of image-text relations in EFL textbooks of both junior and senior high schools.

Teaching goes hand-in-hand with testing or assessment. Tests play an important role in language education. Tests can become the engine for implementing educational policy, and a curriculum change may bring about changes in teaching and testing. The adding of the skill of viewing to the new Curriculum brings about the assessment of visual literacy in EFL tests. A good test "follows and apes the teaching" (Davies, 1968, p. 5), and yet we do not know to what extent the EFL tests in secondary education in China measure the subskills of visual literacy that are required in the new Curriculum.

2. Literature Review

2.1 Visual literacy

The term visual literacy (VL) was coined by John L. Debes in 1969, followed by its first national conference and the foundation of the International Visual Literacy Association (iVLA). The term has now aroused interest in people from such diverse fields as fine arts, graphic design, film studies, communication technology, linguistics and semantics, and education. A multitude of definitions for the concept and the associated skills has arisen from different perspectives. Ausburn and Ausburn (1978, p. 291) define visual literacy as "a group of skills which enable an individual to understand and use visuals for intentionally communicating with others", while Braden and Hortin (1982, p. 169) define visual literacy as "the ability to understand and use images, including the ability to think, learn, and express oneself in terms of images". Whereas a literate person in the traditional sense is able to read and write printed texts, a visually literate person is able to read and write visual language. Reading visual language is "a matter of being able to interpret the visual messages, such as gestures or pictures, produced by others; writing it entails being able to compose meaningful visual messages oneself" (Ausburn & Ausburn, 1978, p. 291). From the above definitions, we may generalize that visual literacy involves two basic skills: understanding (reading) the visual text and using (writing) the visual text, for the purpose of intentional communication. The former involves decoding or deciphering

meaning that is constructed in a visual text produced by others, and the latter involves producing a visual text by oneself in order to encode or construct meaning.

2.2 Visual-verbal interplay

Signs create signs ad infinitum (Eco, 1979). Different modes of signs are interconnected. When we interpret one mode of signs we are in effect creating other modes of signs. Social semiotics sees different semiotic modes, e.g. the verbal and the visual, as "socially and culturally shaped resource for meaning making" (Bezemer & Kress, 2008, p. 171). Meaning no longer lies merely in language because of the occurrence of multiple modes for making sense. Examples of verbal modes are written and spoken language, and those of visual modes are pictures, charts, diagrams, color, font and shapes. The visual and verbal modes have different potentials for making meaning, and they often collaborate to realize complementary or co-dependent intersemiotic meanings in a coherent multimodal text. For example, the visual mode has the potential of invoking the three dimensionality sensed in the perception of objects and of enabling longer memory retention as it taps into emotions to provide direct sensory input (Stöckl, 2004). The virtue of images over words seems to be that they permit socialized knowledge to be closer to immediate apprehension, closer to knowledge by direct acquaintance (Bopry, 1994). In addition, different modes are often used together in one sign for meaning construction. For example, when the three modes of writing, image and color are used in one sign, they play different roles and achieve maximum communication effect: image *shows* what takes too long to read, and writing *names* what would be difficult to show, while color *frames* and *highlights* specific aspects of the overall message (Kress, 2010).

2.3 Martinec and Salway's model of image-text relations

Martinec and Salway (2005) propose an analytical framework of image-text relations by dividing the semantic relations between images and texts into relative status and logico-semantics (Figure 1). Here images refer to the visual mode and texts the verbal mode.

Figure 1. Framework of image-text relations (adapted from Martinec & Salway, 2005, p. 358)

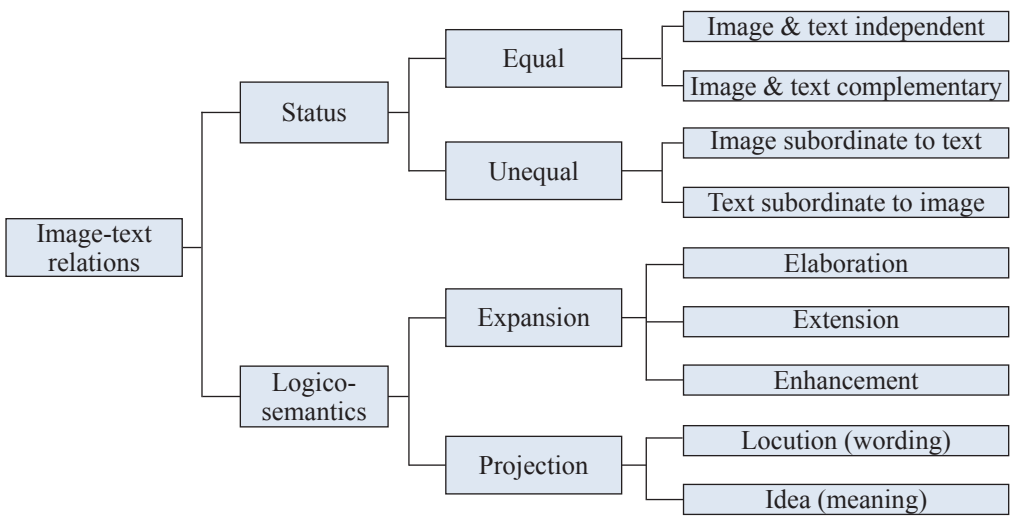


Image-text relations can be equal or unequal in status. Images and texts are considered equal when they are evenly important. The equal status can be independent when one mode exists in parallel with the other or complementary when the two join equally to contribute to the whole. Images and texts are considered unequal when one modifies or is subordinate to the other. It can be the image's subordination to the text or vice versa.

Image-text logico-semantic relations fall into two categories: expansion and projection. Expansion means one mode expands the other mode in meaning construction, and projection means one mode repeats the other mode in meaning construction. Expansion is further divided into three subcategories: elaboration, extension and enhancement. Elaboration refers to the relation where either the image or the text adds detailed information but not always new information to the other; extension refers to the relation where either the image or the text provides new or related information for the other; enhancement relates to a relation where the image or the text enhances the other circumstantially by time, place, reason or purpose. Two types of projection are identified, depending on whether an exact wording is quoted (locution) or an approximate meaning is reported (idea). Locution tends to appear in speech bubbles, while idea is always found in thought bubbles.

This framework of image-text relations is used in the current study for the analysis of visual-verbal interplay in Chinese EFL textbooks.

3. Methodology

3.1 Research questions

This study aims to present and discuss the current situation of visual literacy in secondary EFL education in China, centering around three questions:

(1) What are the ways of visual-verbal interplay for meaning construction in the current EFL textbooks?

(2) What subskills of visual literacy are required in the new Curriculum?

(3) To what extent do the EFL tests assess the subskills of visual literacy required in the new Curriculum?

The answer to the first question helps us evaluate whether different ways of meaning construction are presented in textbooks and whether the visual-verbal interactions in textbooks play their utmost role as input for the cultivation of visual literacy. The combined answers to the second and third questions enable us to judge to what extent EFL tests measure what they are expected to measure concerning the skill of viewing, and hence provide evidence of construct validity. It is hoped that the textbooks provide varied and sufficient input and the tests measure to the greatest degree the construct of visual literacy as expected in the new Curriculum so that an optimal learning condition is set to exert its maximum effect on visual literacy education.

3.2 Data collection and analysis

For the analysis of visual-verbal interplay in textbooks, two sets of EFL textbooks published by People's Education Press are selected as data. These textbooks are

widely used learning materials for secondary EFL education in China. One set of 5 textbooks (*Go for it*) is for junior high school students (Grades 7-9), and the other set of 7 textbooks (*New Senior English for China*) is for senior high school students (Grades 10-12). Each textbook includes five to ten topic-guided units, and each unit consists of several sections focusing on language competence. Comparatively speaking, the five junior high school textbooks are richer in images, composed of visual modes of colorful drawings, photos, cartoon pictures, tables, maps, etc., and the senior high school textbooks are featured by text dominance in quantity. It is noted that there are ten textbooks in the *New Senior English for China* series, and each provincial department of education decides how many textbooks to use for each grade considering the regional difference in education and students' language needs. In a majority of schools Books 1 to 7 are used as learning materials for the purpose of college entrance examination (*Gaokao*) even though the number of textbooks for each grade is different. Therefore, these seven textbooks with no grade distinction are used in this study as illustrative data.

All the images in combination with language are selected from the sections of vocabulary, conversation, reading and exercise in junior high school textbooks, and the section of reading in senior high school textbooks. Data analysis includes both descriptive statistics of frequency and percentage and text analysis of image-text relations.

To work out the construct of visual literacy in secondary EFL education in China, we first make a list of all the subskills related to viewing as specified in the new Curriculum, and based on which we then generalize a bi-dimensional three-level construct of visual literacy.

For the assessment of visual literacy, EFL test papers for *Zhongkao* and *Gaokao* are collected. *Zhongkao*, the high school entrance exam, and *Gaokao*, the college entrance exam, are the most influential and high-stakes national tests for secondary EFL education in China. *Zhongkao* test papers are developed by each city and *Gaokao* test papers by provincial educational department or the Ministry of Education. Here we choose as data 13 *Zhongkao* test papers in 2020 which are developed separately by 13 cities in Jiangsu Province, a province strong in education, and 12 *Gaokao* test papers used in the past three years (2018-2020), including 9 national test papers developed by the Ministry of Education and 3 provincial test papers developed by Jiangsu Province. These test papers are analyzed manually to find out what subskills of visual literacy are measured in EFL tests. Then we match the subskills of visual literacy measured in EFL tests with the construct specified in the new Curriculum so that we can see whether there is a correspondence between the subskills of visual literacy measured and the subskills of visual literacy required or expected. The degree of correspondence reflects the construct validity of these tests.

4. Results

4.1 Visual-verbal interplay for meaning construction in EFL textbooks

On the whole, all the four types of image-text relative status relations are found in the textbooks, indicating that the students are given varied input. The addition of images into textbooks affords different ways of making meaning and shaping knowledge

by adding an additional level of information to be decoded. Meanwhile, we find an uneven distribution of image-text relations in status. Almost half of the images are used to support or illustrate the text by providing greater detail or visual description for the reader (*image subordinate to text*); over a quarter of images interact with the text to provide parallel information in different modes to co-construct meaning (*image-text independent*); a few images dominate in meaning construction (*text subordinate to image*); and images are least often used to be interdependent on the text and complement each other in meaning making (*image-text complementary*). That is to say, not all the ways of meaning construction by the visual mode are sufficiently input into the textbooks.

Of the two categories of image-text logico-semantics relations, expansion occurs more frequently than projection in the textbooks. Of the three types of expansion, elaboration and extension are rich in quantity, which means that the image or the text adds either detailed or new information to the other. The co-construction of meaning by words and images helps learners develop a better understanding.

4.1.1 Image-text relative status relations

In junior high school textbooks the image-text relation is found more unequal (77.41%) than equal (22.59%) in status (Also refer to Huang, 2020). From Table 1, we can see that of 270 cases of image-text combinations, a majority (58.15%) shows image subordination to text, and image-text complementary relation is found the least frequent (10%). In addition, the higher the grade, the fewer the images are. That is, with the increase of age, cognitive and psychological development, teenagers are expected to rely less on images than on texts.

Table 1. Image-text relative status relations in junior high school textbooks

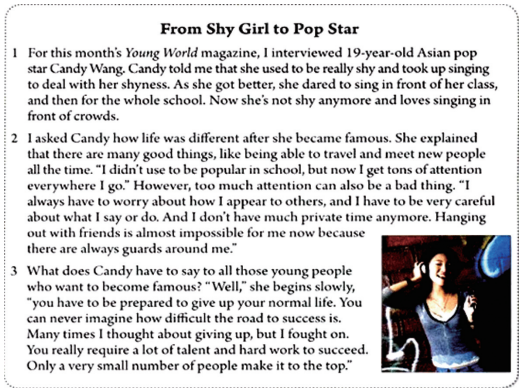
Grade \ Status	Equal		Unequal		Total
	Image & text independent	Image & text complementary	Image subordinate to text	Text subordinate to image	
7	22	13	48	41	124
8	10	7	61	6	84
9	2	7	48	5	62
Total	34	27	157	52	270
Percentage	12.59%	10.00%	58.15%	19.26%	100%

A majority of images are related to only a part of the text, indicating that images are used to support the text and illustrate ideas or concepts. Figure 2 shows an example, in which the reading passage is dominant in meaning construction and the image of a smiling girl singing with confidence is used to illustrate the text (*19-year-old Asian pop star Candy Wang*). The image helps students visualize the character and her experience which may influence them emotionally and increase their own confidence in overcoming shyness.

The relation of text subordination to image occupies 19.26%, showing that a text

is connected with only part of the image. Figure 3 appears in the exercise section in which students are required to identify different colors and match each color with the word. Obviously letters in different colors play a decisive role without which the exercise cannot be done.

Figure 2. Image subordination to text (U4G9)



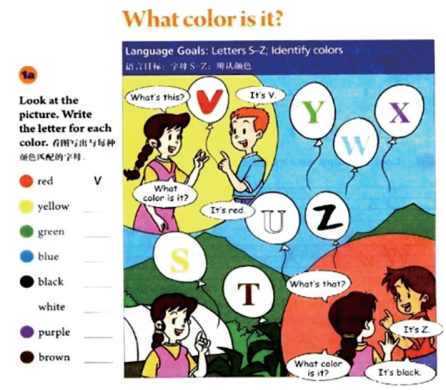
From Shy Girl to Pop Star

1 For this month's *Young World* magazine, I interviewed 19-year-old Asian pop star Candy Wang. Candy told me that she used to be really shy and took up singing to deal with her shyness. As she got better, she dared to sing in front of her class, and then for the whole school. Now she's not shy anymore and loves singing in front of crowds.

2 I asked Candy how life was different after she became famous. She explained that there are many good things, like being able to travel and meet new people all the time. "I didn't use to be popular in school, but now I get tons of attention everywhere I go." However, too much attention can also be a bad thing. "I always have to worry about how I appear to others, and I have to be very careful about what I say or do. And I don't have much private time anymore. Hanging out with friends is almost impossible for me now because there are always guards around me."

3 What does Candy have to say to all those young people who want to become famous? "Well," she begins slowly, "you have to be prepared to give up your normal life. You can never imagine how difficult the road to success is. Many times I thought about giving up, but I fought on. You really require a lot of talent and hard work to succeed. Only a very small number of people make it to the top."

Figure 3. Text subordination to image (U3G7)



Note: U =Unit, G =Grade; G7-9 are junior high school textbooks

In senior high school textbooks, statistics (Table 2) show a weak balance between equal (54.69%) and unequal (45.31%) image-text relations, with image-text independent relation occupying the largest percentage (45.31%), and image-text complementary relation the smallest percentage (9.38%), and a small variation is found in the number of images in different textbooks (Bao, 2020). The smaller number of image occurrences in senior high school textbooks can be explained partly by the fact that only the reading section in each unit is analyzed and partly by students' increasing verbal competence and decreasing reliance on images for decoding meaning.

Table 2. Image-text relative status relations in senior high school textbooks

Status	Equal		Unequal		Total
	Image & text independent	Image & text complementary	Image subordinate to text	Text subordinate to image	
Total	29	6	20	9	64
Percentage	45.31%	9.38%	31.25%	14.06%	100%

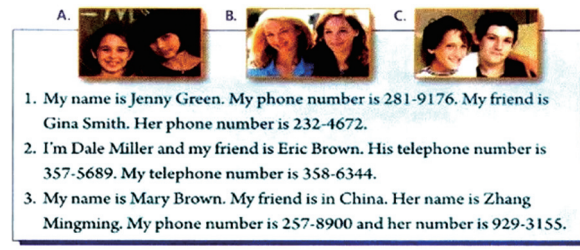
Different from junior high school textbooks, senior high school textbooks show that image-text independent relation (45.13%) ranks top in percentage followed by image-subordination to text relation (31.25%). On one hand, images are often used to support or reinforce the text in high school textbooks, and on the other hand, images and texts can work together to provide parallel information though in different

modes to generate meaning. Illustrations of maps are good examples of image-text independent relation. Figure 4 shows a map of United Kingdom in which the verbal text such as "UNITED KINGDOM", "England", "Oxford", "ATLANTIC OCEAN" and visual resources of colors and shapes co-contribute to an understanding of the geographical location and features of the Great Britain area.

In both junior and senior high school textbooks, the image-text complementary relation occurs the least often. Figure 5 shows an example of three images joining with the text in forming a meaningful exercise which requires students to match the paragraphs with the pictures. The visual information of male/female faces and Chinese/foreign facial features complements the verbal information that helps distinguish male names from female ones and Chinese names from foreign names. The two modes are interdependent and work hand-in-hand for the completion of the exercise.

Figure 4. Image-text independent relation (U4B2)

Figure 5. Image-text complementary relation (U1G7)



Note: B =Book; B1-B7 are senior high school textbooks

4.1.2 Image-text logico-semantics relations

In junior high school textbooks, as shown in Table 3, the expansion relation (71.48%) occurs more frequently than the projection relation (28.52%), and of 270 cases of image-text interplays extension (42.22%) ranks top in frequency, followed by elaboration (29.26%), locution (21.85%), and idea (6.67%). It is noted that enhancement is found to combine with elaboration or extension rather than appear alone, and therefore the number of its occurrences does not enter into statistics of total or percentage.

Table 3. Image-text logico-semantics relations in junior high school textbooks

L-S / Grade	Expansion			Projection		Total
	Elaboration	Extension	Enhancement	Locution	Idea	
7	45	41	8	32	6	124
8	21	42	13	15	6	84
9	13	31	7	12	6	62
Total	79	114	28	59	18	270
Percentage	29.26%	42.22%	/	21.85%	6.67%	100%

In senior high school textbooks (Table 4), the projection relation is not found; over one-third image-text interplays (35.94%) show combinations of elaboration, extension, enhancement or projection, and elaboration ranks top in frequency (48.44%).

Table 4. Image-text logico-semantics relations in senior high school textbooks

L-S	Expansion				Projection	Total
	Elaboration	Extension	Enhancement	Combinations		
Total	31	10	0	23	0	64
Percentage	48.44%	15.62%	/	35.94%	/	100%

In both junior and senior high school textbooks, elaboration and extension are rich in number. Elaboration refers to the relation where either the image or the text adds detailed but not always new information to the other. Figure 6 shows an example. In this exercise section, each image and the word express synonymous meaning to be matched, and the image adds detailed and vivid visual information to the text which helps students relate the word to their life experience

Extension refers to the relation where either the image or the text provides new or related information for the other. Figure 7 shows two images in a reading passage. The first image elaborates the text by adding detailed and visualized information of two sky lanterns made of bamboo and covered with paper rising into the air like small hot-air balloons. The second image of a paper-cut in the shape of the Chinese character 福 (good fortune) provides new related culture-loaded information for the text, hence giving an image-text extension relation.

Figure 6. Elaboration (U2G8)

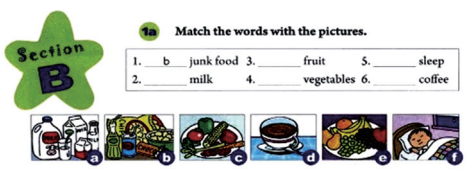
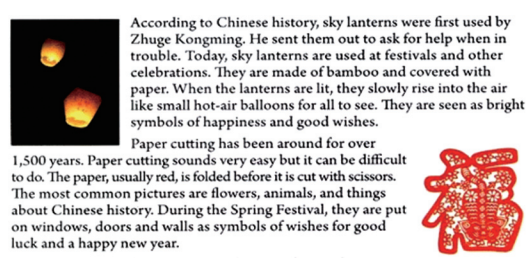


Figure 7. Extension (U5G9)


According to Chinese history, sky lanterns were first used by Zhuge Kongming. He sent them out to ask for help when in trouble. Today, sky lanterns are used at festivals and other celebrations. They are made of bamboo and covered with paper. When the lanterns are lit, they slowly rise into the air like small hot-air balloons for all to see. They are seen as bright symbols of happiness and good wishes.

Paper cutting has been around for over 1,500 years. Paper cutting sounds very easy but it can be difficult to do. The paper, usually red, is folded before it is cut with scissors. The most common pictures are flowers, animals, and things about Chinese history. During the Spring Festival, they are put on windows, doors and walls as symbols of wishes for good luck and a happy new year.

Enhancement occurs when the image or the text enhances the other circumstantially by time, place, reason or purpose. In both junior and senior high school textbooks, enhancement is found to combine with either elaboration or extension. Figure 8 shows a combination of elaboration and enhancement. On the one hand, the picture describes a man lying on the side of the road with a woman shouting for help and a bus driver coming to help, constructing the same meaning as written in the text (elaboration). On the other hand, the text gives the exact time (*9:00 a.m. yesterday*), place (*Zhonghua Road*) as well as reason (*the man had a heart problem*) of the accident, thus enhancing the picture. Figure 9 shows a combination of extension and enhancement in that the graph provides related

information for the text "There is no doubt that the earth is becoming warmer (see Graph 1)." (extension), and at the same time enhances the text by drawing a growth curve over a time span (1860-2000) (enhancement).

Figure 8. Combination of elaboration and enhancement (U1G8)

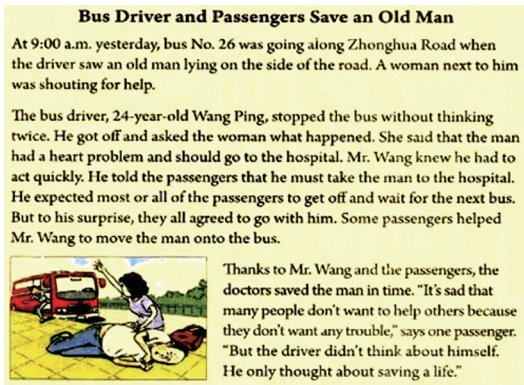
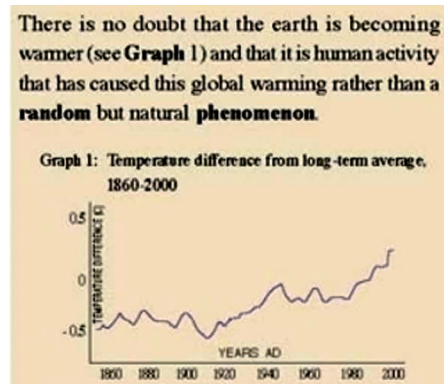


Figure 9. Combination of extension and enhancement (U4B6)



Projection occurs when an event expressed by one mode is reconstructed by the other. Locution (wording projection) tends to appear in speech bubbles, while idea (meaning projection) is always found in thought bubbles. They are often found in comic strips or combinations of texts and diagrams. This partly explains why projection is not found in the reading section of senior high school textbooks. In junior high school textbooks, cases of projection are found in the lead-in section. Figure 10 gives a good example of wording projection that appears in the speech bubble (*I study by making word cards*.) and of meaning projection that appear in three thought bubbles expressing separately the ideas of studying by making word cards, by listening to tapes and by asking the teacher for help. The speech bubble complements the left thought bubble and serves as a model for reading or understanding thought bubbles in this section.

Figure 10. Projection (U1G9)

The analysis of image-text relations in textbooks manifests meaning choices from two semiotic systems, where resources from each system interact to construct

various meanings and possibilities. In the process of reading multimodal texts and understanding image-text relations, Chinese EFL learners are trained to understand meaning conveyed by such non-verbal resources as pictures, graphs and colors in multimodal texts, understand the functions of verbal and visual modes in meaning making and understand how meaning is co-constructed by words, pictures, graphs and colors, and are therefore developing their visual literacy.

4.2 Construct of visual literacy specified in the new Curriculum

As pointed out by Ausburn and Ausburn (1978), today's children are visual children. They have grown up bombarded with visual messages, developed a visual vocabulary prior to a verbal one and learned to cope with complexity and subtlety of visual communication. Treichler (1967) confirmed that about 83% of the information is obtained via visual perception, 11% through hearing and 6% through other senses, and moreover, learners retain 10% of what they read, 20% of what they hear, 30% of what they see, and 50% of what they see and hear. The benefits of multimodal texts and the importance of "viewing" in the new media age are recognized and stated explicitly in the new Curriculum.

Viewing does not merely mean seeing with eyes, which is a sensory perception. Viewing is a cognitive ability to decode and encode visual elements, which is developed by nurture and can be trained and capable of development and improvement. As stipulated in the new Curriculum, viewing is one component of language competence, which is in turn one of the four core competences to be cultivated—language competence, cultural awareness, thinking capacity, and learning ability. Language competence, which is the foundation, involves the ability to understand and express meaning in the socio-cultural context with five language skills of listening, reading, speaking, writing and viewing. Viewing is defined as the skill to understand and use graphs, tables, animations, signs and videos in multimodal texts, which is visual literacy in nature. An understanding of multimodal texts entails not only the traditional reading literacy skills, but also the ability to read information in the graphs and tables and understand meaning in signs and animations.

The new Curriculum describes three levels of language competence to be developed in senior high school education. In addition, taking into account the individual needs as well as the regional difference in educational level in China, the new Curriculum offers three types of courses: obligatory courses, achieved by all high school graduates (graduation level); selective obligatory courses, taken by those sitting for the college entrance exam (*Gaokao* level); elective courses, selected by those with higher-level language needs (advanced level). By analyzing descriptions of level standards and course requirements, we work out a list of 11 sub-skills of visual literacy, or a bi-dimensional three-level construct of visual literacy (Table 5).

Table 5. A bi-dimensional three-level construct of visual literacy in the new Curriculum

Dimension	Level	Description
Understand (or read) visual texts	1	① can selectively record information needed in the process of listening, reading and viewing. ② can understand meaning conveyed by such non-verbal resources as pictures, images, signs, and colors in multimodal texts.
	2	③ can understand the functions of verbal and non-verbal resources (e.g. tables, pictures and signs) in meaning construction of multimodal texts. ④ can predict text content based on graphic information.
	3	⑤ can understand how meaning is co-constructed by words, sounds, pictures and images in films, television, picture books, songs, newspapers, magazines and other media. ⑥ can identify meaning conveyed by typography such as font and size.
Use (or write) visual texts	1	⑦ can use verbal and non-verbal means to describe personal experiences and object features. ⑧ can use icons, images, tables, and formats to transmit information and express meaning in the written production.
	2	⑨ can intentionally use titles, icons, tables, format, font and size to effectively transmit information and express meaning in the written production.
	3	⑩ can create texts in different modes. ⑪ can use such non-verbal resources as images and tables to creatively express meaning.

From the above descriptions, we can summarize expectations of visual literacy in secondary EFL education in China.

1. Visual literacy enables a learner to understand (or read) and use (or write) visual resources or visual texts for intentional communication.

2. Visual resources that are used for secondary EFL education include mainly images, tables, graphs, pictures, icons, signs, animations, colors, font, and size.

3. Visual modes are resources for meaning making in the socio-cultural context. Visual modes work with verbal modes (words) and aural modes (sounds) for understanding and expressing meaning.

4. For high school EFL learners, understanding (or reading) visual texts involves subskills of identifying and recalling key information in the visual text, of understanding meaning constructed in the visual text or co-constructed by combined verbal, aural and visual resources, and of predicting text content based on visual information.

5. For high school EFL learners, using (or writing) visual texts means to be able to intentionally use visual resources to describe personal experiences and object features, to transmit information and express meaning in the written production, and to create visual texts for meaningful communication.

6. In the process of cultivating visual literacy skills, learners' visual thinking capacity is developed hierarchically, ranging from identifying key information, understanding visual texts, through predicting text content, appraising functions of visual resources in meaning (co-)construction, to expressing meaning with visual resources and creating visual texts.

The new Curriculum describes viewing as a skill used separately or in combination with skills of reading and listening for understanding and producing meaning. This is in consistent with English education in other countries such as Australia, New Zealand and Singapore. It is also worth noting that in addition to skill, visual knowledge (e.g. knowledge of the way of presenting static or dynamic images and pictures) and visual strategies (e.g. strategies of obtaining visual resources and of distinguishing different functions of images and words in different texts) are also valued as important in English education in Australia and New Zealand (Yang & Wu, 2019).

4.3 Assessment of visual literacy in EFL tests

Zhongkao test papers measure language skills of reading and writing, language knowledge and use, and mainly consist of sections of reading comprehension, task-based reading, written production, grammar and vocabulary knowledge, and language use. *Gaokao* test papers consist of sections of listening, reading, language use and writing. Both are in the format of multiple-choice questions, blank-filling, cloze and writing. As can be seen in Tables 6-8, images occur less frequently in *Gaogao* test papers than in *Zhongkao* test papers, a majority of images occur in the reading part, and the most frequently used visual resource is the picture (drawings and photos included).

Table 6. Frequency of images in *Gaokao* test papers (N =12)

Test paper / Section			Jiangsu			National 1			National 2			National 3			Total
			2018	19	20	18	19	20	18	19	20	18	19	20	
Reading	Reading Comprehension	Table			1			1	1						3
		Picture	4	3											7
		Ad							1						1
	Task-based Reading	Table	1	1	1										3

As shown in Table 6, in 9 national *Gaokao* test papers, only two tables and one ad are found in the reading comprehension section. The rare use of visual resources in nation-wide tests may be explained by the education imbalance and the consideration of test fairness in China. Jiangsu province is strong in education and therefore uses more images in *Gaokao* test papers. However, an in-depth analysis of each image reveals that only visual reading but not visual writing is assessed, and in most cases images are subordinate to texts playing a supportive role for meaning construction. To be more specific, the seven pictures in Jiangsu test papers occur in two reading passages, one advertising Metropolitan Museum of Arts and the other introducing Buxton touring. The pictures offer different parallel information with the text, presenting an apparent image-text independent relation. However, the information

provided by the pictures is irrelevant to either meaning understanding of the text or the choice of options for item response. The pictures are rather used to activate test-takers' background knowledge about museums or city touring, raise their interest in reading and probably lower their test anxiety. The tables used in the reading section measure test-takers' Level 2 (or *Gaokao* level) visual reading described in the new Curriculum—*can understand the functions of verbal and non-verbal resources in meaning construction*. For example, the table of train timetable functions well by listing departure time, origin, destination and arrival time in four columns, making it easy for test-takers to scan the timetable and locate the specific information needed for the question. In this case the table plays a subordinate and supportive role in reading comprehension.

Compared with *Gaokao* test papers, *Zhongkao* test papers are found to use more images that are more widely distributed in sections of reading, grammar & vocabulary, cloze and written production (Table 7). On average, there is one image in each *Gaokao* test paper, and nine images in each *Zhongkao* test paper. Most images in *Zhongkao* test papers (87.18%) appear in the reading section, leaving a few dispersed in other sections. Visual resources include pictures, icons, tables, posters, notices, ads, and webpages, and over half of them (57.62%) are pictures (Table 8).

Table 7. Frequency of images in *Zhongkao* test papers (N =13)

Test section	Reading		Grammar & Vocabulary			Cloze	Written production	Total
	Reading Comprehension	Task-based Reading	Reading	Grammar & Vocabulary	Vocabulary			
Test format	Multiple-choice	Blank-filling	Blank-filling	Multiple-choice	Blank-filling	Blank-filling	Writing	
Number	79	17	6	4	2	5	4	117
Percentage	67.52%	14.53%	5.13%	3.42%	1.71%	4.27%	3.42%	100%

Table 8. Visual resources in *Zhongkao* test papers (N =13)

Resource	Picture	Icon	Table	Poster	Notice	Ad	Webpage	Total
Number	67	21	16	6	3	2	2	117
Percentage	57.26%	17.95%	13.68	5.13%	2.56%	1.71%	1.71%	100%

An in-depth analysis of each image and its relation with the text indicates that the following five subskills of visual literacy are measured, and the correspondence of each subskill with the one in the construct, i.e. the subskill specified in the new Curriculum, is noted in the parenthesis that follows.

(1) can understand meaning conveyed by such visual resources as pictures, icons, tables, posters, and webpages (subskill①②⑥)

Images are presented in different forms or genres such as tables, posters, notices, ads and webpages, and test-takers have to understand the format or layout of each genre before they can locate information needed. As shown in Figure 11-14, the

ability of reading and understanding the picture, the table, the poster and the webpage (font and size included) is directly measured. In the process of decoding these images, test-takers are expected to selectively record information needed.

(2) can understand the text aided by pictures or images (subskill③④)

Images are subordinate to the text, playing a supportive role for the text comprehension. With the aid of image(s), test-takers find it easier to understand the text. The image supports the text comprehension in different ways.

(2) a. The image helps predict the text theme or content. (Figure 15) (subskill④)

(2) b. The image helps infer word meaning. (Figure 16)

(2) c. The image helps activate background knowledge. (Figure 17)

(2) d. The image helps connect the unknown to the known. (Figure 18)

Figure 11. Understanding a picture

1. Look at the picture on the right. What is the woman probably saying?

A. Come in please. B. Turn round please. C. Stand up please. D. Hold on please.

Figure 12. Understanding a table of weather report

Here is the 6-day weather report in Suqian, Jiangsu Province.

Day	Description	Temperature	Humidity (湿度)
Monday June 29	Partly Sunny	27℃/ 19℃	91%
Tuesday June 30	A. M. Showers	28℃/ 21℃	85%
Wednesday July 1	Partly Cloudy	27℃/ 23℃	79%
Thursday July 2	Mostly Sunny	25℃/ 20℃	89%
Friday July 3	Mostly Cloudy	30℃/ 24℃	87%
Saturday July 4	P. M. Showers	31℃/ 22℃	92%

31. What's the weather like on Wednesday, July 1?

A. Mostly Sunny. B. Partly Cloudy. C. A. M. Showers. D. P. M. Showers.

32. The highest temperature appears on_____.

A. Tuesday B. Monday C. Saturday D. Thursday

33. The humidity on Friday, July 3 is_____.

A. 92% B. 85% C. 79% D. 87%

Figure 13. Understanding a poster

25. This is a poster of _____.

A. a sports meeting B. a charity walk

C. a football match D. a women's race

26. This event will start _____.

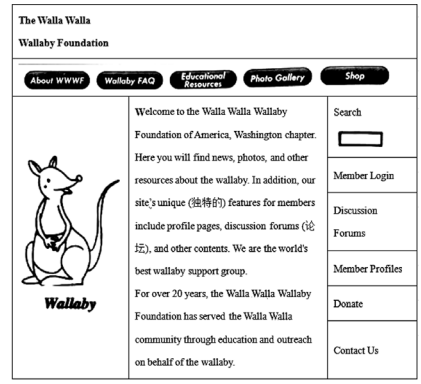
A. early in the morning B. late in the morning

C. early in the afternoon D. late in the afternoon

27. If you take part in this event with your parents, the least cost will be _____.

A. 105 dollars B. 120 dollars C. 165 dollars D. 360 dollars

Figure 14. Understanding a webpage



27. If you want to give money to support WWWF, you can click on the icon "_____".

A. About WWWF B. Donate

C. Discussion Forums D. Educational Resources

Figure 15. Predicting text content

Figure 16. Inferring word meaning

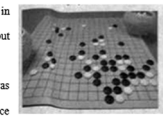
Figure 17. Activating background knowledge

Li Ziqi won Person of the Year in the category of cultural influence(年度文化传播人物)in 2019. She is famous for her videos in which she does the work of a farmer. She owns a great number of fans on social media platforms(社交媒体平台)all over the world.

Figure 18. Connecting unknown to known

The game of *Go* was one of the four greatest artistic types in Chinese culture. It is not only a competitive event of the mind, but also a board game of entertainment.

Created in China more than four thousand years ago. *Go* was introduced to Korea and Japan over 1, 000 years ago, and has since become a favorite activity of many people there. Today, *Go* still serves as a way of cultural exchange among the people in many Eastern and Western countries, as players in these countries take part in many international games every year.



Most pictures in the reading section are found to support the text in meaning construction. Figure 15 shows the use of a logo Hema Xiansheng, a popular fresh market widely spread in Jiangsu Province, to help predict the text content. In Figure 16, the word *kettle* pointed at the container helps infer the meaning of the new word. The picture shown in Figure 17 of a girl doing a farmer's work helps activate the reader's background knowledge of Li Ziqi and the series of videos she is famous for. The word *go* used as a noun is unknown to the test-takers, but the picture (Figure 18) helps them connect the unknown word to the known game of go and thus facilitates the text comprehension.

(3) can understand how the verbal text and the visual text interact to co-construct meaning (subskill⑤)

In addition to its supportive role, the image can complement the text in meaning co-construction. For example, the interplay of the verbal description by the text and the visual description by the picture helps co-construct the meaning of "Welcome!" in sign language, and the key to the question is dependent on a combined reading of images and the text (Figure 19).

Figure 19. Meaning co-constructed by verbal and visual texts

"Oh," she said. "Well, how do you say 'Welcome!' in sign language?"

"Like this." I swept my open right hand in toward my body, palm (掌心) up.

31. How is Welcome!' said in sign language?

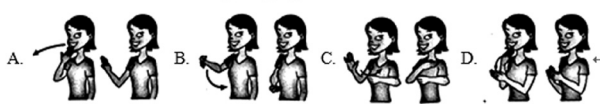


Figure 20. Transforming visual into verbal texts

(4) can understand the functions of verbal and visual resources in meaning construction (subskill③)

As shown in the above examples, verbal and visual resources are used to perform different functions, subordinate or complementary. The test-takers are expected to be able to distinguish the main from the minor, and decode meaning in the situated social-cultural context.

(5) can understand the visual text and transform it into a verbal text (subskill②)

In the written production section, test-takers are given pictures and required to decode the pictures and then encode meaning in a coherent verbal text. Figure 20 gives an example of writing a speech on the theme "striving for happiness" based on the given picture. The writing task involves an understanding (decoding) of activities described in the visual text and encoding of a coherent speech in the verbal text.

From the above analysis of visual literacy subskills measured in EFL test papers, we conclude that visual texts or images are used as input and the six subskills of understanding (reading) visual texts as described in the new Curriculum are directly or indirectly measured. However, the five subskills of using or producing (writing) visual texts are not measured. That is to say, only one dimension of visual literacy is measured, and the ability of producing (writing) visual texts is overlooked in EFL tests. The construct validity is thus low.

5. Discussion and Implications

(1) Visual-verbal interplay for meaning construction in EFL textbooks

"Competence with images is now a prerequisite of competence in life" (Lewis, 2001, p. 60). A visually literate individual needs to have a variety of skills to make meanings of visual texts. Understanding meaning co-constructed by verbal and visual modes is an essential part. Based on Martinec and Salway's (2005) framework of image-text relations, we analyze two sets of EFL textbooks that are widely used for Chinese high schoolers, and find a decreasing occurrence of images in textbooks with the increase of age or language proficiency level. That is, more cognitively, psychologically and affectively mature learners tend to depend less on iconic signs (visual, vocal, tactile, etc.) than on cultural signs (language). This finding is consistent with Sebeok's biosemiotic view and the *natural learning flow* principle (Danesi, 1998). Sebeok's biosemiotic approach suggests that children acquire knowledge and skill through the bodily experience first, then model this knowledge iconically, and finally adapt their own primary models of knowledge to the forms of representation that they learn in the cultural text. A child's learning thus flows naturally from iconicity to cultural symbolism, i.e. "from concrete, sensory modes of understanding to more conventional, abstract modes" (Danesi, 1998, p. 61).

We also find that a majority of images in textbooks are subordinate to texts, playing a supportive or illustrative role, and that image-text elaboration and extension relations are abundant in textbooks, revealing that a great number of images add either detailed or new information to the text. Other image-text relations, e.g. complementary and projection relations, rarely occur in textbooks. The visual mode interacts with other modes for meaning-making. Since bimodal texts manifest meaning choices from two semiotic systems, and each semiotic system has its own affordances, a visual text may create equivalent, complementary, contrasting or even contradictory meanings

to the verbal text (Callow, 2020). Some things are best done by using language, some things are best done by using images, and others are best done by combining language and image (the verbal and the visual) (Hassett & Schieble, 2007). As today's texts are increasingly hybrids of words and images, our instruction should consider the image-text relationship as epistemologically interconnected rather than divided between understanding the visual and the verbal. The inclusion of various image-text relations in textbooks can help EFL learners develop a richer understanding of the various functions of visual resources and how meaning is co-constructed by verbal and visual modes, and thus cultivate their visual literacy. For this purpose, we need a more even distribution of multiple image-text relations in textbooks, and textbook compilers are suggested to present visual input in different relations to the verbal text so as to enhance a wider range of ways of meaning construction.

In addition, new technologies provide nonlinear, interactive, dynamic, mobile and hypertextual features of communication. Textbooks should be featured by combinations of sound, print and images, different types of multimodal texts should be used as materials for literacy education, and students should be provided with varied means for developing their competences, including time to look and think deeply about visual and multimodal texts (Callow, 2008) and discuss how different modes interact with each other in meaning co-construction. Future concern should be the guided design of classroom activities that shape students' visual learning to negotiate multiple levels of meaning in visual and multimodal texts, and research on the effects of these activities or treatments on the development of students' visual reading, writing and thinking.

(2) Construct of visual literacy in the EFL curriculum

From China's national English curriculum, we generalize a bi-dimensional three-level construct of visual literacy and a list of 11 subskills of viewing. Specifically, high school EFL learners are expected to develop such subskills of visual reading as identifying key information in the visual text, understanding meaning constructed in the visual text or co-constructed with verbal and aural texts, and predicting text content based on visual information, and subskills of visual writing as intentionally using visual resources to describe personal experiences and object features, to express meaning in the written production, and creating visual texts for meaningful communication.

The construct of visual literacy or subskills of viewing can guide the selection and compilation of visual texts for the development of teaching materials and tests. However, it is far from an exhaustive list of visual literacy and a more comprehensive or complete list of visual literacy subskills or strategies is needed for better guidance of classroom teaching and assessment. In addition, viewing is defined as a skill in the new Curriculum, and we suggest an expansion of the viewing skill to embrace visual knowledge (e.g. knowledge of visual vocabulary and of visual conventions, Avgerinou, 2009), visual communication strategies (e.g. predicting content based on visual information) and visual thinking (e.g. turning verbal information into visual forms; logical thinking by means of images) so as to better prepare today's children to be competent learners in the digital age.

(3) Assessment of visual literacy in EFL tests

The addition of the skill of viewing into the new Curriculum is supposed to bring about visual literacy measurement in tests. Our study does provide evidence that a few subskills of visual literacy are measured in nation-wide EFL *Zhonggao* and *Gaokao* test papers. However, only one dimension of reading (understanding) visual texts is assessed, and the other dimension of writing (using) visual texts is not evidenced. Visually literate people should know not only how to decipher visual texts but also how to use or create visual texts. Since tests can drive teaching and learning and become powerful determiners of what happens in classrooms (Alderson & Wall, 1993), the ignoring of visual writing ability in large-scale tests will result in negative washback. Also the fact that not all the subskills of viewing specified in the new Curriculum are measured in national syllabus-based tests and not all the subskills of viewing assessed in the tests are specified in the Curriculum indicates on the one hand an urge for a complete list of visual literacy subskills and on the other hand low test validity.

To improve the test construct validity and promote positive washback effect, test designers are suggested to align visual literacy subskills to be measured in tests with those specified in the curriculum and instructed in classroom learning. For example, to teach and assess students' visual writing ability, instructors and assessors can make use of transition devices for students to translate what they have read in the verbal mode into the visual mode of drawings, graphs or tables, or guide students to multimodal compositions of posters, notices, twitters, etc. In the process of choosing, simplifying, abstracting, analyzing, composing, correcting, comparing and combining various visual resources in order to express meaning, children are developing their visual thinking ability as well.

Furthermore, tests are only one form of assessment. To better assess learners' visual literacy and promote their visual learning, summative assessment should be integrated with ongoing formative assessment that involves authentic learning experience in and out of classrooms. As proposed by Callow (2008), assessment should include student-made visual responses (drawing, painting, multimedia) into the texts viewed and discussed and involve students' using a metalanguage in order to support their development as viewers, makers and critics of visual and multimodal texts. Future teachers should develop their own visual competence as well so as to plan effective visual and multimodal teaching and assessment strategies in the classroom.

After all, man is a sign maker, and through the manipulation of signs s/he is changing the Umwelt or the zone between the internal world of the individual and the outside world (Uexküll, 1981), which in turn filters which signs the individual will attend to and which s/he will ignore. Perception is selective, and when the individual uses all sensory modalities to discern the characteristics of what has attracted attention and finds useful information in the world, s/he is actively constructing the surrounding world. This perception process can be mediated by school instruction. An ESL/EFL education within a semiotic framework would concern with how the learner manipulates signs in order to make new meanings. Meaning construction requires that sign-makers choose modes for the expression of what they have in mind, modes which they see as most apt and plausible in the given context, verbal, visual or auditory. The ability to use

different modes of signs affects both the ability to understand meaning created by others and the ability to create meaning by self. In the learning process of communicating with multiple modes of signs, learners are constructing Umwelt, changing the way they perceive the world and developing their mind. Learners should therefore be literate in all forms of meaning making, visually literate included. The inclusion of visual literacy in ESL/EFL education should involve a clearer definition and a complete list of visual literacy knowledge, skills and strategies in the curriculum, a balanced distribution of image-text relations in textbooks, an alignment of visual literacy teaching, learning and assessment and continual development of language teachers' visual competence.

References

Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14*(2), 115-129.

Ausburn, L. J., & Ausburn, F. B. (1978). Visual literacy: Background, theory and practice. *Programmed Learning & Educational Technology, 15*(4), 291-297.

Avgerinou, M. D. (2009). Re-viewing visual literacy in the "bain d' images" era. *TechTrends, 53*(2), 28-34.

Baker, L. (2001). How many words is a picture worth? Integrating visual literacy in language learning with photographs. *Canada Journal of Education, 26*(2), 201-217.

Bao, X. J. (2020). A multimodal study on senior high school English textbooks *New Senior English for China and Advance with English* (Unpublished master's thesis). Hunan Normal University, Changsha.

Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Communication, 25*(2), 166-195.

Braden, R. A., & Hortin, J. A. (1982). Identifying the theoretical foundations of visual literacy. *Journal of Visual/Verbal Languaging, 2*(2), 37-51.

Burkhardt, G. et al. (2003). *EnGauge 21st century skills: Literacy in the digital age*. Naperville & Los Angeles: the North Central Regional Educational Laboratory and the Metiri Group.

Callow, J. (2008). Show me: Principles for assessing students' visual literacy. *The Reading Teacher, 61*(8), 616-626.

Callow, J. (2020). Visual and verbal intersections in picture books: Multimodal assessment for middle years students. *Language and Education*, (1), 1-20.

Danesi, M. (1998). *The body in the sign: Thomas A. Sebeok and semiotics*. New York/Ottawa/Toronto: LEGAS.

Davies, A. (1968). *Language testing symposium: A psycholinguistic perspective*. Oxford: Oxford University Press.

Early, M., Kendrick, M., & Potts, D. (2015). Multimodality: Out from the margins of English language teaching. *TESOL Quarterly, 49*(3), 447-460.

Eco, U. (1979). *A theory of semiotics*. Bloomington: Indiana University Press.

Felten, P. (2008). Visual literacy. *Change: The Magazine of Higher Learning, 40*(6), 60-64.

Gu, Q. Y., & Zang, C. Y. (2011). Input modality effects on second language comprehension and incidental vocabulary acquisition. *Journal of PLA University of Foreign Languages, 34*(3), 55-59.

Hassett, D. D., & Curwood, J. (2009). Theories and practices of multimodal education: The instructional dynamics of picture books and primary classrooms. *The Reading Teacher, 64*(4), 270-282.

Hassett, D. D., & Schieble, M. B. (2007). Finding space and time for the visual in K-12

literacy instruction. *English Journal, 97*(1), 62-68.
Huang, Q. (2020). A comparative study of image-text relations in *Go for it* and *New Standard English* from the perspective of multimodality (Unpublished master's thesis). Jilin University, Changchun.
Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, (32), 241-267.
Kress, G. (1997). Visual and verbal modes of representation in electronically mediated communication: The potentials of new forms of text. In I. Snyder (Ed.), *Page to screen: Taking literacy into the electronic era* (pp. 53-79). Sydney: Allen & Unwin.
Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. London: Routledge.
Lewis, D. (2001). *Reading contemporary picturebooks: Picturing text*. London/New York: Routledge/ Falmer.
Lian, X. P., & Huang, Y. F. (2010). Effects of different input modes on incidental vocabulary acquisition. *Journal of Xi'an International Studies University, 18*(3), 110-113.
Liu, L. L. (2020). A comparative study of multimodal discourse analysis of Chinese and British EFL textbooks of junior high school (Unpublished master's thesis). Sichuan Normal University, Chengdu.
Martinec, R., & Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual Communication, 4*(3), 337-371.
Ministry of Education of the People's Republic of China. (2020). *National English curriculum standards for high school*. Beijing: People's Education Press.
Mirzoeff, N. (1998). *The visual culture reader*. London: Routledge.
O'Neil, K. E. (2011). Reading pictures: Developing visual literacy for greater comprehension. *The Reading Teacher, 65*(3), 214-223.
Stöckl, H. (2004). In between modes. In E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on multimodality* (pp. 9-30). Amsterdam: John Benjamins.
Treichler, T. G. (1967). Are you missing the boat in training aids? *Film Audio-Visual Communication, 48*, 14-16.
Uexküll, T. (1981). The sign theory of Jakob von Uexküll. In M. Krampen et al. (Eds.), *Classics of semiotics* (147-179). New York: Plenum Press.
Xiao, H. Z., Zhou, M. Z., & Liu, Y. T. (2018). Effects of tasks and settings on listening comprehension. *Modern Foreign Languages, 41*(3), 367-376.
Yang, L. N., & Wu, Z. M. (2019). Definition and Construct of viewing ability in the English subject from the perspective of core competences. *Foreign Language Teaching in Schools,* (10), 23-28.
Zhang, Z. (2010). A co-relational study of multimodal PPT presentation and students' learning achievements. *Foreign Languages in China, 7*(3), 54-58.
Zheng, Q., & Xu, Y. (2020). Effects of multimodal presentations on English vocabulary learning anxiety. *Journal of Xi'an International Studies University, 28*(2), 49-53.

About the author

Haizhen Wang (uuwanghz@sina.com) is Associate Professor of Linguistics at the School of Foreign Languages, Soochow University, Suzhou, China. She obtained her PhD in applied linguistics from Nanjing University, China. Her research interests include SLA, language testing and assessment, and English language education.