Iranian Journal of Applied Linguistics (IJAL), Vol. 16, No. 2, September 2013, 137-160

Analysis of Images in Iranian High School EFL Course Books

Mohammad Hassan Tahririan [a]
Professor of Applied Linguistics, Sheikhbahaee University, Isfahan, Iran

Elham Sadri
PhD Candidate in TEFL, Sheikhbahaee University, Isfahan, Iran

Received 23 February 2013; revised 26 July 2013; accepted 12 August 2013

Abstract

Although images are abundant and play significant aesthetic and pedagogical roles in today's EFL course books, they are still one of the less explored areas of research. The present study is an attempt to examine the role and function of images in Iranian high school EFL course books; it also aims to cast a critical eye over their suitability and efficiency. To this end, Kress and Van Leeuwen's (2006) model of visual grammar was adopted. The findings revealed that although the informative and/or illustrative functions of the images make them important resources in Iranian high school EFL course books, their full potential for language learning does not seem to be adequately exploited. This relates to flaws at the representational mode of meaning making, such as outdated portrayal of objects, gender stereotyping, and overdramatizing national identity as well as shortcomings at the interactive mode, such as poor modality due to grayscale printing and plain graphics with no contextualization or pictorial detail. These flaws make the images seem out of today's world and hardly credible for high school students who are already adept at new technology. Since these shortcomings may sacrifice pedagogical objectives, the study calls for changes in both technical quality and underlying messages of images in the course books.

[a] *Email address:* Tahririan@shbu.ac.ir
Corresponding address: Department of Foreign Languages, Sheikhbahaee University, Baharestan, Isfahan, Iran.

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

Keywords: Images; EFL course books; Representational, interactive, and compositional mode of meaning making

Introduction

The old adage 'a picture is worth a thousand words' is echoed in tens of EFL course books which have recently been introduced into the global market, including that of Iran. Most available EFL course books feature eye-catching images, pictures, cartoons and photos that beg to be explored and draw in even the most reluctant learners. Although it may seem a superficial way of judging a book, the reality is that students' experience of using a book and having confidence in it is much affected by how it looks, and whether it contains appealing images. The reader's eyes meet pictures before texts and it is no secret that young readers' first impression is usually shaped by the book's appearance; hence, colorful pictures and alluring illustrations give rise to positive impressions and attitudes. But pictures can do even more than hook learners into the book. As the adage implies, pictures convey information more efficiently and effectively than words do. They can quickly and easily convey much detailed information in visible form; therefore, they contribute a lot to readability and comprehensibility of a book.

Pictures and visual representations have a long history in instructional materials, and their positive value in second language acquisition and EFL has been documented in the literature. It is understood that using visuals in teaching results in a greater degree of learning (Omaggio, 1979). Pictures in EFL course books trigger students' interest and help them engage in language learning in more creative and critical ways (Canning, 1997, 1998; Curtis & Bailey, 2001). Also, it is acknowledged that pictures are universal stimuli that provide a starting point for language sharing in EFL classroom; they provide students with tools to self-regulate their language learning processes and help teachers make their instruction more explicit and catchy (Canning-Wilson, 1999). A body of research has been conducted on the role of various types of visuals, for example, pictures, images, photos, animations, and video clips as *situational contexts* in conceptual processing and language comprehension (Barsalou, 2003; Bransford & Johnson, 1973; Di Carlo, 1994; Hegarty, 2004, to name but a few). Particularly, some studies have focused on the beneficial effect of images and visuals in providing contextual clues for reading comprehension (Omaggio, 1979), writing process and composition tasks (Bailey, 1995), and speaking and listening tasks (Mueller, 1980; Rong, 2002). Another advantage of including pictures in EFL course books is that they provide

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

excellent testing prompts; pictures lend themselves to be developed into tests to manipulate language structures, vocabulary, functions, situations, and skills to determine what the learner has acquired (Canning-Wilson, 1999).

Aside from adding flair and appeal to course books, and regardless of their facilitative and instructional role in language learning, pictures take on a crucial yet less noticed function in EFL course books. They work as a powerful factor in learners' construction of knowledge about the real world; they can immerse learners into a new world that cannot otherwise be created in the classroom environment (Canning-Wilson, 1999). Pictures bring outside world into the classroom thus making situations appear more authentic and helping students experience perspectives other than their own. More importantly, characters, places and situations portrayed by images in foreign language learning course books introduce learners to what the world may look like through the eyes of its speakers, and provide them with opportunities to develop their understanding of the target culture, mindset, and worldview (Mendoza & Reese, 2001). Consequently, the ways images in EFL course books portray people in the target society and depict their lifestyles, physical features, communicative manners, and cultural behaviors directly affect learners' judgment and foster specific impressions.

For all these reasons, it must be viewed as an essential necessity to analyze and explain the function and characteristics of images in different EFL course books. Surprisingly, however, this significant challenge has not so far received much attention. As for studies conducted by Iranian scholars, the existing investigations have examined many of current EFL course books used in school curricula in Iran from different aspects and perspectives (e.g. Ansary & Babaii, 2002; Azizfar, Koosha, & Lotfi, 2010; Yarmohammadi, 2002); nevertheless, no study has focused on the images in these course books while they are definitely worth of a thorough investigation. The present study takes this gap as its point of departure and, drawing on the framework proposed by Kress and Van Leeuwen (2006), attempts to investigate the images in Iranian high school EFL course books in order to shed light on their role, function, and status, as well as their suitability and appropriateness. The following research questions are central to this inquiry:

- What is the status and significance of images in Iranian high school EFL course books?

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

- What are the interpretations and implications of the images in Iranian high school EFL course books?
- To what extent do the images in Iranian high school EFL course books suit foreign language learning objectives?

Theoretical framework

The theoretical framework guiding the study was Kress and Van Leeuwen's (2006) 'grammar of visual design'. Kress and Van Leeuwen (henceforth KvL) place themselves at variance with Barthes' (1977) widely-accepted view that meanings of images are related to and mostly dependent on language. They argue for autonomy of images from verbal texts and postulate that images communicate meanings "through their own specific forms and independently of words" (KvL, 2006, p. 17). KvL's theory has its roots in Halliday's (1978, 1994) Social Semantics and Systemic Functional Linguistics. By emphasizing that meaning can be projected via semiotic resources other than language (e.g. images, architecture, music, mathematical symbolism, gesture, clothing, etc.), Halliday (1978) exemplifies the start of a semiotic approach which broadened the narrow focus of written language in linguistics. He argues against the traditional separation between language and society and postulates that the study of signs and messages must be situated within the context of social relations and processes. Along similar lines, KvL propose a grammar of visual design to present a socially-based theory of visual representation. The use of the term 'grammar' implies that they attempt to examine ways in which what is depicted in images is developed into a coherent, meaningful whole; in much the same way that grammarians examine how words are combined into clauses, sentences, and whole texts. In the same way that the grammar of language helps readers make sense of complex sentences in texts, the grammar of visual design has a critical role in reading the images because it reveals implicit regularities and invisible patterns that govern the visual representations; and hence, helps viewers interpret, negotiate, and make meaning from information presented in the form of an image. KvL (2006) assume that visual mode draws upon the same semantic system as does language; therefore, the grammar of visual mode is determined by the same metafunctions that organize the language (i.e., ideational, interpersonal, and textual). With differences in nomenclature, KvL (1996, 2006) adopt the three Halliday's (1978) metafunctions to create a grammar of visual design assigning representational, interpersonal and compositional meanings to images. Figure 1 portrays the correspondence between Halliday's metafunctions and KvL's modes of visual meaning.

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

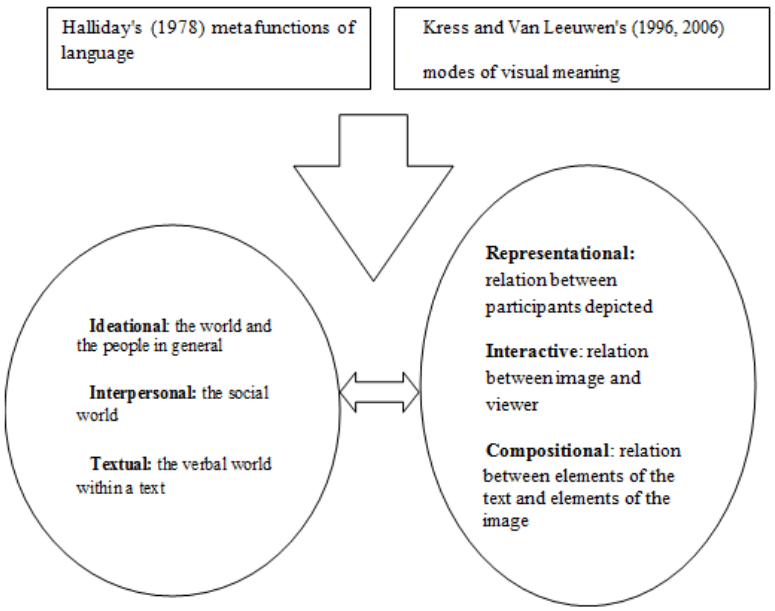


Figure 1: Correspondence between Halliday's (1978) metafunctions and KvL's (1996, 2006) modes of visual meaning

The representational mode of visual analysis involves looking at the participants, whether animate or inanimate, the activities or attributes represented by the participants, and circumstances surrounding the participants or embeddings in which the representations are developed. Interactive meaning is constructed between the represented and the viewer and is realized in the ways that the images attempt to address their potential viewers in interactional terms. The compositional mode of visual analysis requires the study of aspects of page layout which determine the extent to which the visual and verbal elements achieve a sense of coherence in the whole text. Taken together, KvL's (1996, 2006)grammar of visual design is the first systematic and comprehensive account of visual structures which examines the ways in which images simultaneously communicate meanings regardless of the text to which they are attached. This suggests a provisional framework for describing and evaluating the images in media, art, and printed materials.

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

Method

Materials

The materials used in the study were three high school EFL course books currently used in Grades 1 to 3 in Iranian secondary schools. The course books have been compiled by the Course Book, Curriculum Development and Planning Department of the Ministry of Education of the Islamic Republic of Iran. They were developed in the 1980s and revised in 1990s.

The study encompassed all the images in the three EFL course books; that is a total of 296 images. The information as to the number of images, number of pages and the distribution of images in each course book appears in Table1.

Table1

Frequency and distribution of images in Iranian high school EFL course books

EFL course book	Number of images	Number of pages
Book I	116	130
Book II	88	106
Book III	92	102

Data analysis

The analysis of images was performed based on KvL's (2006) model of visual grammar. KvL's model covers an enormous range of aspects and types of visuals from children's drawings to book illustrations, photo-journalism to fine art, as well as three-dimensional forms such as sculpture and toys. Therefore, KvL's visual grammar is not a one-size-fits-all toolkit delimitated in ready-made categories that could equally be applied to all visual representations. Consequently, it was necessary to scrutinize the broad analytical features proposed in KvL's model and adopt those features that seemed to bear more relevance to the subject of the analysis (i.e., images in EFL course books). What follows is a detailed account of the adopted categories of KvL's (2006) framework which guided the data analysis procedure. These features are summarized in Table 2 as well.

Representational mode: according to KvL (2006), the representational mode relates to the ways in which visual structures represent people, animals, and places; their actions, experiences, and relationships; as well as the characteristics and qualities attributed to them. Thus, to analyze this mode of meaning making in the images under study, in the first place the type (in terms of being human, non-

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

human, male, female, child, and adult) and frequency of presence of the participants in the images were figured out. Then the relationships and characteristics of the participants were investigated to reveal the covert meanings (e.g., sociocultural connotations, gender stereotypes) represented by the images.

Interactive mode: According to KvL (2006), the interactive mode of visual meaning refers to ways in which the visuals address, communicate, and connects with the viewers. The image-viewer interactions in the images under study were analyzed based on these three aspects: distance, perspective, and modality.

Distance (or social distance) relates to the size of the visual frame. KvL (2006) suggest that the amount of participant's body shown in the visual frame determines the degree of social distance and intimacy between the participants and the viewers. More specifically, a close shot which shows the head and shoulders of the participants implies an intimate relationship, a medium shot which cuts off participants approximately at waist conveys a familiar or social connection, and a long shot frame which shows full human figure demonstrates that the social relation between the visual and the viewer is a public, largely impersonal one.

Perspective specifies the degrees of involvement and power. This is realized through various choices of visual horizontal and vertical angles. Put differently, the perspective or point of view from which viewers and participants interact denotes different levels of involvement, abstraction, power, or inferiority. More specifically, KvL (2006) propose that along the horizontal axis, a frontal plane associates with an attitude of involvement whereby the viewer is invited to become part of the world depicted in the image. On the other hand, the use of an oblique angle conveys a sense of detachment and connotes that what the viewer sees is not part of his or her world. The vertical axis and its variants (e.g. high, low, or eye-level) signify power at different levels and in different relationships (KvL, 2006). For example, whenever a represented participant is pictured from a high angle, he or she is seen from the perspective of inferiority to the viewer. Conversely, if the represented participant is shown from a low angle, he or she is said to have power over the viewer. Finally, if the image is at eye-level, it encodes a relation of equality between the viewer and the participants.

Modality refers to ways of modulating reality in visual representations. KvL (2006) define modality as the criteria against which viewers express judgments

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

about the truth or credibility of visual structures. According to KvL (2006), two critical factors in credibility or naturalistic modality of visuals are color and contextualization. They suggest that the use of color is of great importance in expressing visual modality in that the more the color is abstracted from naturalistic presentations, the lower is the modality. They contend that in order for visuals to have high modality, they should have: high color saturation rather than black and white; diversified colors rather than monochrome; and modulated colors rather than unmodulated flat colors. Naturalistic modality is influenced by the contextualization of the image as well; that is, by its background and the extent to which the settings are represented in detailed and itemized manners. Generally speaking, KvL (2006) maintain that the presence of background in an image increases its modality from a naturalistic perspective whereas the absence of background lowers it.

Compositional mode: Compositional meanings have to do with the principles of layout, or the ways visuals are composed or structured, and the kinds of meanings conveyed by the various possible compositions. These include single mode forms, as in the composition of just a single visual, and multiple mode forms, as in the composition of a text comprising one or more visuals and accompanying verbal text (such as EFL course books in the present research). Based on this, the compositional mode was analyzed in terms of two features: (a) text-image status and (b) information value.

Text-image status refers to relative relationship between the text and its co-occurring image. KvL (2006) suggest that in texts on humanistic subjects such as English course books, images may have the following relations with their co-occurring texts:

- Information: the written text may become less important, with the message articulated primarily in the visual mode
- Illustration: the visual representation is an account or evidence of the truth of the claims made in the linguistic part
- Decoration: the visual and linguistic components have no immediate semantic relationship and the visual mode is redundant

Another aspect of the layout, defined by KvL (2006) as information value, has a major role in conveying compositional meanings. The information value involves considerations about the placement of image elements or pictorial zones (left/right;

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

top/bottom). KvL (2006) suggest that if verbal-visual compositions follow a left-right structure, the left hand element conveys the meaning of a familiar and self-evident piece of information (or the 'Given') and the right hand element of the layout demonstrates a piece of new information which should be attended to (or the 'New'). On the other hand, the information value of atop/bottom positioning carries different features in that the upper section of a verbal-visual composition is believed to contain what is presented as 'Ideal' (idealized or generalized essence of information) whereas the lower section what is 'Real' (more practically oriented information).

Table 2

Modes and realizations of visual meanings based on KvL's (2006) model

Mode	category		sub-category
Representational	Participants	human	Age
			Gender
			sociocultural portrayal
		non-human	
Interactive	Distance	close up	
		medium shot	
		long shot	
	Perspective	horizontal	Frontal
			Oblique
		vertical	High
			Low
			eye-level
	Modality	Color	
		Contextualization	
Compositional	Information Value	left/ right	
		Top/bottom	

Results

Representational Mode

The first step in analyzing this mode of meaning making was determining the type of the represented participants and the frequency of their presence. The results indicate that of the 296 images found in the three books under study, 190 images have human participants as their major theme, and 106 images center on representation of non-human participants (53 objects, 17 animals, and 35 settings and locations).The finding that the human participants are more frequent than the non-humans shows that the images mirror a realistic picture of the audiences' lives

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

where humans and their actions and social relations have a central place. Table 3 summarizes the information about the distribution of participants by type and frequency of presence.

Table 3

Type and frequency of presence of the participants in the images

Participant			*N*	%
human (64 %)	Child		92	31
	Adult (33 %)	female	27	28
		male	71	72
non-human (36%)	Object		53	18
	Animal		17	5.7
	Locations		36	12.2

It was observed that non-human participants include objects (18 %), animals (5.7 %), and locations (12.2 %) which depict typical animals (e.g., cats, dogs, horses, etc.) and common ordinary objects (e.g., pen, book, car, etc.) and familiar common places (e.g., classroom, library, home, park, etc.) in naturalistic form. Nevertheless, the choice of objects and locations seem somewhat outstripped and outdated by the Iranian lifestyle at the present time. The represented objects (e.g. clocks, chairs and tables, stationeries, kitchen, home appliances, and many more) are portrayed in ways that could hardly associate with the current life-ways of the young audience of EFL course books. None of the images portray the signs of recent technological advancements in current Iranian society, for example, computers, cell phones, electric home appliances, and the likes. The outdated representation in the EFL course books under study is most obvious in the case of 'vehicles'. Vehicles are the most frequently observed category of object participants in the books, since 31 out of 53 images (57 %) center on vehicles as their major subject. The vehicles are represented in different forms; that is: car (16), bus (4), bicycle (4), train (2), taxi (1), airplane (1), ship (1), boat (1), and snowplow (1). However, these portrayals are irrelevant to the realities of present Iranian lifestyle. Another conspicuous instance of representing objects in outdated ways is the case of 'television sets'. Television sets are depicted in 10 images and watching television seems to be the most favorite activity of families. Out of 24 images which depict family members together (i.e., mother, father, siblings, etc), in 6 images (25 %) family members are depicted as watching television. This significant role assigned to television is indeed true of the majority of Iranian

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

families nowadays; nevertheless, the kind of antiquated television set presented in those images is uncommon and out of the ordinary. What is more, out of 10 images representing television sets, in 8 images the program being watched is not depicted (or clear); interestingly however, in the 2 images which the program is discernable, what is being watched is a puppet show (*madrese-ye-mooshhaa*) which used to be aired on national television many years ago. Figure 2 displays examples of such behind the time representation of television.

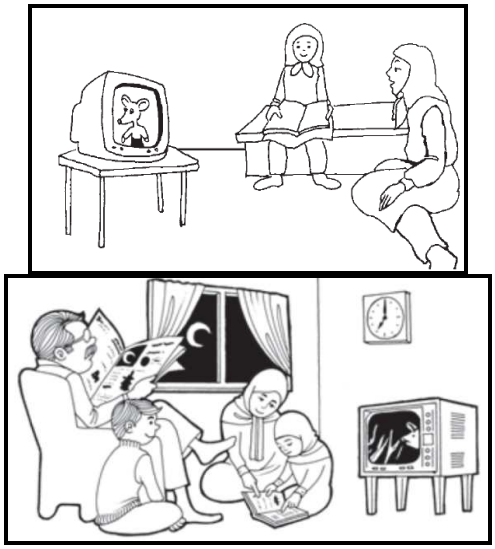


Figure 2: Outdated portrayal of ways of living: the case of television sets and programs (left: Book II, p.37; right: Book I, p.5)

It was found that of the images representing humans, the subject matter of 92 images (31 %) is mostly children and the subject of 98 images (33 %) is adults. In terms of adults' gender distribution, 71 images (72 %) represent males and 27 images (28 %) represent females. Aside from the number of presence of males which is significantly more than females (around 2.5 times), other patterns of biased and stereotyped gender depiction were also observed. One aspect of such unequal and stereotyped representations is the social and domestic roles ascribed to males and females. The results of this part of the analysis are presented in Table

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

4.In terms of social roles, males are represented as involving in a broad array of different occupations whereas females are depicted as being engaged in a very limited number of occupations that are conventionally deemed as more feminine. In addition to job possibilities, gender stereotyping is also detected in the depiction of genders in social actions and public places. For instance, of 5 images which picture the setting of 'restaurant' none includes females; as if eating out is a social activity specified to men. Also, of the 5 images dealing with the concept of 'travelling', none represents females, again giving the impression that travelling alone by train, airplane, or bus is only men's job!

Furthermore, it was found that when the two genders are represented in domestic settings, females are often portrayed as being involved in doing household chores while males are often sitting, watching television, or reading. Interestingly, of the three images which depict men as involving in 'washing' activity, two are about washing cars (which in the meantime is a sex-based type of activity) and the other image illustrates a man looking daggers at a pile of dishes prompting this response: 'Mr. White doesn't dislike eating cakes; he dislikes washing the dishes!' (Book III, p. 30)

Table 4
Domestic and social gender depiction in the images of Iranian EFL high school course books

Gender	Occupation (Social role)	Indoor activity (Domestic role)
Female	teacher 3 doctor 1 nurse 1 cashier 1 librarian 2	working in the kitchen 3 washing the dishes 2 washing/ spreading the clothes 2 buying/ carrying home groceries3 watering flowers 1 watching television 1 reading 1 eating 1 helping/talking with children 4

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

Male	shop-keeper 3 teacher 2 farmer 5 worker 5 police 2 driver 4 doctor 1 boss 2 athlete 3	clerk 3 scientist 2 tourist 2 thief 2 waiter 3 hunter 2 animal trainer 6 shoe maker 1 mechanic 1 carpenter 1	reading newspaper/books 5 being a host 2 being sick 2 washing a car 2 washing the dishes 1 setting out to work 2 saying prayers 1 making toys out of wood 1 watching television 2 doing sports 4 playing with children 1

In terms of appearance, all of the human participants (100 %) are portrayed in accordance with Iranian cultural and religious norms; that is they are represented in full-covering attire (females in hijab) in all the situations either inside or outside their houses. The participants' appearance does not convey any messages including appreciation of a particular occupation or social class over others, or attitudes in favor of certain political or ethnical affiliations. In other words, the appearance of the participants voices and appreciates the cultural identity and religious norms of Iranian society at the national level. However, in at least two instances this portrayal seems to blow out of proportion. The first instance of exaggeration in giving prominence to Iranian social identity is the cases of participants who are depicted with Iranian appearance and clothing but addressed with Western names. In fact, a total of 60 human participants are assigned names including 32 Iranian names (e.g., Ali, Maryam, Mr. Taban, etc.) and 28 Western names (e.g., Bob, Kim, etc.), but in all these cases (100 %) the participants are represented as Iranians. Figure 3 provides an example of such cases.



Figure 3: Incongruous representation of participants: Iranian characters with Western names (Book III, p. 17)

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

Other instances of going to extremes in portraying Iranian identity are the two cases in book II (p. 35 and p. 95) where this question is asked: 'which man is from Iran/Iranian?' the picture cue for answering this question demonstrates Iranian men with full beard and men from Eastern Asia with slanted and small eyes. It gives the impression that any Iranian man has mustache and beard in the same way that any Chinese guy has slanted eyes!

Interactive Mode

Distance: the degree of social distance and intimacy between the participants and the viewers was examined based on the amount of body shown in the images or close up, medium shot and long shots. From among 190 images representing human participants, 132 images (70 %) were in the form of long shots and 58 images (30 %) were in the form of medium shots. Close up shots were not observed in the images. According to KvL (2006), close up frames demonstrate personal, intimate relations between the images and the audience while long shot frames indicate impersonal, and medium shot frames imply social relations. Thus, it may be inferred that the absence of close up shots in books under investigation indicate a social distance between the participants and the viewers. However, this implication should be taken with caution considering the insights from other studies (e.g., inference Lewis, 2001; O'Toole, 1994; Royce, 2007; Van Leeuwen & Jewitt, 2001) which establish that close-ups may take on a different meaning in educational resources and especially pedagogic books. It has been argued that close-up shots are infrequent in picture-based text books (unless they appear on the front covers) and long and medium shots are prominent in such books as a way of revealing characters. Furthermore, O'Toole's (1994) work on the language of visuals confirms that in many cases long and medium frames more than indicating detachment or engagement, suggest that the object is there for the viewer's contemplation and deliberation. Thus, it seems logical to conclude that the frame size in the images under study render impersonal and objective relationship with the audience in order to foster their reflection and thought. Figure 4 provides typical examples of frame size in the images under investigation.

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

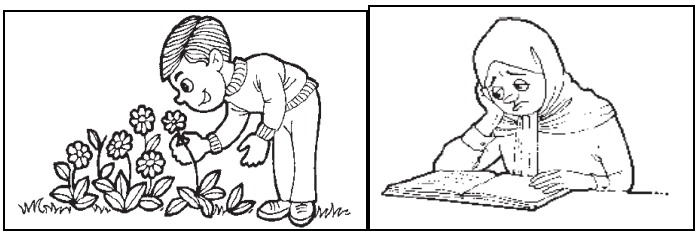


Figure 4: Typical examples of frame size in the images (Left: example of long-shot frames, Book I, p.22; Right: example of medium-shot frames, Book I, p. 59)

Perspective: the results of this stage of analysis revealed that the participants in the images are seen neither from above nor below but from a frontal eye-level perspective. In other words, all the images (100 %) are represented in frontal perspective (along horizontal axis) and eye-level angle (along vertical axis).According to KvL (2006), frontal plane denotes involvement and inclusion of the viewer in the world of the image. Also, eye-level perspective communicates power equality between the images and the audience. The fact that there is no power difference established between the represented participants and the young audience implies that the students would identify and make connections with the images. Figure 5exemplifies the frontal eye-level perspective in the images under study.

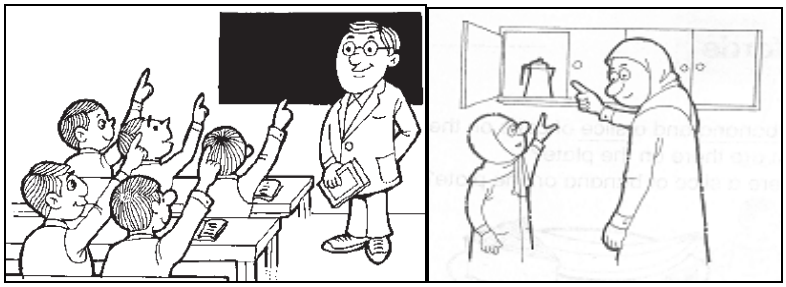


Figure 5: Typical examples of frontal eye-level perspective in the images (Left:, Book I, p.22; Right: Book II, p. 2)

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

Modality

To examine the modality or the degree of credibility of images to the audiences of the EFL course books, two categories of color and contextualization were taken into account. Regarding color, even a quick leaf through the books under attention shows that all the images in the three course books (100 %) are monochrome and plain. In books I and II the images are printed in grayscale (i.e., black and white) and in book III a slight tinge of blue color is added to the print. According to KvL (2006) the absence of color might serve to convey certain meanings in abstract art, photography, cinematography, and so forth, but in the naturalistic code which aims to represent real-life events and actions – as is the concern of the images under study— the absence of color lowers the degree of naturalness of the representations; and hence, adversely affects the quality of being trusted and believed in. KvL (2006, p.160) distinguish a continuum of color modality for naturalistic coding orientation which runs from "full color saturation to the absence of color; that is black and white". Accordingly, the color modality of the images in the EFL course books falls within the lowest modality range.

Another analyzed sub-category of modality is contextualization (i.e., the degree to which a scene is presented in details).KvL's (2006) define a continuum for contextualization which moves from full contextualization (maximum modality) to ellipsis (i.e., reducing the contextual elements to the most minimal) and ends in full decontextualization (lowest modality). In the books under study, no image enjoys full contextual representation; 154 of the images (52 %) fall within the range of ellipsis and 142 images (48 %) represent no background and are blank in terms of contextualization. The images with blank contextualization represent participants in blank backgrounds while no element is added to indicate the setting or the location in which the depicted actions take place. Figure 6 provides some typical examples of blank contextualization in the images under study.

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

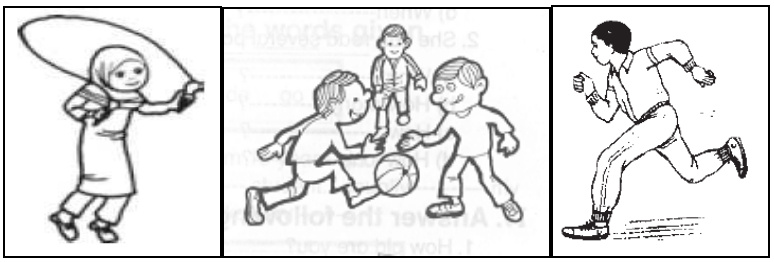


Figure 6: Typical examples of blank contextualization in the images (Left:, Book I, p.8; Middle: Book I, p.3; Right: Book II, p. 20)

In the case of ellipsis, the participants, actions, and settings are displayed in most trivial and minimal details. For example, as Figure 7 illustrates, conceptualization of 'cage' is realized by means of just a few vertical lines with no other elements to provide more depth and detail to the scene. Overall, it could be concluded that the investigated images suffer from poor contextual modality because the minimal and simplistic settings depicted in them give the impression that the scene is taking place in void and the participants are abstracted away from reality.



Figure 7: Examples of minimalist contextualization in the images; the case of 'cage'(Book II, pp. 19, 36, 88, respectively from left)

Compositional Mode

Text-image status: the status of the images was examined based on the relative relation of each image and its corresponding text. As is evident from Table 5, the majority of the analyzed images (99 % in book III, 41.4 % in book II, and 59.6 in book I) were of informative type, that is the type of image that carried most of the information load of the text/image pair. The reason for the significant frequency of

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

informative type of image is that a great number of exercises in EFL course books are picture-based questions, thus the image carries the required information to prompt an answer. Images of illustrative type were also observed (38.8 % and 21.3 % in books I and II, respectively). This type of image accompanies the text to put more emphasis on or provide evidence for the claims made in the linguistic part. Like informative images, illustrative images accompany language exercises but they do not serve as response cues but play a complementary role and help the students relate to and engage in doing the exercises more effectively. The least frequent type of images are decorative ones (19.8, 19.1, and 1 % in books I,II, and III, respectively). Contrary to the previous two types, decorative images do not provide response cues for the exercises and do not carry significant information; they just serve to make the layout more attractive. On the whole, in terms of text-image status, the results suggest that images are of prominent salience in the books to provide information or explanation for the texts to which they are attached.

Table 5
Frequency and distribution of text-image status in the images of the course books

EFL Course book	Information		Illustration		Decoration	
	N	%	N	%	N	%
Book I	48	41.4	45	38.8	23	19.8
Book II	53	59.6	19	21.3	17	19.1
Book III	91	99	0	-	1	1

Information value: the information value of images was analyzed in terms of left/right and top/bottom structuring. The positioning seems to be dependent on the image size: smaller images would fit in the small place in the right hand of the page, but larger images which need more space are placed below the texts. A total of 157 images (53 %) follow left/right structuring where the text is placed in the left side of the page and the image appears in the right side. In KvL's (2006) terms, in these cases the texts are the Given and the images are the New. This sounds convincing considering the salience of the images in terms of providing necessary information for students to comprehend the questions and do the exercises. From among all the images, 121 (41 %) follow a top/button pattern in which the texts are at the top (Ideal) and the images appear below the texts (Real). Unsworth (2008) notes that the Ideal/Real distinction in *textbooks* conveys specific meanings; the top part deals with the more generalized, abstract, conceptual information, while the

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

bottom part deals with the specific, concrete, observable information. The images are thus justifiably positioned below the texts to provide the observable clue for students to comprehend the attached texts. In a small number of images (18 images; 6 %) the top-bottom structure is the reverse; the image appears at the top and the text is in the lower position. This applies to the case of 'reading passages' of the lessons where the image is placed in the top half of the leader page serving as a snapshot which captures the salient meanings of the passage.

Discussion

Having reported the results of the study as to the different meanings conveyed by the images in representational, interactive, and compositional aspects, we now put the findings together to address whether the images live up to their intended pedagogical goals. Taking such a holistic picture requires weighing the merits and the disadvantages identified in the images under study. To begin with the merits, as the results at the compositional mode indicate, in terms of text-image status, the images demonstrate remarkable significance in text-image pairs in that they mostly serve informative or illustrative functions. Put another way, the images are not redundant and ornamental but are of rich functional value. Other compositional features-- left/right and top/button structuring-- appear to be in line with this functional salience. The images are placed either in right or New (rendering new information which is crucial to the point of the message) or bottom (more specific and practically orientated information) positions.

The salient functional and informational values of the images are further supported by adopting a factual stance at the representational level. The fact that the images cover a broad array of human and non-human participants, ordinary actions, and every-day events represent realistic pictures of the world that is coherent with the students' real life. At present, non-realistic visual arts are becoming ubiquitous in educational materials (Chen, 2010; Martin, 1985) and fictitious creatures such as talking animals and personified objects endowed with sense and sensibility have gained popularity in books, visual aids, and many educational resources. Although playful cartoons have the merit of provoking students' motivation and interest in learning the foreign language, it is argued that the pervasive use of fantasy in educational settings runs the risk of promoting a "childist" ideology (Martin, 1985; p. 55) which underestimates students' ability to learn the language from texts that deal seriously with a range of real contexts.

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

Thus, adopting a factual and realistic style in the course books under scrutiny seems to be the best choice for high school adolescents.

The number of child and adult participants in the images falls within the same range (31 % and 33 %, respectively) which again has the advantage of making connections between the represented participants and the students. The represented children are grouped together as classmates, friends, and family members and are depicted as being involved in the types of activities that adolescents at the same age normally do (e.g., studying, doing sports, watching television, etc). Hence, the young audience of the EFL course books can identify with the represented participants as their peer group who share the same type of interests and habits. The results of the analysis of the interactive mode provide additional evidence for this conclusion. Although the medium and long shot frames in the images imply some social distance which is due to objectivity and fostering contemplation, the use of frontal angle perspective indicate that the represented participants and the viewers are approximate and coherent. In this respect, the choice and representation of participants is indeed appreciable; such depictions foster students' identification and involvement with the images by enhancing a sense of group identity and solidarity (Van Leeuwen & Jewitt, 2001).

Another favored characteristic of the images revealed at the representational mode is the appreciation of Iranian sociocultural identity. Today, with the rise of globalization and the spread of the on- going processes of cultural globalization, teaching the culture has gained prominence. New insights into TESOL/EFL pedagogy (e.g., Kamaravadivelu, 2008, 2012) contend that the long-established sociological concepts of assimilation, multiculturalism, and hybridity do not adequately address the political, religious and cultural complexities that cultural globalization has brought about. In this situation, it is essential that English learners experience a process of social and political awakening wherein they are directed to question the traditional assumptions about relative cultural positioning and are allowed to maintain their individual and collective identities and become aware of the social realities around them (Kamaravadivelu, 2012). Notwithstanding some farfetched portrayals, the fact that the images represent the practices and attitudes in localized ways is potentially of great value.

Unfortunately however, these merits are dwindled away by numerous flaws. The first and probably the most important consideration relates to the modality of

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

the images. As KvL (2006) note, modality rests on culturally and historically determined standards of what counts as real. The interpretation of reality for young students is remarkably different from that of past generations; today's fast growing technology, modern color prints, and computer enhanced images have constituted a kind of standard for visual modality for the present time culture. In this respect, the images in the high school EFL course books seem to be lagging behind today's technological advancements. The images are black and white and drawn in a very plain and simplistic form; there is no trace of technological advancement and social practices; and also the objects and entities are represented in outdated ways. Lack of proper contextualization adds more fuel to the flat and monochrome images. They are undoubtedly not favored by adolescents who are already familiar with high technology and are besieged with colorful and stunning pictures in other EFL materials. Hence, these images do not seem to bea good medium for the representational depiction of the way the world looks in real. Another flaw recognized in the analysis of the representational mode relates to stereotyped gender depiction. The results confirm that females are under-represented in the images and are overloaded with sex-linked occupations (in social sphere) and traditional stereotyped activities (in domestic settings).

Nonetheless, the flaws observed at different levels of meaning construction might find explanation in financial and economic reasons. Reasonably, incorporation of color, graphic art, photographs, and other advanced pictorial representations seem to be abandoned to trade off the cost of course books and thus make their price affordable for all students, especially those living in remote rural regions. In fact, Iranian high school EFL course books have been revised in 1997/1998 and since then they have been republished repeatedly over the years without revisions or modifications. But winds of change have been strongly felt since then and it is high time to be out with the old and in with the new. Iranian school curricula in general and EFL course books in particular will suffer if they do not catch up with current trends of technology and pedagogy. In fact, many of the above-discussed shortcomings will be rectified only if the EFL course books are extensively revised and refined.

Conclusion

Drawing on KvL's (2006) model of visual grammar, this study gives an account of the images in Iranian high school EFL course books. The findings confirm that at representational level, the images depict the participants in factual and familiar

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

ways which have a reasonable correspondence with 15 to17 year-old high-school students. With regard to the interactive meanings characterized in terms of social distance and perspective, the images render messages of equality and objectivity which warrant the association of the students with the images. At the compositional level, the status of the images exhibits salience over the texts accompanying them because they communicate paramount informative/illustrative functions as they provide contextual cues for exercises and verbal units. Besides, the images are mostly placed in the right (New) or bottom (Real) layout position which indicates their concrete and tangible role as clues to provide the students with real facts and new information required to make sense of the attached texts.

However, these valuable features are impinged by shortcomings observed at the representational and interactive modes, such as: poor modality due to grayscale printing and low-quality and plain graphic design, stereotyped messages, incongruous depictions of Iranian individuals with Western names, and out of date depiction of present Iranian lifestyle. In sum, Iranian high school EFL course books seem to be lagging behind the technological advancements of the current time. Thus, in their present format, the images not only hinder the intended pedagogical objectives of the English course books but also may demotivate and disinterest the students. Therefore, there is a definite need for immense changes in the quality of technical presentation of images as well as their ideological underpinnings. Nevertheless, strategies that editors should adopt to exploit the full potential of images to cater for the pedagogical objectives of EFL course books, and at the same time remain within the mainstream paradigms of Iranian culture, warrant scholarly attention and further exploration.

Notes on Contributors

Mohammad Hassan Tahririan *is a Professor of applied Linguistics at Sheikhbahaee University who has widely published in International journals and supervised many MA and PhD students. His major areas of interest include EFL material development and ESP.*

Elham Sadri *is a PhD candidate in TEFL at Sheikhbahaee University. Her areas of interest include text analysis and ESP.*

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

References

Ansary, H., & Babaii, E. (2002). Universal characteristics of EFL/ESL textbook: A step towards systematic textbook evaluation. *The Internet TESL Journal, 2*, pp. 1-8.

Azizifar, A., Koosha, M., & Lotfi, A. R. (2010). An analytical evaluation of Iranian high school ELT textbooks from 1970 to the present. *Procedia-Social and Behavioral Sciences, 3*, 36-44.

Bailey, M. (1995). The Impact of Integrating Visuals in an Elementary Creative Writing Process. Retrieved from http://www.eric.ed.gov/ERICWebPortal

Barsalou, L. W. (2003). Situated simulation in the human conceptual system.*Language and Cognitive Processes, 18*-5/6, 513-562.

Barthes, R. (1977). *Image music text*. UK : Harper Collins.

Bransford, J. D., & Johnson, M. K. (1973).Considerations of some problems of comprehension., In W. G. Chase (Ed.), *Visual information processing* (pp. 383-438). New York: Academic Press.

Canning, C. (1997). Theoretical use of Visuals. In C. Canning & J. Koester, *Illustrated Visual Aids for Academic English, 1*(1), 2-4.

Canning, C. (1998). Visual support and language teaching, *TESOL Arabia News, Volume 5* (4), 3-4.

Canning-Wilson, C. (1999). *Using Pictures in the EFL/ESL Classroom*. Retrieved fromhttp://www.eric.ed.gov/ERICWebPortal/search

Chen, Y. (2010). The semiotic construal of attitudinal curriculum goals: Evidence from EFL textbooks in China. *Linguistics and Education, 21*(1), 60-74.

Curtis, A., & Bailey, K. M. (2001). Picture your students talking: Using pictures in the language classroom. *ESL Magazine, 4*(4), 10-11.

Di Carlo, A. (1994). Comprehensible input through the practical application of video texts in second language acquisition. *Italica, 71*(4), 465-483.

Halliday, M. A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold

Halliday, M. A. K. (1994). *An Introduction to Functional Grammar*. London: Edward Arnold.

Hegarty, M. (2004). Dynamic visualizations and learning: Getting to difficult questions. *Learning and Instruction, 14,* 343-351.

Kress, G., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design (1st Edition)*. London: Rutledge.

Kress, G., & Van Leeuwen, T. (2006). *Reading images: The grammar of visual design (2nd Edition)*. NY: Rutledge.

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]

Kumaravadivelu, B. (2008). *Cultural globalization and language education*. Yale University Press.

Kumaravadivelu, B. (2012). *Language teacher education for a global society: A modular model for knowing, analyzing, recognizing, doing and seeing*. NY: Routledge.

Lewis, D. (2001) Reading Contemporary Picture Books. In G. Kress and T. van Leeuwen (Eds.), *Grammar of Visual Design* (pp. 145-670). London: Routledge.

Martin, J. R. (1985). Process and text: two aspects of human semiosis. *Systemic perspectives on discourse, 1*(15), 248-274.

Mendoza, J., & Reese, D. (2001).Examining multicultural picture books for the early childhood classroom: Possibilities and pitfalls. *Early Childhood Research and Practice, 3*(2), 1-31.

Mueller, G. (1980). Visual Contextual Clues and Listening Comprehension: Anexperiment. *Modern Language Journal, 64,* 335-40

Omaggio, A. (1979). Pictures and Second Language Comprehension: Do they Help? *Foreign Language Annals, 12*, 107-16.

O'Toole, M. (1994).*The Language of Displayed Art*. London: Leicester University Press

Rong, W. A. N. G. (2002). The Application of Visuals-based Listening Exercises in the Teaching of Listening and Speaking Course. *Media in Foreign Language Instruction, 1,* 004.

Royce, T. D. (2007). Intersemiotic complementarity: a framework for multimodal discourse analysis. In T. D. Royce & W. L. Bowcher (Eds.), *New Directions in the Analysis of Multimodal Discourse* (pp. 63-110). Mahwah: Lawrence Erlbaum.

Unsworth, L. (2008). Multiliteracies and metalanguage: Describing image/text relations as a resource for negotiating multimodal texts. In D. Leu, J. Corio, M. Knobel, & C. Lankshear (Eds.), *Handbook of Research on New Literacies*. New Jersey: Erlbaum.

Van Leeuwen, T., & Jewitt, C. (Eds.). (2001). *The handbook of visual analysis*. Sage.

Yarmohammadi, L. (2002). The evaluation of pre-university textbooks. *The Newsletter of the Iranian Academy of Science*, *18*, 70-87.

[Downloaded from ijal2.khu.ac.ir on 2026-11-25]