

Engaging Young Learners Through Human Images: A Cross-Cultural Comparative Study of English Textbooks in the US and China

Hai Nin Yeoh

Universiti Malaya, Malaysia

Email: violetyeohhn@gmail.com

ORCID: <https://orcid.org/0009-0008-0472-4002>

Huey Fen Cheong

Universiti Malaya, Malaysia

Email: hfcheong@um.edu.my

ORCID: <https://orcid.org/0000-0001-9274-4442>

Rui Jin

Universiti Malaya, Malaysia

Email: rkim79804@gmail.com

Abstract:

ELT materials produced by Western native speakers are often highly valued in EFL/ESL contexts. However, such materials may embody particular practices and beliefs about social interaction and education that may not adequately account for cultural differences across EFL/ESL contexts. This study examines how locally produced American and Chinese primary school English textbooks interact with learners via human images. Drawing on Kress and van Leeuwen's Visual Grammar and Hofstede et al.'s cultural dimensions, this study analysed the interactional metafunction, focusing on gaze, social distance, and power relations in 144 human images from two sixth-grade textbooks. Findings show that the American textbook more frequently employs direct gazes and close-up shots, inviting personalised engagement and reflecting student-centred pedagogical orientations associated with individualism. In contrast, the Chinese textbook predominantly uses indirect gazes and medium-to-long shots, positioning human figures as objects of observation within broader social contexts and reflecting more collectivist and teacher-centred educational orientations. Power relations are generally balanced in both textbooks. These findings show that ELT materials do not just teach the English language, but also embed cultural differences, particularly in social interaction (where language is used) and education. The interaction between textbooks and learners also signals teacher-student interaction. All these highlight the ongoing tension between the push for globalisation and the call for deglobalisation, with the latter often advocating greater cultural inclusivity and sensitivity in ELT practices and materials.

Keywords: ELT textbooks, visual grammar, cultural comparison, learner engagement

1. Introduction

ELT materials produced by Western native speakers are often highly valued in EFL/ESL contexts, despite potential cultural misalignment in non-Western contexts (Lai, 2014; Weninger & Kiss, 2013). Si's (2019) study revealed the continued dominance of English-as-a-native-language (ENL)-informed ELT materials in the Chinese EFL context, highlighting insufficient attention and effort to include



linguistic and cultural diversity for broader global varieties, particularly the learners' own cultures. Similarly, Yeoh and Cheong (2025) criticised that Western-authored literature taught in ESL classrooms embedded cultural assumptions that may not fully align with the local ESL contexts. Amid ongoing discussions surrounding cultural differences in education between Western and Eastern contexts, as well as recent diplomatic tensions between the US and China, this preliminary study compares an American and a Chinese English textbook produced and used in their respective countries. This study aims to provide insights into cultural differences in teaching practices and the creation of teaching materials, contribute to broader global educational understanding, and support more informed cross-cultural awareness in the Western and Eastern educational contexts.

Past studies on language textbooks predominantly prioritised verbal features and non-human images (e.g., Keles & Yazan, 2023; Uzum et al., 2021). From a social semiotic perspective, human images constitute an important mode of meaning-making. This echoes the term “visual grammar” in a widely used framework in social semiotics – Kress and van Leeuwen’s (2020) Visual Grammar. Although previous studies have examined images in English language textbooks, less attention has been given to cross-cultural comparisons and how visual interactional patterns may reflect different educational and cultural contexts.

To address the limited attention given to cross-cultural visual analyses of human images in primary school English language textbooks, where imagery is highly relied upon for engaging young learners, this study examines how human images in American and Chinese English textbooks engage learners differently. This echoes McConachy’s (2018) view that textbooks mediate social interaction through verbal cultural representations, which can vary across different cultures.

This study, therefore, examines how human images in locally produced primary school English textbooks construct interaction with learners in different cultures using three key features: gaze, social distance, and power relations. It addresses two research questions:

- 1) How do the interactional features (gaze, social distance, and power relations) of human images differ between American and Chinese English textbooks?
- 2) How do these differences reflect the cultural values of their educational contexts?

2. Literature Review

2.1 Visual images in English textbooks

Past studies have analysed various features in textbooks, e.g., instructions, verbal texts, and images (Lucy et al., 2020). According to Elmiana (2019), images in EFL textbooks are attractive features that can help foster students’ motivation in language learning: “to engage effectively with the exercises and provide emphasis for the assertions made in the linguistic parts of the text” (p. 1). Beyond language acquisition, past studies show that these visuals teach social interaction. Elmiana (2019) reports that images in Indonesian EFL textbooks are highly idealised and interactive, constructing meanings designed to personalise learner engagement with the tasks. Purwaningtyas’ (2020) study demonstrates that visual images in Indonesian EFL textbooks project cultural tolerance and gender equity by balancing professional roles between male and female characters. Mantei and Kervin (2014) found that



picture book illustrations allow young learners to make personal, text-to-self connections reflecting their home and community experiences. In China, Chen (2010) demonstrated how multimodal resources like dialogue balloons allow the textbook's editor voice to interact with character and reader voices, fostering dialogic engagement.

However, cross-cultural research focusing on how human images construct learner engagement remains limited. While Elmiana (2019) provides insights into Indonesian EFL textbooks, the lack of cross-cultural comparison makes it difficult to interpret if those visual features reflect unique national values. Cross-cultural comparisons make these culturally specific interactional choices visible and interpretable.

2.2 Social interaction in the US and China

This study employs Kress and van Leeuwen's (2020) Visual Grammar, focusing on the interactional metafunction: eye contact, social distance, and power relations. Regarding eye contact, Uono and Hietanen (2015) report cross-cultural differences in direct gaze perception, with Western participants showing greater sensitivity to direct cues than Eastern participants in experimental settings. Stanley et al. (2013) found cultural differences in gaze allocation during emotion recognition tasks, with American and Chinese participants attending differently to facial and contextual information. Similarly, social distance has also been associated with cultural variation in interpersonal spacing and relational preferences (Beaulieu, 2004).

Regarding power relations, lower power distance is generally associated with Western individualist cultures such as the United States, while higher power distance is more commonly associated with Eastern collectivist cultures such as China (see Zhang, 2013). The Hofstede model shows systematic differences across cultures in individualism/collectivism and power distance. According to the more recent update from Hofstede et al. (2010), the United States is reported to score 91 on individualism, while China scores 20. For power distance, the United States scores 40, while China scores 80. Societies with lower power distance tend to be more egalitarian in social relations, whereas higher power distance societies tend to be more hierarchical, reflected in relations such as between authority figures and subordinates (Hofstede et al., 2010).

2.3 Social interaction in educational contexts in the US and China

Grant et al. (2013) observed that American teachers prioritise individualised instruction and lesson differentiation, whereas Chinese teachers lean heavily toward cohesive, whole-group instruction. Numerous studies suggest that the Western individualist education (e.g., the US) is often described as more student-centred, where students are encouraged to participate actively, ask questions, and contribute to classroom discussion (Nelson, 2000). In contrast, the Eastern collectivist education (e.g., in China) is often described as more teacher-centred, where teachers often dominate the interaction in the class with the goal of passing knowledge to the collective students. Zhang (2013) notes that Chinese students tend to perceive classroom relations as hierarchical, with teachers occupying a higher position, and students may prefer seeking help from peers rather than directly interacting with their teachers.



To date, past studies have primarily analysed the interaction among real individuals in physical educational settings. This study extends the literature by examining interactional meaning embedded in textbook human images.

3. Methodology

This preliminary study focuses on one English textbook each from the US and China, used in primary schools of prominent states/provinces, California and Beijing, respectively. This study employs Hofstede et al.'s (2010) cultural dimensions, focusing on individualism versus collectivism, and Kress and van Leeuwen's (2020) Visual Grammar, emphasising the interactional metafunction: how images engage viewers through gaze (direct/indirect), social distance (close-up, medium, long shots), and power relation (high, equal, low angles). Human images (53 American, 91 Chinese) were collected from two nationally approved sixth-grade English textbooks used in the US (Wonders: Reading/Writing Workshop; McGraw Hill, 2016) and China (English Subject; People's Education Press, 2013). These images were coded and quantified based on the interactional features and interpreted through Hofstede et al.'s (2010) cultural lens.

For analytical clarity, the paper tends to explain cultures in broad generalisations with simplified binaries of East–West and US–China, particularly from the cultural aspect of individualism and collectivism. We recognise that regional samples are not fully representative and that treating cultural orientations as a strict binary can overlook individual or subcultural diversity along the cultural continuum (Hofstede et al., 2010; McSweeney, 2002). However, while recognising cultural diversity is crucial to understanding the cultural landscape, enabling discussion of cultural differences is equally important. See Cheong (2020, 2021) on intangible cultural diversity, tangible cultural differences, the demand for tangible knowledge, and how academic favouritism towards diversity and anti-generalisation stems from the hegemonic Western individualism that has been ruling academia.

4. Findings and Discussion

4.1 Eye contact: More direct gazes in the American English textbook and indirect gazes in the Chinese English textbook

Table 1. Gazes of Human Images in the American and Chinese English Textbooks

Textbook	Frequency (Percentage)		
	Direct gaze	Indirect gaze	Total
The American English textbook	31 (58%)	22 (42%)	53 (100%)
The Chinese English textbook	16 (18%)	75 (82%)	91 (100%)



Table 1 shows the number of human images with direct and indirect gazes in the American and Chinese English textbooks, respectively. Human images in the American English textbook often use direct gazes (58%), creating eye contact and potentially increasing viewer engagement, while the Chinese English textbook favours indirect gazes (82%), where characters look away, offering information rather than demanding interaction (see Figure 1).



Figure 1. Examples of Direct Gaze in the American English Textbook and Indirect Gaze in the Chinese English Textbook

American textbook images, both individual and groups, often show direct gazes with positive or neutral expressions. This aligns with individualist design conventions linking direct viewer address to active learner engagement. Even groups maintain direct eye contact, suggesting a stronger orientation towards viewer address in visual design. In contrast, Chinese English textbook images avoid eye contact with viewers. This visual choice may reflect broader cultural differences in gaze behaviour reported in previous cross-cultural studies (Uono & Hietanen, 2015).

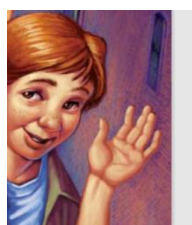
Exceptions include 42% indirect gazes in the American textbook, typically used for internal reflection or storytelling, and 18% direct gazes in the Chinese English textbook, often appearing in instructional or question-based contexts where learner response is invited.

4.2 Social distance: More close-up shots in the American English textbook and medium-to-long shots in the Chinese English textbook

Table 2. Camera Shots of Human Images in the American and Chinese English Textbooks

Textbook	Frequency (Percentage)			
	Close shot	Medium shot	Long shot	Total
The American English textbook	22 (42%)	16 (30%)	15 (28%)	53 (100%)
The Chinese English textbook	10 (11%)	44 (48%)	37 (41%)	91 (100%)

Table 2 shows the distribution of camera shots, corresponding to the social distance between the human images and students/readers. The American textbook features more close-up shots (42%), often featuring children of a similar age to the readers, which may encourage emotional identification and simulate peer interaction (see Figure 2).



Close-up shot
(American textbook)



Medium and long shot
(American textbook)



Long shot
(Chinese textbook)



Close-up shot
(Chinese textbook)

Figure 2. Examples of Camera Shots in the American and Chinese English Textbooks

An example shows a girl's head and hand greeting students, fostering familiarity, especially when combined with direct gaze. In contrast, the Chinese English textbook uses medium (48%) and long shots (41%) to portray group activities and context, which positions the characters as objects of observation, consistent with collectivist values favouring social harmony and teacher-guided learning.

Some variation exists: the American textbook uses long shots (28%) and medium-to-long shots (30%) to show teamwork or child-adult dynamics. Close-up shots (11%) in the Chinese English

textbook mostly appear in vocabulary lessons, showing emotions like “joy” and “thought”, but often serve informational rather than engaging purposes, indicating that distance does not always equate detachment.

4.3 Power relation: Equal angles/power in both American and Chinese English textbooks, with low angles only in the American textbook

Table 3. Camera Angles of Human Images in the American and Chinese English Textbooks

Textbook	Frequency (Percentage)			
	High angle	Low angle	Equal angle	Total
The American English textbook	2 (4%)	8 (15%)	43 (81%)	53 (100%)
The Chinese English textbook	4 (4%)	0 (0%)	87 (96%)	91 (100%)

Table 3 presents the distribution of camera angles. Most images in both textbooks are presented from equal angles, 81% in the American and 96% in the Chinese English textbook. This likely reflects standard textbook design conventions for reader comfort rather than intentional power dynamics (see Figure 3).



Equal angle
(American textbook)



Low angle
(American textbook)



Equal angle
(Chinese textbook)



High angle
(Chinese textbook)

Figure 3. Examples of Camera Angles in the American and Chinese English Textbooks

The American textbook uses low angles (15%) to depict characters as brave and admirable, inviting students to look up to them. The Chinese English textbook’s few high-angle shots (4%) offer wider

group views, reinforcing collective activity and framing subjects as visual information rather than interaction partners.

5. Conclusion

This study reveals distinct visual interaction styles shaped by cultural values. The American textbook personalises interaction through direct gazes and close-ups, reflecting Western individualism and active, student-centred learning. The Chinese English textbook favours indirect gazes and medium-to-long shots, reinforcing collectivist, social learning, and passive observation. Both use mostly equal angles, likely reflecting design conventions rather than power dynamics.

These findings echo broader cultural differences between Western individualism and Eastern collectivism, aligning with educational practices, the roles of teachers, and materials. While offering insights into cultural patterns in textbooks, this preliminary study acknowledges the complexity of cultures and cautions against overgeneralising US–China/East–West binaries. Broader research across diverse contexts is needed to understand how visual design shapes and reflects cultural approaches to education. Culture must be compared to be observable and distinguishable from other cultures, and thus, claimable and justifiable as a culture.

All in all, these findings show that ELT materials do not just teach the English language, but also embed cultural differences, particularly in social interaction (where language is used) and education. The interaction between textbooks and learners also signals teacher-student interaction. All these highlight the ongoing tension between the push for globalisation and the call for deglobalisation, with the latter often advocating greater cultural inclusivity and sensitivity in ELT practices and materials.

Declarations and Acknowledgement:

The author declares there is no conflict of interest.

References

- Beaulieu, C. (2004). Intercultural study of personal space: A case study. *Journal of Applied Social Psychology*, 34(4), 794–805.
- Chen, Y. (2010). Exploring dialogic engagement with readers in multimodal EFL textbooks in China. *Visual Communication*, 9(4), 485–506.
- Cheong, H. F. (2020). *A problematised critical approach: Constructions of metrosexualities in the UK and Malaysia* [Doctoral Thesis, Lancaster University]. <https://doi.org/10.17635/lancaster/thesis/955>
- Cheong, H. F. (2021). Let the culture speak: How gender is socially/culturally constructed by Western individualism and Eastern collectivism. *Gender, Place & Culture*, 28(2), 299–304. <https://doi.org/10.1080/0966369X.2020.1789072>



- Elmiana, D. S. (2019). Pedagogical representation of visual images in EFL textbooks: A multimodal perspective. *Pedagogy, Culture & Society*, 27(4), 613–628.
- Grant, L. W., Stronge, J. H., & Xu, X. (2013). A cross-cultural comparative study of teacher effectiveness: Analyses of award-winning teachers in the United States and China. *Educational Assessment, Evaluation and Accountability*, 25(3), 251–276. <https://doi.org/10.1007/s11092-013-9170-1>
- Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software for the mind* (3rd ed.). McGraw-Hill.
- Keles, U., & Yazan, B. (2023). Representation of cultures and communities in a global ELT textbook: A diachronic content analysis. *Language Teaching Research*, 27(5), 1325–1346.
- Kress, G., & van Leeuwen, T. (2020). *Reading images: The grammar of visual design* (3rd ed.). Routledge. <https://www.taylorfrancis.com/books/mono/10.4324/9781003099857/reading-images-gunther-kress-theo-van-leeuwen>
- Lai, H.-Y. T. (2014). Learning English as an international language: EFL learners' perceptions of cultural knowledge acquisition in the English classroom. *Asian Social Science*, 10(1), 1. <https://doi.org/10.5539/ass.v10n1p1>
- Lucy, L., Demszyk, D., Bromley, P., & Jurafsky, D. (2020). Content analysis of textbooks via natural language processing: Findings on gender, race, and ethnicity in Texas US history textbooks. *AERA Open*, 6(3), 2332858420940312.
- Mantei, J., & Kervin, L. (2014). Interpreting the images in a picture book: Students make connections to themselves, their lives and experiences. *English Teaching: Practice and Critique*, 13(2), 76–92.
- McConachy, T. (2018). Critically engaging with cultural representations in foreign language textbooks. *Intercultural Education*, 29(1), 77–88.
- McGraw Hill. (2016). *Wonders Reading/Writing Workshop, Grade 6*. McGraw-Hill Education. <https://www.mheducation.com/prek-12/product/wonders-reading-writing-workshop-grade-6-mcgraw-hill/9780076765737.html>
- McSweeney, B. (2002). Hofstede's model of national cultural differences and their consequences: A triumph of faith-a failure of analysis. *Human Relations*, 55(1), 89–118.
- Nelson, G. (2000). Individualism-Collectivism and Power Distance: Applications for the English as a Second Language Classroom. *The CATESOL Journal*, 12(1). <https://doi.org/10.5070/B5.36463>
- People's Education Press. (2013). *English Subject: Grade 6, Volume 1*. People's Education Press.
- Purwaningtyas, T. (2020). Didactic symbol of visual images in EFL textbook: Multi-modal critical discourse analysis. *Pedagogy: Journal of English Language Teaching*, 8(1), 51–63.
- Si, J. (2019). English as a native language, World Englishes and English as a lingua franca-informed materials: Acceptance, perceptions and attitudes of Chinese English learners. *Asian Englishes*, 21(2), 190–206. <https://doi.org/10.1080/13488678.2018.1544700>
- Stanley, J. T., Zhang, X., Fung, H. H., & Isaacowitz, D. M. (2013). Cultural differences in gaze and emotion recognition: Americans contrast more than Chinese. *Emotion*, 13(1), 36–46. <https://doi.org/10.1037/a0029209>
- Uono, S., & Hietanen, J. K. (2015). Eye contact perception in the west and east: A cross-cultural study. *PLOS ONE*, 10(2), e0118094.
- Uzum, B., Yazan, B., Zahrawi, S., Bouamer, S., & Malakaj, E. (2021). A comparative analysis of cultural representations in collegiate world language textbooks (Arabic, French, and German). *Linguistics and Education*, 61, 100901.
- Weninger, C., & Kiss, T. (2013). Culture in English as a foreign language (EFL) textbooks: A semiotic approach. *TESOL Quarterly*, 47(4), 694–716.



- Yeoh, H. N., & Cheong, H. F. (2025). That's Not My Gender!: Gender stereotyping in western-authored children's literature taught in Malaysian primary ESL classrooms. *Linguistics and Education*, 89, 101463. <https://doi.org/https://doi.org/10.1016/j.linged.2025.101463>
- Zhang, Y. (Leaf). (2013). Power Distance in Online Learning: Experience of Chinese Learners in U.S. Higher Education. *International Review of Research in Open and Distributed Learning*, 14(4), 238–254. <https://doi.org/10.19173/irrodl.v14i4.1557>

