Proceeding of ICSPHS 2026 Symposium: Critical Perspectives on Global Education and Psychological Development
DOI: 10.54254/2753-7048/2026.HT29533

# *A Brief Review of Diagnosing and Repairing Misinterpretation in IELTS TFNG via ChatGPT Scaffolds*

**Xingrui Hu**

*Faculty of Education Studies, Universiti Putra Malaysia, Selangor, Malaysia*
*gs70731@student.upm.edu.my*

***Abstract.*** This study reviews the recent literature on how Chinese EFL graduate learners explain and correct micro-reading errors regarding scope, polarity, attribution, temporality, and lexical-semantic drift when using ChatGPT as scaffolding to complete IELTS True/False/Not Given (TFNG) reading passages. Based on Kintsch's Construction-Integration Model, this study models the reading comprehension process as a verification process rather than recall. This study positions ChatGPT as an assistant for evidence-seeking, and learners are guided to extract evidence, paraphrase, refine, and check modifiers before making a decision. This study also aims to construct an event-level "misunderstanding remediation" coding framework and develop an actionable toolkit including prompt templates and verification checklists. The findings are expected to provide pedagogical and interface suggestions for evidence- based reading in AI-assisted language learning, such as reducing automation bias.

***Keywords:*** Evidence-based reading, EFL comprehension verification, Cognitive processing in L2 reading, AI-assisted language learning, Misinterpretation and repair

## 1. Introduction

In the IELTS Reading test, True-False-Not Given (TFNG) questions assess learners to verify statements based on text, distinguishing between confirmation, contradiction, and lack of information. The most difficult part is conceptual confusion caused by "Not Given" when inference is needed [1-2]. Misunderstandings at the micro-levels such as scope and quantification, negation, speaker attribution, temporal or conditional relationships, and lexical- semantic drift, pose significant error risks [3-4]. Nowadays, GenAI, like ChatGPT is widely used to scaffold EFL learners' reading skills. It can support comprehension, evidence verification and adaptation for TFNG reading tasks [5]. Most of the previous studies focused on output accuracy for comprehension verification and reasoning in L2 reading. However, there is limited research investigating process-level engagement with AI feedback in L2 reading. This study reviews the literature on how Chinese EFL graduate learners interact with ChatGPT-generated main points and evidence during IELTS reading tasks. Specifically, this study investigates which key points trigger verification, how learners correct misconceptions generated by ChatGPT, and the relationship between decision-making and confidence levels. Ultimately, it seeks establish evidence-priority scaffolding that promotes verifiable reasoning and reduces automation bias in AI- assisted reading [6].

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-7048/2026.HT29533

## 2. Literature review

### 2.1. Definition of reading comprehension

Reading comprehension is a construction and verification process in which the meaning of a text is built and tested by the reader, revised and further constructed. According to Kintsch's Construction-Integration Model, comprehension is “the activation of propositional representations of the text, which are ultimately integrated into a coherent structure representing the meaning of the text” [7]. The accuracy of the meaning of a text depends on whether the activated propositional representations remain consistent with textual information and whether the qualifiers, scope, and relationships are properly preserved [7]. Empirical research showed that when textual information conflicts with prior context, readers would usually spend more time reanalyzing and repairing meaning. When learners skip this verification leads to misunderstandings [8]. For EFL readers, the incidence of misunderstandings would be more frequent as their limited linguistic knowledge would damage the verification process [3,9].

### 2.2. Micro-level misinterpretations and learner repair strategies

The empirical research has identified several recurring micro-level sources of misinterpretations among EFL learners. First, when qualifying words that are in textual evidence but are omitted from propositions (only, at least, up to, most), it would result in the logical boundary expansion or contraction of scope and quantifier drift [7,10]. Second, confusion between negation and polarity would ask learners to do extra mental operations. If verification is omitted, misunderstandings would be retained in memory and result in erroneous judgments [11-12]. Third, when attending to local information, EFL learners would usually experience misunderstandings of pronoun mismatches/omissions resulting in information processing misunderstanding [10,13]. Fourth, when conjunctions mismatch semantic relationships, learners would experience processing conflicts due to unstable logical mapping and would consequently slow down reading speed [4]. Finally, when logical competition results in misclassifying similar words as logically equivalent and doing 'good-enough processing,’ learners would experience lexical inference errors [10].

These micro-level misunderstandings arise because EFL readers, when conducting online syntactic analysis, lack sufficient depth and precision in their syntactic knowledge, relying instead on lexical semantic cues [3], which reduces their ability to verify propositions accurately against textual evidence. This logical competition would become salient in IELTS Reading TFNG tasks as they flexibly transform reading comprehension into evidence-based verification [14]. However, verification errors in EFL readers are not random; they are systematic results of logical competition between form-level parsing and higher-order inference. Therefore, effective comprehension would result in restarting verification through rereading, modifier checking, or repeated sentence comparison. To overcome such verification breakdowns, learners typically engage in repair strategies such as rereading, modifier checking, and cross-sentence comparison [8,13]. These verification misunderstandings are the direct causes of this study’s focus: how can Chinese EFL graduate learners detect and rectify these micro-level misunderstandings in IELTS TFNG tasks with ChatGPT scaffolding.

DOI: 10.54254/2753-7048/2026.HT29533

### 2.3. ChatGPT as a cognitive scaffold for evidence-based reading verification

The application of artificial intelligence in language learning is becoming more and more common. At the reading comprehension level, this application has changed the way EFL readers understand texts. This study views ChatGPT as a cognitive scaffold that materializes verification procedures [5,8]. Wang et al.'s study shows that personalized scaffolding based on ChatGPT can enhance learning achievement and motivation without blowing up mental effort but slightly increase productive mental load through ZPD-aligned difficulty (particularly conducive to verification-oriented reading) [5]. With step-by-step prompts and feedback, learners can use AI-generated cues to test their verification process, making verification an explicit process instead of an implicit one.

In the context of IELTS TFNG reading, ChatGPT can support "proof-before-verdict" reasoning by guiding evidence extraction, paraphrase refinement, and qualifier checking [5]. Critically for EFL readers, whose on-line parsing tends to rely more on lexical–semantic cues and less on detailed syntactic representations [3], such scaffolds make validation steps explicit (e.g., locating warrants, preserving scope/negation, resolving anaphora), thereby compensating for reduced verification sensitivity [5,8]. However, Klingbeil et al. show that learners may over-rely on AI-generated propositions even when these diverge from source-text logic, revealing automation bias and pseudo-coherence in comprehension tasks [15]. Hence, AI support must be framed as a source-based verification scaffold rather than an answer provider.

For EFL readers, interaction with ChatGPT can render verification sensitive (i.e., aware of logical scope, negation, and referential consistency) during reading verification. Compared with native readers, Chinese EFL learners misinterpret more TFNG items by integrating syntactic and semantic cues [3,14]. Therefore, the AI serves a compensatory role in externalizing many verification steps (e.g., evidence retrieval, paraphrase refinement, scope checking), which are implicit in traditional reading processes [5,8].Therefore, it is necessary to examine how Chinese EFL graduate learners engage with ChatGPT to diagnose and repair micro-level misunderstandings in TFNG reading, revealing both the potential and limits of AI-mediated evidence alignment in academic reading contexts. Therefore, it is necessary to explore how Chinese EFL graduate learners utilize ChatGPT to diagnose and rectify micro-misunderstandings in TFNG tasks, to uncover the potential and limitations of AI-mediated evidence within academic reading contexts.

## 3. Discussion

Recent research focuses on the process of EFL reading comprehension, where understanding depends on constructing propositions and verifying them against textual evidence. Micro-level risks —such as scope/quantification, polarity, attribution, temporal/conditional cues, and lexical semantic drift—are highly likely to occur when verification processes are neglected. Research also indicates that while AI tools like ChatGPT can support reading outcomes and motivation through structured prompts and assessment patterns when intervening in reading ability, certain potential risks exist.

Although these studies provide a starting point for exploring how comprehension happens under AI mediation, they remain mostly outcome oriented, we still lack evidence tracking the verification process, specifically, how Chinese EFL learners initiate, repair, and complete verification in IELTS TFNG tasks and whether ChatGPT can really improve students' verification-based reading process while misleading fluent comprehension. By focusing exclusively on reading proficiency gains, current research confines structural interpretation to test scores, failing to specify process-level, observable consistency indicators for evidence. This limitation prevents sufficient leverage for instructional and interface design decisions.

DOI: 10.54254/2753-7048/2026.HT29533

## 4. Conclusion

This study investigates how Chinese EFL graduate learners explain and correct micro-level misunderstandings when using ChatGPT scaffolding to complete IELTS TFNG reading tasks. Conceiving ChatGPT as an evidence-seeking assistant, this study aims to build an evidence-first workflow integrating candidate evidence extraction, interpretation refinement, qualifier checking, and verification prior to final judgment. By developing an event-level "misunderstanding-to-repair" coding scheme and actionable toolkit (including prompt templates, verification checklists, and evidence cards) aiming to make the verification process observable and teachable as routine practice. Future research can (i) extend low-burden process tracing to quantify verification intensity and confidence calibration, (ii) test generalization beyond TFNG and across genres, and (iii) compare model/interface variants to enhance proof-seeking while mitigating pseudo-coherence and over-reliance.

## References

[1] Hughes, A., 2003. Testing for Language Teachers. Cambridge University Press.
[2] Alderson, J.C., 2000. Assessing Reading. Cambridge University Press.
[3] Clahsen, H., Felser, C., 2006. Grammatical processing in language learners. AppliedPsycholinguistics27, 3–42. https: //doi.org/10.1017/s0142716406060024
[4] Zufferey, S., Mak, W., Degand, L., Sanders, T., 2015. Advanced learners' comprehension of discourse connectives: The role of L1 transfer across on- line and off-line tasks. Second Language Research 31, 389–411. https: //doi.org/10.1177/0267658315573349
[5] Wang, X., Zhong, Y., Huang, C., Huang, X., 2024. ChatPRCS: A Personalized Support System for English Reading Comprehension Based on ChatGPT. IEEE Transactions on Learning Technologies 17, 1722–1736. https: //doi.org/10.1109/tlt.2024.3405747
[6] Parasuraman, R., Sheridan, T.B., Wickens, C.D., 2000. A model for types and levels of human interaction with automation. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 30, 286–297. https: //doi.org/10.1109/3468.844354
[7] Kintsch, W., 1998. Comprehension: A Paradigm for Cognition. Cambridge University Press.
[8] Singer, M., 2013. Validation in Reading Comprehension. Current Directions in PsychologicalScience22, 361–366. https: //doi.org/10.1177/0963721413495236
[9] Bernhardt, E., 2005. PROGRESS AND PROCRASTINATION IN SECOND LANGUAGE READING. Annual Review of Applied Linguistics 25, 133–150. https: //doi.org/10.1017/s0267190505000073
[10] Sanford, A.J., Graesser, A.C., 2006. Shallow Processing and Underspecification. Discourse Processes 42, 99–108. https: //doi.org/10.1207/s15326950dp4202_1
[11] Carpenter, P.A., Just, M.A., 1975. Sentence comprehension: A psycholinguistic processing model of verification. Psychological Review 82, 45–73. https: //doi.org/10.1037/h0076248
[12] Schroyens, W.J., Schaeken, W., D'Ydewalle, G., 2001. The processing of negations in conditional reasoning: A meta-analytic case study in mental model and/or mental logic theory. Thinking & amp; Reasoning 7, 121–172. https: //doi.org/10.1080/13546780042000091
[13] Sanford, A.J., Garrod, S.C., 1998. The role of scenario mapping in text comprehension. Discourse Processes 26, 159–190. https: //doi.org/10.1080/01638539809545043
[14] Moore, T., Morton, J. and Price, S., 2012. Construct validity in the IELTS academic reading test (Vol. 11). IELTS Research Report 4.
[15] Klingbeil, A., Grützner, C., Schreck, P., 2024. Trust and reliance on AI — An experimental study on the extent and costs of overreliance on AI. ComputersinHumanBehavior160, 108352. https://doi.org/10.1016/j.chb.2024.108352