ISELT 224
Interdisciplinary Studies in English Language Teaching

Pages 224 to 242

# When and Why Does Typographical Emphasis Help? A Mixed-Methods Study of Genre-Proficiency Moderation and Iranian EFL Learners' Perceptions

Parviz Ajideh [1], Mohammad Zohrabi [1], Farzad Ranjbar [2]*

[1] Professor, Department of English, Faculty of Persian Literature and Foreign Languages, University of Tabriz, Tabriz, Iran

[2] Department of English, Faculty of Persian Literature and Foreign Languages, University of Tabriz, Tabriz, Iran

10.22080/iselt.2026.33270.1238

**Received**
October 14, 2026
**Accepted**
November 8, 2026
**Available online**
November 13, 2026

**Keywords:**
EFL, Learner Perceptions, Mixed Methods, Reading Comprehension, Typographical Emphasis.

**Abstract**

Typographical emphasis markers such as boldface, italics, and underlining are common in second-language reading materials, but evidence does not show when they help most or how learners use them. This convergent mixed-methods study examined how text genre and learner proficiency moderate the effects of emphasis and how 48 Iranian English as a foreign language (EFL) learners (24 lower-intermediate, 24 advanced) experienced the markers. The quantitative strand was a 640-participant, 4 × 4 × 2 factorial reading-comprehension experiment; the qualitative strand used semi-structured interviews analyzed thematically. Emphasis effects differed significantly across genres and across proficiency levels, but the three-way interaction of genre, proficiency, and emphasis was not significant. The largest gains occurred for advanced learners reading expository texts with bold emphasis, whereas narrative texts showed little benefit. Thematic analysis identified four patterns: awareness and noticeability, perceived helpfulness and strategic reading, preference for boldface, and distraction when emphasis was overused. Interviews broadly matched the quantitative pattern, particularly the greater strategic use reported by advanced learners. Together, the strands suggest that emphasis first makes information more noticeable and can then support selective attention and strategic reading, depending on genre demands and readers' ability to use the cue.

## 1. INTRODUCTION

Teachers and textbook writers mark texts all the time. Key terms appear in bold, new vocabulary in italics, and the line a student must not miss gets underlined. These conventions are so familiar that we rarely ask what they do for the person reading (Bringhurst, 2013). In a classroom where English is taught as a foreign language (EFL) the question carries more weight, because learners

* **Corresponding Author:** Farzad Ranjbar, Instructor, Fajr 1 dormitory, University of Tabriz, Tabriz, Iran, **Email**: farzadranjbar92@gmail.com

Copyright © 2027 The Author(s). Published by the Faculty of Persian Literature and Foreign Languages at the University of Mazandaran. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license.

reading in English often work with limited vocabulary and slower word recognition, and a cue that helps them decide where to spend attention could change how much of a passage they take in (Grabe, 2009; Koda, 2005).

Research that bears on this question comes from several fields, and it suggests that boldface, italics, and underlining can influence second-language reading under some conditions (e.g., Kim, 2025; Rassaei & Jabbarpoor, 2025; Wu et al., 2024). Studies of multimedia learning find that cues which mark important information help learning a little (Richter et al., 2016; Schneider et al., 2018). Second-language researchers, working from the idea of noticing, have asked whether visually marked grammar forms are more likely to be picked up (Lee & Huang, 2008; Sharwood Smith, 1993). Theory in this area, including work on signaling and cognitive load, suggests that a visual cue is most valuable when readers have to decide for themselves which information deserves attention (Mayer, 2021; Sweller et al., 2011).

Two questions follow from this, and neither has received much attention. The first is for whom, and in what kind of text, emphasis works. Genre and proficiency have seldom been examined together as moderators of typographical emphasis. A narrative supplies its own thread of characters and events, while an expository text leaves more of the organizing to the reader (Graesser et al., 2011). A learner whose word recognition is still effortful has less capacity to spare than one whose recognition is automatic (Perfetti & Stafura, 2014). The second question is how learners themselves experience emphasis. Test scores can show that an effect occurred, but they cannot show whether learners noticed the markers, how they read them, or whether they felt helped or hindered. Studies that pair an experimental comprehension measure with learners' own accounts are rare.

This article takes up both questions. It reports two strands of a larger doctoral project with Iranian EFL learners: a factorial experiment on how proficiency, genre, and emphasis interact, and an interview study of how learners perceived the markers. The research questions were as follows.

**RQ1:** What are Iranian EFL learners' perceptions of the effectiveness of typographical emphasis in reading?

**RQ2:** Are there significant interaction effects among proficiency level, text genre, and typographical emphasis on learners' reading comprehension outcomes?

The two questions were examined within a convergent mixed-methods design, not as unrelated quantitative and qualitative exercises. The quantitative analysis identifies the pattern of interaction, and the interviews give participants' accounts of noticing, strategy use, preferences, and distraction. The two strands are interpreted together (Creswell & Creswell, 2018). The theoretical framework draws on the Signaling Principle (Mayer, 2021), Cognitive Load Theory (Sweller et al., 2011), the Noticing Hypothesis (Schmidt, 1990), and metacognitive accounts of strategic reading (Baker & Brown, 1984). These perspectives make different predictions about why a visual cue may or may not help, and the design allows those predictions to be examined with both outcome and learner-report data.

## 2. LITERATURE REVIEW

### What Typographical Emphasis Is

Typography is the craft of arranging written language so that it can be read. Over recent decades researchers have come to treat it as part of the perceptual conditions in which reading takes place, not as decoration added afterward (Bringhurst, 2013). Two older terms keep the discussion straight. Legibility is how easily single letters and words are recognized; readability is how easily

a stretch of text can be read and understood (Dyson, 2004). A page can be legible and still hard to read if its organization is poor.

Typographical emphasis markers belong to this larger system. In this article, the term covers three devices that leave the words themselves unchanged: boldface which thickens the letterforms, italics which slant them, and underlining which adds a line beneath them. The conventions grew out of printing practice. After the movable type spread in fifteenth-century Europe, printers varied weight, style, and capitalization to set headings, titles, quotations, and foreign words apart from running text, and the three devices settled into familiar roles: bold for importance, italics for terms and titles, and underlining mostly in handwriting and, later, for hyperlinks (Bringhurst, 2013). Textbook authors adopted these conventions with little evidence about their effects on readers.

What the three markers share is that they change how prominent a word looks without changing what it means. That property makes them attractive for teaching. A teacher can steer attention toward a definition, a claim, or a cause without simplifying the passage or adding a gloss, so learners still meet authentic language. Whether steering attention actually improves comprehension is a separate matter, and the answer depends on how the cue works on the reader.

## How Emphasis Might Guide Attention

Reading is selective. Skilled readers do not spend the same effort on every word; they decide where to look, how long to stay, and when to go back, and eye-movement research shows that these decisions respond to word difficulty, sentence structure, and the visual form of the text (Rayner et al., 2016). Attention is the gateway in this process, because information that is not attended is unlikely to be processed deeply or remembered (Posner & Petersen, 1990).

Visual salience offers one explanation of why a bold word can win attention. Salience is the degree to which an element stands out from its surroundings (Itti & Koch, 2001). In feature integration theory, properties such as brightness, orientation, and shape are registered early and in parallel across the visual field, and only afterward does attention bind them into objects (Treisman & Gelade, 1980), so a word with heavier strokes can be picked out before it is read. Later work added that attention is not driven by the stimulus alone: in Guided Search, bottom-up salience combines with top-down goals and expectations (Wolfe, 2021). This matters here, because readers who have spent years with textbooks have learned that marked words usually matter. An emphasis cue may thus work twice, once through perceptual contrast and once through what readers believe it means.

Two instructional theories turn this into a prediction about learning. The signaling principle holds that cues which highlight essential material help learners by cutting down on search and pointing to organization (Mayer, 2021). The cue adds no information; it changes how the existing information is processed. Cognitive Load Theory reaches a similar view from the side of working memory: a text that leaves readers to work out for themselves what is important spends capacity that could go to understanding (Sweller et al., 2011). Both theories also carry a warning. A signal helps only while it stays selective, because the more words are marked, the less any single marked word stands out.

These accounts have limits that matter for what follows. They explain how emphasis might direct attention. They do not show that a noticed word is understood, and they say little about whether the benefit is the same for every reader and every kind of text.

## Earlier Research on Signaling and Typography

The strongest evidence for signaling comes from multimedia learning. Meta-analyses of signaling, one on text-picture relations and one on media more broadly, report positive, but modest, effects on learning (Richter et al., 2016; Schneider et al., 2018), and a meta-analysis that compared highlighting made by learners with highlighting provided by instructors (Ponce et al., 2022) belongs to the same line of work. Older studies of organizational signals such as headings and summaries concluded that readers use signals to see how ideas are structured (Lorch, 1989; Lorch & Lorch, 1995). More recent work on digital highlighting adds a caution. Mason et al. (2024) found that active highlighting was not effective in itself during digital reading, although what students chose to highlight contributed to their literal comprehension and learning. The quality of the marking, and not the mere presence of a cue, seems to matter. Most of this research used multimedia materials or whole-text signals, not the word-level marking that boldface or underlining provides.

Studies of typography proper come from design, psychology, and human-computer interaction. They deal mostly with typeface, size, spacing, and layout, and their outcome measures are usually reading speed, preference, or eye movements (Dyson, 2004). Some of this work found that readers' stated preferences ran ahead of measurable differences in performance (Bernard et al., 2003). Results are also hard to compare, because typographical variables are usually changed together. Much of the work used proficient first-language readers on short laboratory tasks, so it says little about learners who divide their attention between decoding a foreign language and understanding it.

Work that examines word-level marking in connected text is recent and small. Wu et al. (2024) used event-related potentials to study how font emphasis affects reading comprehension. Kim and Clariana (2017) had 88 Korean learners of English at mixed proficiency read one of two versions of an expository science passage and concluded that coherent text signals strongly influenced comprehension, and Kim (2025) examined verbal and nonverbal text signals in second-language expository comprehension. These studies point in a useful direction. They leave open how learners experience emphasis themselves, and whether the benefit holds across kinds of text and levels of proficiency.

## Textual Enhancement and Noticing in Second-Language Reading

In applied linguistics, typographical emphasis has mostly been studied under another name. Schmidt's (1990) Noticing Hypothesis proposes that language forms must be consciously registered before they can be learned, and Sharwood Smith (1993) suggested that making forms perceptually prominent, which he called input enhancement, might help learners notice them. Textual enhancement is the written version of the idea: bold, italics, underlining, capitals, color, or larger type applied to target forms inside a text.

Results have been mixed. Lee (2007) examined textual enhancement together with topic familiarity and measured both reading comprehension and learning of the English passive, and Simard (2009) compared different enhancement formats. Enhanced forms sometimes attracted more attention, but gains in noticing did not always carry over to learning. Lee and Huang's (2008) meta-analysis of visual input enhancement found small effects overall, with variation across studies, and a more recent synthesis pooled 29 primary studies with 1,749 participants (Rassaei & Jabbarpoor, 2025). Much of this literature has asked whether enhancement works and has paid less attention to the conditions under which its effect grows or shrinks.

Eye tracking has helped to separate attention from learning. Studies such as Lee and Révész (2018), who tracked learners' eyes as they watched textually enhanced captions, make it possible to measure attention directly instead of inferring it from test scores, so that whether a learner looked at a form can be kept apart from whether the learner learned from it. The Noticing Hypothesis already implies this distinction: noticing is a condition for learning, not a guarantee of it. Recent studies have made enhancement respond to the reader. Jung et al. (2025) highlighted target collocations only when a learner's gaze reached them, recorded eye movements, and tested learning immediately and two weeks later. They concluded that highlighting can be a useful focus-on-form technique in task-based reading.

Nearly all of this work concerns grammar or vocabulary, and the usual outcome is a recognition or recall test on a target form. Whether emphasis changes comprehension of a whole text is a different question, since comprehension depends on building a coherent representation of the text and not on noticing a single form. Proficiency has been looked at here too. Lee (2021) compared two enhancement techniques and learner proficiency in a study of noticing, which fits the general expectation that learners with less automatic language processing lean more on external cues. Beyond a few studies like this, moderators have received little systematic attention.

Read critically, this body of work has weaknesses that matter here. Signaling research relies largely on adult first-language readers and on multimedia material, so its findings may not carry over to adolescents reading in a foreign language. Textual enhancement studies usually measure noticing or recall of single forms, which is a long way from understanding a passage. Experiments that do measure comprehension often rely on a single text, so differences between genres are tangled up with differences between topics. Moreover, almost no study asks learners what they made of the cues. Meta-analytic averages hide these differences, which is why moderators and learner reports deserve direct attention.

## Genre and Proficiency as Possible Moderators

Texts differ in how much work they leave to the reader. Genre is usually defined as a socially recognized type of text with a stable purpose, organization, and conventions (Hyland, 2004; Martin & Rose, 2008). Narratives follow characters, goals, and events in time, so readers can build a causal chain that holds the text together, which is one reason narratives are usually found easier to comprehend than informational texts (Best et al., 2008; Graesser et al., 2011). Expository texts organize information around concepts and their relations, bring in unfamiliar terms, and expect readers to supply structure and background knowledge (Best et al., 2008), and argumentative texts add claims, evidence, and counterargument that readers must tell apart. In Kintsch's (1998) construction–integration model, the amount of integration a text demands varies with its organization. Studies of Iranian EFL learners have compared literal and inferential comprehension across expository and narrative texts (Saadatnia et al., 2017). If emphasis is useful because it helps readers find structure, it should matter most where structure is hardest to find.

Proficiency changes what the reader brings to the page. In second-language reading, it is among the strongest correlates of comprehension (Jeon & Yamashita, 2014). The lexical quality hypothesis explains part of the reason: precise, well-connected word knowledge makes recognition fast and leaves capacity for inference and integration (Perfetti & Stafura, 2014), and automaticity theory makes the same point about lower-level processes (LaBerge & Samuels, 1974). Readers at a lower level have less capacity to spare, and the usual prediction is that they would lean harder on external cues. A competing prediction comes from metacognitive accounts, which hold that skilled readers monitor their understanding and decide where to put attention (Baker & Brown, 1984); advanced second-language readers are generally reported to use such strategies more often

and more effectively (Grabe, 2009; Koda, 2005). Research with Iranian EFL learners points the same way: Ajideh et al. (2024) gave 111 intermediate-level undergraduates pre- and post-reading activities and reported gains in reading comprehension together with gains in self-regulation and reading motivation. The activities were instructional, not typographical, but the finding fits the idea that outside support works through the reader's own regulation of attention and effort. Applied to emphasis, a cue is only as useful as the reader's ability to interpret and use it. The two predictions point in opposite directions, so proficiency should be tested, not assumed.

Genre and proficiency may also act together. Narrative may let lower-proficiency readers rely on background knowledge and familiar structure, while expository and argumentative texts place heavier demands on vocabulary and discourse integration (Best et al., 2008; Graesser et al., 2011). A cue might, then, help most where an able reader meets a demanding genre. We know of no study that has tested this jointly with typographical emphasis.

### Learners' Own Accounts

Comprehension scores cannot show whether learners noticed the markers, how they interpreted them, whether they changed the way they read, or whether they found the markers helpful or distracting. Noticing is not understanding (Schmidt, 1990), and the signaling and cognitive load literatures warn that heavy marking can backfire (Mayer, 2021; Sweller et al., 2011). Learners' reports may, therefore, explain why a cue works for some readers and texts and not for others. They may also diverge from performance, as the typography research cited above suggests. Interview data are rare in this area, which is why the present study pairs them with an experimental measure.

### The Present Study

In sum, existing studies suggest that typographical emphasis can aid reading, but they do not show which readers benefit or in which kinds of text. The present study addresses two gaps. It tests whether proficiency, genre, and emphasis interact in their effects on the reading comprehension of Iranian EFL high school learners, and it asks how learners perceive emphasis in terms of noticing, usefulness, strategy use, preference, and distraction. A convergent design brings the two kinds of evidence together.

## 3. METHODS

### Design

This article reports two strands of a larger convergent parallel mixed-methods dissertation project (Creswell & Creswell, 2018). The quantitative strand used a 4 (emphasis: control, bold, underline, italic) × 4 (genre: argumentative, expository, descriptive, narrative) × 2 (proficiency: lower-intermediate, advanced) between-subjects factorial design ($N$ = 640; $n$ = 20 per cell). The dissertation also tested the overall effect of emphasis and compared the three markers with one another. Those analyses are reported in a companion article from the same project (Author, in preparation) and are mentioned here only where they help to interpret the interactions. The qualitative strand consisted of semi-structured interviews with 48 participants, two from each of the 24 non-control cells, and focused on learners' awareness of the markers, their perceived usefulness, preferences, strategy use, and difficulties. In outline, about 1,200 students took a placement test, 640 were assigned to conditions and read one passage, and 48 were later

interviewed. The two strands were collected in the same study period and brought together through a joint display.

## Participants

The target population was Iranian EFL learners aged 16 to 18 in Grades 10 to 12 of public high schools in Tabriz, the capital of East Azerbaijan Province in northwestern Iran. Tabriz was chosen for its demographic diversity and because the researcher had institutional access there. Public schools in the city follow the national curriculum set by the Ministry of Education, so learners' exposure to English instruction is fairly uniform. The schools were selected purposively, not at random. Five public high schools (three girls' and two boys' schools) were chosen because they were accessible, their administrations agreed to take part, and together they gave a reasonable picture of the city's public-school students. Within these schools, participants were recruited by convenience sampling: students whose parents consented and who were present in regular class periods formed the pool that took the placement test. Random assignment came only after proficiency screening, as described below.

Participants were between 16 and 18 years old (mean age ≈ 17.2 years), and their first language was Persian. They had received between 6 and 8 years of formal English instruction. Learners who had lived for more than 6 months in an English-speaking country, who had a diagnosed reading disability, or who had an uncorrected visual impairment were excluded.

Proficiency was determined with the Oxford Placement Test (OPT; Allan, 2004), described under Instruments, which was administered to approximately 1,200 students during regular class periods. Students scoring outside the two target bands were thanked and not included. Eligible participants ($N$ = 640; 320 per proficiency level) were then randomly assigned, within each proficiency level, to one of the 16 emphasis-by-genre conditions by means of a computer-generated randomization sequence stratified by gender, so that each cell contained approximately equal numbers of female and male participants where school composition allowed. Assignment was concealed from the participants and test administrators until just before testing. Table 1 summarizes the characteristics of the sample. Further details of the quantitative sample are provided in the companion article (Author, in preparation).

**Table 1: Participant Characteristics**

| Characteristic | Quantitative sample ($N$ = 640) | Interview subsample ($n$ = 48) |
|---|---|---|
| Age | 16–18 years (mean ≈ 17.2) | Drawn from the same population |
| Gender | Approximately equal numbers of female and male students | Approximately balanced |
| Academic level | Grades 10–12 in public high schools | Same as the full sample |
| English learning experience | 6–8 years of formal instruction | Same as the full sample |
| First language | Persian | Persian |
| Proficiency (OPT) | Lower-intermediate (scores 30–39), $n$ = 320; advanced (scores 48–54), $n$ = 320 | Lower-intermediate, $n$ = 24; advanced, $n$ = 24 |
| School context | Five public high schools (three girls' and two boys' schools) | Same schools |

The qualitative strand drew on the 480 participants in the 24 non-control cells (24 cells × 20 participants). Control-condition participants were not interviewed because they had not

encountered typographical emphasis and could not comment on it. From each non-control cell, two participants were selected at random, which gave 48 interviewees (24 lower-intermediate and 24 advanced learners, approximately balanced by gender). They were, therefore, drawn from the 480 learners, not from the full pool of 640. No sample-size formula was applied. The number follows from the design: two interviewees per cell gives equal coverage of every emphasis, genre, and proficiency condition, and it was judged to provide enough material for thematic analysis.

## Instruments

### *Oxford Placement Test*

The OPT is a standardized measure of general English proficiency. The version used has 60 multiple-choice items in two parts: Part 1 (questions 1–40) was taken by all students, and Part 2 (questions 41–60) was started only on the supervisor's instruction. The items are short notices, cloze passages, and sentence-completion questions on grammar and vocabulary, and there is no listening section. The test takes about 30 minutes. It was given in groups during regular class periods, with the same timing, instructions, and conditions in all five schools; proctors read the standardized instructions aloud, and the answer sheets were scored with the official key, one point per item, for a maximum score of 60. The scoring guidelines divide scores into six bands: 0–17 beginner (A1), 18–29 elementary (A2), 30–39 lower-intermediate (B1), 40–47 upper-intermediate (B2), 48–54 advanced (C1), and 55–60 very advanced (C2). Learners scoring 30–39 were classified as lower-intermediate and those scoring 48–54 as advanced; the other bands were excluded. Earlier validation work reports internal consistency of $\alpha$ = .89–.94 and test-retest reliability of $r$ = .91, and describes concurrent validity with other standardized proficiency tests as well established (Allan, 2004).

### *Reading Comprehension Tests*

Four base passages, one for each genre, were developed for the study. They were adapted from reading sections of the Iranian university entrance examination, Cambridge English examination materials at PET and FCE level, and authentic texts simplified for lower-intermediate readers. The sources were chosen to give passages similar to the reading tasks learners meet in school and in standard examinations, so that the passages would differ mainly in genre. Each genre followed a written specification: argumentative texts state a claim and support it with evidence and counterarguments; expository texts explain a topic through comparison, cause and effect, or problem and solution; descriptive texts depict a person, place, object, or event through sensory detail; and narrative texts tell a story with characters, plot, and setting. Each passage, 250–300 words long, was prepared in four typographic versions (control, bold, underline, and italic), which gave 16 passage versions.

Difficulty was controlled in two ways. The passages were of similar length, and their vocabulary was matched with the New General Service List (Browne et al., 2013): at least 90% of the words in every passage came from the first 1,000-word band and the rest from the second and third bands, and proper nouns and genre-specific terms were glossed. No readability formula was applied, so the passages were matched on length and word frequency, but not on syntactic complexity. The information to be emphasized was chosen by two applied linguists with expertise in teaching English as a foreign language (TEFL) who independently identified four to six information units in each passage, such as topic sentences, definitions, claims, and causal relationships, and resolved disagreements by discussion. The same units were marked in all three

emphasized versions, so the versions differed in the marker and in nothing else. Each passage was followed by ten multiple-choice questions (four literal, four inferential, and two on the main idea or text structure) adapted from standard examination formats. Each correct answer earned one point, so scores ranged from 0 to 10. Item construction and psychometric properties are reported in the companion article.

### *Interview Protocol*

The interview protocol had ten open-ended questions. It was developed from the literature on reading strategies and typographical enhancement (e.g., Norton, 2018) and from the research questions. The questions covered noticing, perceived helpfulness, preferences among markers, changes in reading strategy, and distraction. Examples asked whether participants noticed words or sentences that were presented differently, which type of emphasis they found most or least helpful, and whether any emphasis seemed distracting or unnecessary. Three experts (two TEFL professors and one applied linguist) reviewed the protocol for content validity, and it was then piloted with five EFL learners who did not take part in the main study.

## Procedure

The data were collected over ten weeks in the five schools. The first two weeks went to recruitment and permissions. In week 3 the OPT was given to all students whose parents had consented, in regular class periods, and took about half an hour. The reading sessions ran from week 4 to week 7, and assignment to conditions was concealed from participants and administrators until just before testing. Each group attended a single 60-minute session in a quiet classroom. The first five minutes were an orientation in which participants were told to read the passage once and answer the questions without outside help. The next 45 minutes were for reading the assigned passage, with or without typographical emphasis, and answering the ten questions, and the last ten minutes were a short break and debriefing. There was no teaching or feedback during the session; the manipulation consisted only of the typographic version of the passage each participant received.

Interviews took place in weeks 8 to 10. Immediately after the reading session, two participants from each of the 24 non-control cells were selected at random, with an eye to variety in gender and in test performance, and invited to an individual interview. Interviews were scheduled at the participants' convenience, usually in a free period or after school, and held in a private room. Each began with a few minutes of rapport building and an explanation of the purpose, with consent for audio recording, and then followed the protocol for about 15–20 minutes. Interviews were conducted in Persian. The recordings were transcribed verbatim in Persian and translated into English by a bilingual research assistant, and a second translator back-translated 20% of the transcripts to check the accuracy.

## Data Analysis

Scores were analyzed with a three-way between-subjects analysis of variance (ANOVA). No data were missing and no outliers were found. Shapiro–Wilk tests were significant in 31 of the 32 cells, which is expected for bounded integer scores from 0 to 10 with 20 observations per cell, and the analysis went ahead because the $F$ test is robust to such departures from normality (Tabachnick & Fidell, 2019). Levene's test was significant, $F(31, 608) = 1.76$, $p = .007$, so Games–Howell procedures were used for pairwise comparisons. Simple-effects analyses were used to interpret the significant genre × emphasis and proficiency × emphasis interactions. Because the three-way

interaction was not significant, the 32-cell results were treated descriptively, to identify conditions with comparatively large or small gains, and not as evidence of a distinct genre–proficiency combination.

The interview transcripts were analyzed with Braun and Clarke's (2006) six-phase approach to thematic analysis: familiarization, initial coding, searching for themes, reviewing themes, defining and naming themes, and writing the report. The approach was hybrid. Initial deductive codes came from the interview questions and the research questions, and other codes, such as distraction and habits from first-language reading, were developed from the interview material. Two coders independently coded 25% of the transcripts ($n$ = 12) and reached substantial agreement ($\kappa$ = .84); they resolved disagreements by discussion, and the lead researcher then coded the remaining transcripts. The two strands were compared through a joint display that placed the main interaction patterns beside the four qualitative themes and classified each relationship as convergence, complementarity, or dissonance (Creswell & Creswell, 2018).

### Trustworthiness and Rigor

The quantitative validity and reliability procedures are reported in the companion article, including internal, external, and construct validity, Kuder–Richardson Formula 20 (KR-20) reliability, test-retest reliability, and the manipulation check. For the qualitative strand, credibility was supported through expert review of the interview protocol and piloting. Dependability was addressed through the documented six-phase coding procedure and inter-coder agreement ($\kappa$ = .84). Back-translation checks and an audit trail linking themes to coded extracts were used to support confirmability.

### Ethical Considerations

Ethical approval was obtained from the researcher's university research ethics committee and from the Tabriz Department of Education, and school principals gave permission after the purpose, procedures, and confidentiality arrangements were explained. Parents or guardians gave written consent, and students gave assent. Participation in both the reading task and the interviews was voluntary, and students could decline or withdraw without penalty. Interview data were anonymized during transcription and stored securely.

## 4. RESULTS

### RQ1: Learners' Perceptions of Typographical Emphasis

The interviews involved 48 participants, two from each of the 24 non-control cells (24 advanced and 24 lower-intermediate learners). The transcripts were analyzed thematically, and four themes were identified: awareness and noticeability, perceived helpfulness and strategic reading, preference for boldface, and distraction and overuse. Percentages refer to the participants within a proficiency group (24 per group), and the quotations are translated from Persian.

#### *Awareness and Noticeability*

Nearly all advanced learners (96%, $n$ = 23) reported noticing the typographical markers, often describing bold words as immediately visible. As one advanced learner put it, "I saw the bold words immediately. They kind of popped out from the rest of the text." Among lower-intermediate learners, 67% ($n$ = 16) reported noticing the markers. Several of those who did not mention them explained that they were concentrating on understanding the basic meaning of the text and paid

little attention to its formatting, as one lower-intermediate participant put it: "I was trying to figure out what the sentences meant. I didn't really look at which words were different."

### *Perceived Helpfulness and Strategic Reading*

Advanced learners often described using the emphasized material deliberately. Some said that they read the bolded material first to establish a general idea before returning to the passage ("I read the bolded parts first to get the main idea, then I went back and read the whole paragraph more carefully"), while others used the emphasis as a reason to slow down and examine a sentence more carefully ("Underlining made me stop and think: 'This must be important.' So I paid more attention to those sentences."). In all, 83% of the advanced learners ($n$ = 20) explicitly said that the emphasis helped them read or understand the passage more efficiently. The lower-intermediate group was less consistent: 50% ($n$ = 12) described the markers as helpful, 38% ($n$ = 9) reported no effect or some distraction, and 13% ($n$ = 3) were uncertain. One lower-intermediate participant captured this ambivalence directly: "The bold words helped me know what the question was about. But sometimes there were too many bold words and I got confused."

### *Preference for Boldface*

Across the two proficiency groups, boldface received the strongest preference, followed by underlining and, then, italics. In all, 88% of the advanced learners ($n$ = 21) and 75% of the lower-intermediate learners ($n$ = 18) identified boldface as the most effective marker. The participants commonly referred to its visual contrast and ease of detection: "Bold is the best – you cannot miss it. Italic is too soft; it looks almost the same as regular text." Underlining was generally regarded as useful, although some learners found the line visually interfering with letters that extend below the baseline ("Underline is okay, but when there are many underlined words, the lines get messy"). Italics were more often described as easy to overlook, with one advanced learner admitting, "I don't think italic helps at all. I didn't even notice it until you asked me."

### *Distraction and Overuse*

Thirty-eight percent of the lower-intermediate learners ($n$ = 9) reported that emphasis could become unhelpful when it was used too densely. Their comments generally concerned visual clutter or uncertainty about what deserved attention when many items were marked, echoed in one participant's remark that "when every sentence has a bold word, it means nothing. I stopped believing that bold meant important." Advanced learners rarely described direct distraction, but several warned that emphasis should be used selectively. One participant summarized this concern by saying, "If you highlight everything, you highlight nothing."

Across the four themes, advanced learners more often reported noticing the markers (96%, $n$ = 23, against 67%, $n$ = 16, among lower-intermediate learners) and described using them to plan their reading. Distraction was reported more often by lower-intermediate learners, although advanced learners also warned that too much emphasis would defeat its purpose.

## RQ2: Interaction of Proficiency, Genre, and Typographical Emphasis

### *Omnibus Interaction Structure*

The three-way ANOVA produced significant genre × emphasis and proficiency × emphasis interactions, $F(9, 608) = 4.88$, $p < .001$, $\eta p^2 = .067$, and $F(3, 608) = 8.07$, $p < .001$, $\eta p^2 = .038$, respectively. The three-way genre × proficiency × emphasis interaction was not significant, $F(9,$

608) = 1.21, $p$ = .288, $\eta p^2$ = .018. Genre and proficiency each changed the size of the emphasis effect, but the data did not provide evidence that their effects combined into a statistically distinct three-way interaction. Appendix A gives the interaction terms in full, and the pattern is examined in more detail below.

By the usual benchmarks, the genre × emphasis interaction was medium in size and the proficiency × emphasis interaction fell between small and medium (Cohen, 1988). Put plainly, the benefit of emphasis was not the same in every genre or for advanced and lower-intermediate learners alike. That the three-way term was not significant means that the genre pattern did not differ reliably between the two proficiency levels. With 20 participants per cell, however, this test may have limited power to detect small differences, so the result should not be read as proof that no such interaction exists.

### *Pattern Across the 32 Conditions*

Table 2 summarizes the most and least effective conditions within each proficiency × genre combination, based on the simple-effects analyses reported in the companion article. Appendix B gives the corresponding tests.

**Table 2: Most Effective Emphasis Marker by Proficiency and Genre**

| Proficiency | Genre | Most effective | Least effective |
|---|---|---|---|
| Advanced | Argumentative | Bold / Underline (tied) | Control |
| Advanced | Expository | Bold / Underline (tied) | Control |
| Advanced | Descriptive | Bold / Underline (tied) | Control |
| Advanced | Narrative | None (ns) | — |
| Lower-intermediate | Argumentative | Bold | Italic / Control |
| Lower-intermediate | Expository | Bold / Underline (tied) | Control |
| Lower-intermediate | Descriptive | Bold | Italic |
| Lower-intermediate | Narrative | None (ns) | — |

For advanced learners, the emphasis affected comprehension in the argumentative, expository, and descriptive texts. The effect was the largest in expository texts, $F(3, 76) = 27.66$, $p < .001$, $\eta p^2 = .52$, where the bold condition scored 2.40 points above the control condition. Argumentative texts followed, $F(3, 76) = 17.40$, $p < .001$, $\eta p^2 = .41$, and then descriptive texts, $F(3, 76) = 13.08$, $p < .001$, $\eta p^2 = .34$. In narrative texts, advanced learners showed no significant effect of emphasis, $F(3, 76) = 0.96$, $p = .417$.

Lower-intermediate learners also showed significant emphasis effects in the argumentative, expository, and descriptive texts, but the effects were smaller than those of advanced learners (all $p$s < .01). Their narrative texts again produced no significant effect, $F(3, 76) = 1.67$, $p = .181$. The narrative gains were small across the marker conditions ($d$ = 0.33–0.67).

As Table 2 shows, bold, alone or tied with underline, was the most effective marker in every cell where emphasis had an effect, and the control or italic condition was the least effective. These descriptive differences are not evidence of a distinct genre–proficiency combination, because the three-way interaction was not significant.

### Integrating the Quantitative and Qualitative Findings

Table 3 presents a joint display connecting the quantitative moderation pattern to the qualitative themes.

**Table 3: Joint Display: Quantitative Moderation and Qualitative Themes**

| Quantitative pattern | Qualitative theme | Relationship |
|---|---|---|
| Advanced learners show larger emphasis benefit ($\eta p^2$ = .41 vs. .23) | Advanced learners report greater awareness and more deliberate strategic use | Convergence |
| Bold produces the largest, most consistent gain | Both groups independently rank bold as most effective | Convergence |
| No genre shows a reliable narrative benefit | Learners describe narrative texts as self-explanatory without markers | Convergence |
| Bold and underline are equivalent for advanced learners, but not for lower-intermediate learners | Lower-intermediate learners report underline as visually "messy"; advanced learners do not | Complementarity |
| Emphasis effect sizes are large, but not universal | Minority (mostly lower-intermediate) report emphasis as distracting when dense | Complementarity |

The joint display shows a close match between the two strands. Advanced learners showed the larger quantitative effect and described the more deliberate use of the markers; boldface was among the most effective markers in Table 2 and the most common preference in both groups; and the lack of a reliable narrative benefit agreed with the learners' descriptions of narratives as needing little extra visual guidance. The interviews also added something the scores could not show: some learners, mostly lower-intermediate, found dense emphasis distracting.

## 5. DISCUSSION AND CONCLUSION

### Overview

This study asked whether typographical emphasis works the same way for every reader and every text, and how learners themselves experience it. Genre and proficiency each changed how much emphasis helped, but the data gave no reliable sign that their effects combine into a distinct pattern. Advanced learners reading expository and argumentative texts gained the most, narratives gained little, and in the interviews, learners described noticing the markers, using them to plan their reading, preferring boldface, and finding heavy emphasis distracting.

### Genre and Proficiency as Separate Moderators

The two significant interactions fit a view in which the emphasis is worth the most when the text needs guidance and the reader can use it. Genre is the clearer part of the story. Narratives carry their own organization in characters, goals, and events in time, so readers get structural support from the text itself (Best et al., 2008; Graesser et al., 2011), and an added cue has little left to do. Expository and argumentative texts leave more of the organizing to the reader, and that is where marking the central concept or claim gave the advanced learners their largest gains. Descriptive texts fell between the two.

Proficiency worked in a less obvious way. A simple capacity account would predict that lower-level readers gain more, since they have less spare attention (LaBerge & Samuels, 1974; Perfetti & Stafura, 2014). The results ran the other way. The interviews help to explain why: the advanced learners were the ones who described using the markers deliberately, and the benefit of a cue depends on the reader's ability to interpret it as well as on whether it is seen (Baker & Brown, 1984; Grabe, 2009). Findings on proficiency in the enhancement literature, such as those of Lee (2021) on noticing, do not translate directly to the comprehension of connected text, where the demands differ.

The three-way interaction was not significant, so the 32-cell pattern cannot be used to argue that one combination of genre and level is special. The cells were modest in size ($n$ = 20), and a test of a three-way term in such a design may miss small effects. The safest reading is that genre and proficiency each moderate the effect of the emphasis, and the largest and smallest gains in Table 2 show the directions of those two effects.

## What the Interviews Add

Test scores show that emphasis can matter. The interviews show how learners met it. Most learners in both groups reported noticing the markers, but noticing is not the same as understanding: a learner may see that a word is bold without knowing why, and the enhancement literature has repeatedly found that attention does not guarantee learning (Lee & Huang, 2008). What seemed to matter, judging from the interviews, was whether learners treated the marked words as a guide, going to them first and, then, returning to the rest of the text. That matches metacognitive accounts of reading, in which skilled readers regulate attention according to their goals (Baker & Brown, 1984; Grabe, 2009), and it suggests that emphasis acts as an external scaffold whose value depends on the strategy the reader brings to it.

The preference for boldface was shared across the two proficiency groups, and boldface also appears among the most effective conditions in Table 2. Typography research has cautioned against assuming that what readers like is what helps them (Bernard et al., 2003), so the match deserves mention, though it does not show that preference produced the gains. A plausible reason for both is perceptual clarity, since bold changes the weight of a word without distorting its shape. That remains an explanation and not a tested one.

The theme with the clearest practical meaning was overuse. Lower-intermediate learners reported distraction more often, and advanced learners also saw that too much emphasis would defeat its purpose. This agrees with cognitive load theory and with the signaling literature, in which a cue helps only while it stays selective (Mayer, 2021; Sweller et al., 2011). If every sentence carries a marked word, the contrast that gives the cue its meaning is gone.

## Bringing the Two Strands Together

Read together, the strands converge. Advanced learners showed the larger gains and described the more deliberate use; boldface was both among the most effective markers and the most preferred; narratives showed little gain; and heavy emphasis appeared in the interviews as a source of distraction. These points suggest a sequence in which visual distinction makes a word more likely to be noticed, noticing directs selective attention, and strategic processing of the marked information then feeds into comprehension. The sequence is an interpretation grounded in theory, not an observed mechanism. The study had no eye tracking, reading-time measure, or direct measure of cognitive load, so it shows an association between typographical conditions and comprehension, while the steps in between remain plausible, but untested.

The results sit comfortably with findings that signaling effects are modest and conditional (Richter et al., 2016; Schneider et al., 2018) and that textual enhancement research is inconsistent when conditions vary (Lee & Huang, 2008; Rassaei & Jabbarpoor, 2025). The learners' insistence on selective marking also echoes recent work on highlighting, in which the quality of what was marked mattered more than the act of marking (Mason et al., 2024). The study adds two things. It moves the question from single-form noticing, as in recent gaze-based enhancement work (Jung et al., 2025), to the comprehension of connected texts, and it shows that learners' own reports can help explain why a cue helps one group and not another. Work on text signals with second-language readers (Kim, 2025; Kim & Clariana, 2017) and on reading support among Iranian learners (Ajideh et al., 2024) points the same way.

## Theoretical Implications

For signaling and cognitive load theories, the results suggest that the cue and the reader cannot be separated. A signal does not carry its own effect; the effect arises from how perceptual prominence meets a text's organization and a reader's strategies. For the Noticing Hypothesis, the findings support the view that noticing is a necessary step, but not a sufficient one, for comprehension, and that what learners do after noticing deserves as much study as the noticing itself (Schmidt, 1990). More generally, the data favor treating genre and proficiency as separate sources of variation in the value of a cue, not as one combined condition.

## Pedagogical Implications

Teachers and materials writers can draw several suggestions from the study, with the caution that it examined short printed passages. Emphasis should mark conceptual or discourse importance, such as a definition, a claim, or a cause, and not difficulty or visual interest. It should be used sparingly, because the contrast between marked and unmarked words is what carries the signal. It can probably be reduced in narratives, where the text already provides organization. Lower-intermediate learners may need help in reading the cue, and a few questions, such as why this idea has been marked and what role it plays in the paragraph, would turn a formatting feature into a reading strategy.

## Limitations

Several limits apply. The participants were high school learners from one city in Iran at two proficiency levels, so the results may not extend to other levels, ages, or educational systems. Only three markers and four genres were studied, in short printed passages with a multiple-choice measure of immediate comprehension, and digital reading, long-term retention, and reading speed were not examined. The passages were matched on length and word frequency, but not on a readability index, so some differences in difficulty between genres may remain. With 20 participants per cell, power to detect small three-way effects was limited. The interviews took place after the reading task, so learners' accounts may differ from what they actually did while reading. The interview data were also translated from Persian, although back-translation checks were used to limit error.

## Recommendations for Future Research

Future studies could combine typographical manipulation with eye tracking or reading-time measures to test the sequence proposed above, and add think-aloud or stimulated recall to see how learners use marked words. Longitudinal designs could ask whether repeated exposure to

selectively emphasized texts lets learners find important information more independently. Larger samples, more proficiency levels, and more genres would allow the joint role of genre and proficiency to be tested with adequate power, and training studies could examine whether teaching learners to interpret emphasis changes the results for lower-intermediate readers.

## Conclusion

Typographical emphasis is not a uniform aid. In this study its value depended on the genre of the text and the proficiency of the reader, and learners' own accounts suggested why: they used the cue when they could read its purpose and found it distracting when it was overused. The practical message is a modest one. Mark what is important, mark little of it, and expect the cue to matter most in texts that leave readers to find the structure for themselves.

## References

Ajideh, P., Zohrabi, M., & Khojand, B. (2024). Exploring the effectiveness of pre- and post-reading activities on the development of reading motivation and self-regulation as essentials for reading comprehension. *Journal of Modern Research in English Language Studies, 11*(3), 25–49. https://doi.org/10.30479/jmrels.2024.19503.2277

Allan, D. (2004). *Oxford placement test 2: Test pack*. Oxford University Press.

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (Vol. 1, pp. 353–394). Longman.

Bernard, M. L., Chaparro, B. S., Mills, M. M., & Halcomb, C. G. (2003). Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text. *International Journal of Human-Computer Studies, 59*(6), 823–835. https://doi.org/10.1016/S1071-5819(03)00121-6

Best, R. M., Floyd, R. G., & McNamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology, 29*(2), 137–164. https://doi.org/10.1080/02702710801963951

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bringhurst, R. (2013). *The elements of typographic style* (4th ed.). Hartley & Marks.

Browne, C., Culligan, B., & Phillips, J. (2013). *The new general service list*. http://www.newgeneralservicelist.org

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dyson, M. C. (2004). How physical text layout affects reading from screen. *Behaviour & Information Technology, 23*(6), 377–393. https://doi.org/10.1080/01449290410001715714

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press. https://doi.org/10.1017/CBO9781139150484

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher, 40*(5), 223–234. https://doi.org/10.3102/0013189X11413260

Hyland, K. (2004). *Genre and second language writing*. University of Michigan Press.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194–203. https://doi.org/10.1038/35058500

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning, 64*(1), 160–212. https://doi.org/10.1111/lang.12034

Jung, J., Révész, A., Stainer, M. J., Pellicer-Sánchez, A., Chung, Y., & Shi, D. (2025). The impact of gaze-contingent textual enhancement on L2 collocation learning from computer-mediated reading tasks. *TESOL Quarterly*. Advance online publication. https://doi.org/10.1002/tesq.3404

Kim, K. (2025). How the relationship between verbal and nonverbal text signals influences second language expository text comprehension: Knowledge structure analysis. *Educational Technology Research and Development, 73*, 15–34. https://doi.org/10.1007/s11423-024-10405-6

Kim, K., & Clariana, R. B. (2017). Text signals influence second language expository text comprehension: Knowledge structure analysis. *Educational Technology Research and Development, 65*, 909–930. https://doi.org/10.1007/s11423-016-9494-x

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293–323. https://doi.org/10.1016/0010-0285(74)90015-2

Lee, B. J. (2021). The effects of proficiency and textual enhancement technique on noticing. *System, 96*, Article 102407. https://doi.org/10.1016/j.system.2020.102407

Lee, M., & Révész, A. (2018). Promoting grammatical development through textually enhanced captions: An eye-tracking study. *The Modern Language Journal, 102*(3), 557–577. https://doi.org/10.1111/modl.12503

Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning, 57*(1), 87–118. https://doi.org/10.1111/j.1467-9922.2007.00400.x

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*(3), 307–331. https://doi.org/10.1017/S0272263108080479

Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review, 1*(3), 209–234. https://doi.org/10.1007/BF01320135

Lorch, R. F., & Lorch, E. P. (1995). Effects of organizational signals on text-processing strategies. *Journal of Educational Psychology, 87*(4), 537–544. https://doi.org/10.1037/0022-0663.87.4.537

Martin, J. R., & Rose, D. (2008). *Genre relations: Mapping culture*. Equinox Publishing.

Mason, L., Ronconi, A., Carretti, B., Nardin, S., & Tarchi, C. (2024). Highlighting and highlighted information in text comprehension and learning from digital reading. *Journal of Computer Assisted Learning, 40*(2), 637–653. https://doi.org/10.1111/jcal.12903

Mayer, R. E. (2021). *Multimedia learning* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781108894333

Norton, C. W. (2018). *Typographic emphasis and contrastive focus: An eye tracking study* [Doctoral dissertation, University of Leeds]. White Rose eTheses Online. https://etheses.whiterose.ac.uk/22135/

Perfetti, C. A., & Stafura, J. Z. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*(1), 22–37.

https://doi.org/10.1080/10888438.2013.827687
Ponce, H. R., Mayer, R. E., & Méndez, E. E. (2022). Effects of learner-generated highlighting and instructor-provided highlighting on learning from text: A meta-analysis. *Educational Psychology Review, 34*(2), 989–1024. https://doi.org/10.1007/s10648-021-09654-1
Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*(1), 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325
Rassaei, E., & Jabbarpoor, S. (2025). Effects of textual enhancement on L2 development: A meta-analysis. *International Review of Applied Linguistics in Language Teaching*. Advance online publication. https://doi.org/10.1515/iral-2025-0118
Rayner, K., Schotter, E. R., Masson, M. E. J., Potter, M. C., & Treiman, R. (2016). So much to read, so little time: How do we read, and can speed reading help? *Psychological Science in the Public Interest, 17*(1), 4–34. https://doi.org/10.1177/1529100615623267
Richter, J., Scheiter, K., & Eitel, A. (2016). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educational Research Review, 17*, 19–36. https://doi.org/10.1016/j.edurev.2015.12.003
Saadatnia, M., Ketabi, S., & Tavakoli, M. (2017). Levels of reading comprehension across text types: A comparison of literal and inferential comprehension of expository and narrative texts in Iranian EFL learners. *Journal of Psycholinguistic Research, 46*, 1087–1099. https://doi.org/10.1007/s10936-017-9481-3
Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129
Schneider, S., Beege, M., Nebel, S., & Rey, G. D. (2018). A meta-analysis of how signaling affects learning with media. *Educational Research Review, 23*, 1–24. https://doi.org/10.1016/j.edurev.2017.11.001
Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*(2), 165–179. https://doi.org/10.1017/S0272263100011943
Simard, D. (2009). Differential effects of textual enhancement formats on intake. *System, 37*(1), 124–135. https://doi.org/10.1016/j.system.2008.06.003
Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer. https://doi.org/10.1007/978-1-4419-8126-4
Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5
Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review, 28*(4), 1060–1092. https://doi.org/10.3758/s13423-020-01859-9
Wu, Y., Luo, C., Wang, Z., Xie, H., Huang, Y., & Su, Y. (2024). A further specification of the effects of font emphasis on reading comprehension: Evidence from event-related potentials and neural oscillations. *Memory & Cognition, 52*, 225–239. https://doi.org/10.3758/s13421-023-01457-9

# Appendices

## *Appendix A: Full Interaction Results for the Three-Way ANOVA*

**Table A1: Interaction Terms from the Three-Way Between-Subjects ANOVA**

| Source | SS | *df* | MS | *F* | *p* | ηp² |
|---|---|---|---|---|---|---|
| Proficiency × Emphasis | 17.92 | 3 | 5.97 | 8.07 | < .001 | .038 |
| Genre × Emphasis | 32.49 | 9 | 3.61 | 4.88 | < .001 | .067 |
| Proficiency × Genre × Emphasis | 8.04 | 9 | 0.89 | 1.21 | .288 | .018 |
| Error | 450.30 | 608 | 0.74 | | | |

***Note.*** **SS = sum of squares; MS = mean square; ηp² = partial eta squared. The main effects and the proficiency × genre interaction answer other research questions of the dissertation and are not reported in this article.**

## *Appendix B: Simple-Effects Tests of Emphasis*

**Table B1: Simple Effects of Emphasis Within Each Proficiency and Genre Combination**

| Proficiency | Genre | *F* (3, 76) | *p* | ηp² |
|---|---|---|---|---|
| Advanced | Argumentative | 17.40 | < .001 | .41 |
| Advanced | Expository | 27.66 | < .001 | .52 |
| Advanced | Descriptive | 13.08 | < .001 | .34 |
| Advanced | Narrative | 0.96 | .417 | .04 |
| Lower-intermediate | Argumentative | 7.43 | < .001 | .23 |
| Lower-intermediate | Expository | 8.89 | < .001 | .26 |
| Lower-intermediate | Descriptive | 14.28 | < .001 | .36 |
| Lower-intermediate | Narrative | 1.67 | .181 | .06 |

***Note.*** **Each row is a one-way test of emphasis (four conditions, *n* = 20 per condition) within one proficiency and genre combination. For lower-intermediate learners, the effects in the argumentative, expository, and descriptive texts were significant at *p* < .001.**