

Assessing Nation's Four Strands Framework: A Critical Content Analysis of Iranian High School English Textbooks

¹Alireza Karbalaei*

²Maryam Aghaalian

Research Paper

IJEAP-2602-2190

Received: 2026-02-26

Accepted: 2026-05-27

Published: 2026-09-30

Abstract: The quality of textbooks significantly influences the effectiveness of teaching and learning in language education. During 2013, the Iranian Ministry of Education proposed a new series of English textbooks for high school students including Prospect and Vision, both developed based on the principles of Communicative Language Teaching (CLT). The present study aims to evaluate the aforementioned textbooks through the lens of Nation's Four Strands framework which advocates for a balanced language curriculum incorporating components such as meaning-focused input, meaning-focused output, language-focused learning, and fluency development. To this aim, a mixed-methods design was used by random sampling and multiple data collection instruments including a structured evaluation checklist aligned with Nation's framework, as well as content analysis of textbook tasks and activities. Quantitative data were collected through checklists completed by 22 Iranian English as a Foreign Language (EFL) teachers and analyzed utilizing SPSS. In addition, qualitative data were extracted from systematic content analysis of the textbooks. The results revealed an uneven distribution of the four strands across the textbooks, with insufficient attention to language-focused learning and fluency development, despite a relatively balanced allocation of activities. Accordingly, such textbooks fail to present a fully balanced curriculum as defined by Nation's framework, despite their potential to serve as primary instructional resources. Thus, teachers should adapt and supplement the materials with additional resources tailored to their learners' needs and classroom contexts.

Keywords: Fluency Development, Language-focused Learning, Meaning-focused Input, Meaning-focused Output, Textbook Evaluation

Introduction

Textbooks serve as fundamental pillars for effective teaching within K-12 English as a Foreign Language (EFL) contexts. As the primary instructional resource in many EFL classrooms, textbooks shape lesson structure, learning activities, and the language input provided to learners. The strategic selection and thoughtful integration of these materials significantly enhance the language acquisition experience, resulting in more engaging and effective learning environments (Tomlinson, 2011, as cited in Kalfut, 2024). Language learning, as a multifaceted process, encompasses several critical components including teachers' pedagogical expertise, learners' cognitive and linguistic abilities, sociocultural context of instruction, as well as the appropriateness and quality of textbooks. The strategic selection and thoughtful integration of these instruments significantly enhance the language acquisition experience, resulting in fostering more engaging and effective learning environments (Tomlinson, 2011, as cited in Kalfut, 2024). Language learning, as a multifaceted process encompasses several critical components including teachers' pedagogical expertise, learners' cognitive and linguistic

¹Assistant Professor of TEFL (Corresponding Author), Email: Karbalaei2008@gmail.com; Department of English Language Teaching, Farhangian University, Tehran, Iran.

²MA in TEFL, Email: maryamaghaaliyan@ymail.com; Department of English Language Teaching, Farhangian University, Tehran, Iran.

abilities, sociocultural context of instruction, as well as the appropriateness and quality of applied instructional materials.

Textbook have traditionally served as the primary teaching resources in the majority of English Language (EL) classrooms, especially in EFL contexts. Richards (1998) substantiated this assertion, noting that many educators rely heavily on textbooks to guide instructional content and classroom activities, particularly in EFL contexts where textbooks often structure the curriculum. Such a phenomenon manifests a broader pedagogical reality where coursebooks, developed by subject matter specialists, are viewed by educators as facilitators of effective teaching and learning processes. Therefore, the majority of teachers concentrate their instructional efforts on such materials which determine language learning objectives, as well as functioning as comprehensive lesson plans for students and educators (Karbalaie, 2025).

The institutional significance of language textbooks extends beyond individual classrooms to encompass entire national educational systems. Cunningsworth (1995) identified the multifunctional role of textbooks as efficient repositories for sources of pedagogical opinions and learning activities, material exposure, and reference materials which support student learning while bolstering the confidence of inexperienced educators. Students extract the majority of their linguistic input from these systematically designed materials, extending beyond direct teacher-learner interactions. However, textbook implementation varies significantly among the educators. Although teachers may adapt or supplement textbook materials to accommodate specific learner needs, these resources still largely determine the structure and range of learning opportunities available to students, particularly in centralized educational systems. While experienced educators with subject matter confidence may employ textbooks as collections of topics to select proper content, less experienced teachers often rely on textbooks to guide their instructional delivery, highlighting the critical significance of textbook appropriateness and quality.

Implementing textbooks requires careful consideration of inherent deficiencies, despite their acknowledged significance in language education. In contemporary language pedagogy at all of the educational levels, selecting materials which optimally address learners' specific needs has emerged as critically essential. The challenge of identifying exceptional textbook resources remains even when these materials are developed by educational professionals, underscoring the complexity of matching pedagogical resources with diverse learner requirements (Cunningsworth, 1995). Comprehensive textbook evaluation processes should identify emphasized linguistic aspects, determine curricular alignment, and examine whether materials meet students' expectations and needs (Ahmadi & Derakhshan, 2016).

Effective language instruction requires course designers and educators to possess comprehensive understanding of fundamental teaching and learning principles, using such methodologies to align with instructional conditions, educator competencies, and learner characteristics. Various teaching and learning activities can be implemented based on the above-mentioned foundational principles, ensuring pedagogical flexibility while maintaining educational rigor (Nation, 1996). Such foundational understanding proves indispensable while investigating whether instructional materials create balanced opportunities for language development.

Recent developments in language pedagogy have witnessed the integration of Nation's strands of language learning into contemporary syllabus-design standards, providing critical support to course developers engaged in teaching English as a second or foreign language (ESL/EFL). Such standards facilitate the construction of well-balanced language courses which address the comprehensive needs of language learners. Nation (2009) emphasized the significance of incorporating all of the strands of learning including meaning-focused input, meaning-focused output, language-focused learning, and fluency development. He advocated for equitable time allocation across these domains, claiming that "equal time should be devoted to each of the strands" to ensure holistic language development.

Thus, a well-structured foreign or second language course should create approximately equivalent learning opportunities through the following strands.

- *Meaning-focused input*: Language acquisition through listening and reading comprehension activities
- *Meaning-focused output*: Language development through speaking and writing production
- *Language-focused learning*: Systematic acquisition through purposeful attention to linguistic features and structures
- *Fluency development*: Mastery through advanced performance levels in all of the skill areas, exceeding conventional competency benchmarks (Nation, 2003).

The aforementioned framework establishes the theoretical foundation for addressing the effectiveness of instructional resources in contemporary EFL contexts, especially regarding their capacity to discuss all of the essential learning dimensions while accommodating diverse learner needs and pedagogical requirements. The remaining sections of this study focus on existing empirical research on textbook evaluation methodologies by assessing the alignment of current Iranian EFL textbooks with Nation's Four Strands framework. This study presents the results by systematic evaluation of widely used textbooks, examines the implications of imbalances recognized within these materials, and provides evidence-based recommendations for optimizing resource selection and development in K-12 EFL environments. The present study investigates a significant research gap concerning the practical application of theoretically sound pedagogical frameworks to real-world instructional materials, especially within specific national educational contexts.

Literature Review

Theoretical Foundations: Evolution from Methods to Principles

The landscape of second language acquisition (SLA) has undergone a significant paradigmatic shift, moving away from rigid method-based approaches toward flexible, principle-driven frameworks. Nation (2018) notes that historical comparisons of teaching methods often failed due to misinterpretation, prompting a focus on underlying principles that offer adaptability and continuity. Within this paradigm, the Four Strands framework emerged as a guiding model for time allocation in language courses, advocating for a balanced integration of meaning-focused input, meaning-focused output, language-focused learning, and fluency development.

While this framework provides a valuable heuristic for curriculum design, it is not without limitations. Nation (2007) himself acknowledges that the 25% time allocation is somewhat arbitrary, explicitly noting that “the balance may need to be adjusted according to learner proficiency, available instructional time, and specific curricular goals” (p. 2). In practice, this ideal frequently collides with structural constraints in EFL settings with limited weekly contact hours, where high-stakes examination pressures often push instructors to prioritize language-focused learning over fluency or meaning-focused output (Arab, 2020; Yaghoubinejad et al., 2021). Moreover, the four strands rarely operate in isolation during actual classroom instruction; they naturally overlap and scaffold one another—for example, vocabulary instruction embedded within extensive reading tasks, or grammar practice integrated into communicative role-plays—making rigid proportional enforcement pedagogically impractical. Consequently, rather than treating the framework as a prescriptive universal standard, it must be critically adapted to local educational realities, shifting the evaluative focus from strict strand compliance to contextual flexibility and strategic integration. To situate these principles within their scholarly development, the evolution of the framework followed a clear chronological trajectory. Nation (1996) initially proposed the foundational concepts, which were subsequently elaborated and formalized in 2007 (Nation, 2007). These ideas were later integrated into comprehensive curriculum design guidelines (Nation & Macalister, 2010), with subsequent reflections further refining their practical application (Nation, 2018). This historical progression sets the stage for a deeper examination of how the framework was conceptualized and operationalized in practice.

Genesis and Conceptualization of the Four Strands Framework

Nation's work on the Four Strands framework draws from his extensive teaching experience in the mid-1990s (Nation, 1996) and was formally elaborated in 2007 (Nation, 2007). This conceptual breakthrough represented a constructive idea for curriculum design, providing concrete implications for allocation of time to specific techniques and activities within a language course. The Four Strands framework can integrate traditionally unfashionable yet beneficial activities such as rote learning through grammar-translation, substitution tables, and word cards, provided that they occupy a constrained proportion of course time. Such inclusivity ensures that learners' and teachers' diverse hopes can be accommodated without disrupting the overall balance of learning opportunities. Nation (2018) attributed the framework development to his broad research interests spanning vocabulary learning, speaking activities, extensive reading, and fluency development, indicating that any viable framework should incorporate these multifaceted foci. The theoretical underpinnings of the Four Strands draw heavily from established SLA research. Further, Nation (2000) believed that effective material design involves utilizing pedagogical principles, recognizing learner needs, and investigating the learning environment (Widiatmoko, 2014). This trajectory focuses on applying these design principles to textbook development, resulting in incorporating the four strands. The framework acknowledges the significance of meaning-focused input for language skill acquisition (Ellis, 2012; Renandya, 2013; Snow, 2021), essential role of meaning-focused output (Swain, 1995; Salaberry, 1997; Izumi, 2002; Toth, 2006; Saarinen, 2024), and practice of integrated skills teaching. Nation (2007) developed this language teaching framework with his expertise in applied linguistics and pedagogical methodology to identify key areas where language teachers should focus their efforts. The Four Strands framework has since become the cornerstone of his theoretical contributions (Nation & Yamamoto, 2012).

Detailed Explanation of the Four Strands

The Four Strands framework advocates for a balanced approach to language learning, suggesting that equal significance (approximately 25% of class time) should be devoted to each of the four strands. This balanced allocation seeks to enable teachers to develop well-rounded curricula which foster flexible and effective language skill development (Nation, 2007). However, Nation (2007) found that imbalances persist in practice. For example, Indonesian teachers often overemphasize language-focused learning, dedicating 70-80% of course time to this strand (Iswandari & Paradita, 2019). The core components of the framework are as follows.

Meaning-focused input: This strand involves learning through listening and reading, where the learner's primary attention is directed toward comprehending the messages and ideas conveyed by the language. Activities include shared and extensive reading, watching media, engaging in conversations, and listening to stories (Hinkel, 2006). Effective meaning-focused input requires that learners know a significant portion of the content, show a strong motivation to understand the input, encounter only a few unfamiliar language features (ideally 95-98% known vocabulary or 1-2 unknown words per hundred words) (Hu & Nation, 2000), can employ background clues to comprehend unfamiliar elements, and process large quantities of input (Nation, 2008). Gregg and Krashen (1986) supported this strand by his Comprehensible Input Hypothesis which emphasizes that optimal language acquisition occurs when learners encounter challenging yet comprehensible language. Furthermore, Nation (2007) reported that texts should contain no more than one unknown word per fifty words to maintain focus on meaning rather than decoding

Meaning-focused output: This strand emphasizes learning through writing and speaking, where learners focus on transmitting messages and ideas to others. Meaning-focused output, which is rooted in Communicative Language Teaching (CLT), encompasses activities such as letter and essay writing, storytelling, giving instructions, and diary keeping (Nation, 2007). The approach is premised on the idea that active language use in communicative tasks fosters grammatical and lexical knowledge, leading to more fluent and precise language use (Nation & Yamamoto, 2012). Key provisions include encountering minimal unknown language, benefiting from extensive opportunities to write and speak,

focusing on message transfer, speaking and writing about familiar topics, and using communication strategies to compensate for knowledge gaps (Nation, 2009). Swain (1985, 1995) theoretically supported this strand by his Output Hypothesis which identifies three key functions including metalinguistic reflection, noticing knowledge gaps, and hypothesis testing through language production (Widiatmoko, 2014; Swain & Lapkin, 1998).

Language-focused learning: This strand involves direct attention to language features and items including strategy practice, grammar exercises, vocabulary study, discourse analysis, and phonological instruction (Long, 2017; Ellis, 1990). Language-focused learning should contain no more than 25% of course time, despite its common teacher-centered orientation. This approach facilitates noticing language features within meaningful contexts through practices such as spaced repetition and intensive writing/reading. Instructional provisions incorporate focus on relevant and simple features, repeated exposure, intentional feature-noticing tasks, sustained analytic engagement, and integration of acquired competencies into the other three strands (Nation & Newton, 2009). Research advocates the efficacy of intentional vocabulary learning for lasting acquisition (Nation, 2001; Boers et al., 2006) and focused grammar instruction (Williams, 2005) although scholarly discourse persists regarding whether focus on form yields optimal results within or outside a meaning-focused context (Ellis, 2006).

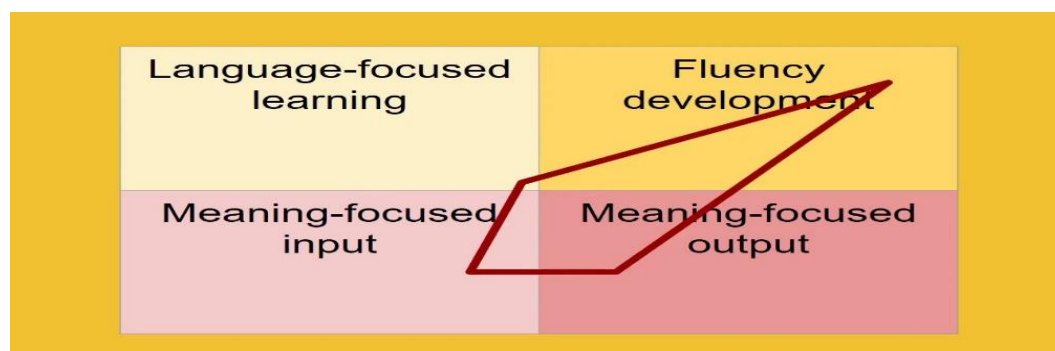
Fluency development: This strand seeks to develop the ability to utilize language effortlessly across all of the four skills. Fluency development emphasizes naturalness and accuracy over mere speed, reinforcing known content through meaning-focused tasks. Techniques include skimming, listening to familiar content, as well as speed and repeated reading. The time-on-task principle underpins this strand, claiming that enhanced deliberate practice improves performance (Nation & Yamamoto, 2012). Conditions for effective fluency development include benefitting from abundant input/output opportunities, operating at a faster-than-normal pace, focusing on meaning transfer, and applying known content (Nation & Newton, 2009). Experts recommend dedicating 25% of course time to fluency activities to create confidence and proficiency (Nation, 2008).

Balancing the Four Strands

Obtaining a balanced curriculum necessitates the proportional distribution of learning time among the four strands. Direct study of language items (language-focused learning) should not surpass 25% of total time, while at least 25% should be dedicated to fluency development. A well-structured curriculum allocates 75% of time to the three meaning-focused strands, ensuring holistic skill development (Karbalaie, 2025). Nation (2018) reported that effective teaching involves more than classroom delivery. Crucial tasks include planning courses which balance learning opportunities across the four strands and organizing classrooms to ensure equitable participation. This balanced approach plays a critical role in fostering comprehensive language proficiency.

Figure 1

Four Strands and Their Relationship



Textbook Evaluation: Bridging Theory and Practice

Textbooks play a crucial role in language teaching, serving as critical resources for learners and educators (Harwood, 2005). Effective evaluation proves indispensable for pinpointing strengths and weaknesses, enabling teachers to optimize learning outcomes. Nation and Macalister (2010) indicated that evaluation determines curriculum design quality, with success hinging on examining cost-effectiveness, teacher perceptions, as well as student satisfaction and learning achievements. In the Iranian EFL context, textbook evaluation plays a key role in shaping high school students' language proficiency. A large number of studies have been conducted on the strengths and weaknesses of Iranian English textbooks.

Early evaluations of CLT-based textbooks in Iran revealed mixed results. Ebrahimi and Sahragard (2017) reported positive attitudes toward CLT-designed textbooks in Isfahan while documenting demands for enhanced quality, emphasizing the impact of teacher proficiency and planning. However, subsequent studies represented more significant shortcomings. Others (e.g. Kaffash, Yazdanmehr, & Ghanizadeh, 2018) studied Prospect 3 employing Cisar's scale, emphasizing its failure to achieve the purported communicative objectives, as reported by students and teachers. Some others (e.g., Goodarzi, Weisi, & Yousofi, 2020) identified critical deficiencies in the Prospect series, despite its improvements over predecessors, especially in design, activities, and methodology. In another study, Goodarzi (2020) uncovered significant deficiencies in Prospect series organization, skill integration, and alignment with CLT principles, especially regarding Communicative, Cognitive, and Creative (CCC) aspects utilizing Ellis' dual framework.

This systemic imbalance is not isolated to specific high school materials; broader research on English education in Iran reveals a persistent overreliance on traditional, form-focused methodologies and instructional resources that fail to systematically address learners' diverse communicative needs (Ranjbar & Soodmand Afshar, 2021). Consequently, evaluating whether newer textbook series actually achieve a balanced curriculum remains a critical imperative. However, some studies found more positive alignments. Arab (2020) highlighted a balanced approach applying the Four Strands framework to Prospect 3 while teachers followed the manual, supporting its alignment with CLT and its contribution to fluency development. In addition, Sadeghzadeh (2021) argued that teachers' positive evaluations regarding Vision textbooks in terms of Aims, Design, and Language Content contrasted with weaker ratings in Skills, Topics, and Methodology.

Research on the Vision series presents a complex picture. Further, Derakhshan (2021) critiqued Vision 1 for its constrained opportunities to foster cultural awareness applying Peircean semiotic theory. Furthermore, Kheradmehr (2022) asserted that all of the Vision textbooks demonstrated low standards across the five Cs (ACTFL), while Vision 3 was slightly preferred. According to Sadrabadi Haqiqi (2023), international textbooks such as Interchange 2 and Top Notch 2 involved more international cultural items than Vision 2. In another study, Jamshidian et al. (2024) declared that the Vision series inadequately presented grammatical knowledge and overemphasized intonation, hindering CLT effectiveness. Additionally, Baqeri (2024) found that Family and Friends One outperformed Prospect One in multiple areas.

Statement of Problem

Critical elements of CLT including skill integration and strategy instruction remain frequently overlooked, despite numerous studies on textbook evaluation in Iran (Goodarzi, 2020; Jamshidian et al., 2024). An excessive focus on Iranian cultural contexts can constrain the sociocultural dimensions inherent in CLT (Kheradmehr, 2022; Derakhshan, 2021). Various evaluation frameworks have been used, encompassing Cunningsworth's (1995) model, Nation and Macalister's (2010) Four Strand framework, ACTFL five Cs, and Ellis' dual framework. While Arab (2020) examined Prospect 3 applying the Four Strands framework and highlighted a balanced approach, no study has yet conducted a systematic multi-book analysis across both the Prospect and Vision series using detailed quantitative content analysis and teacher perceptions. The gap emerges as pivotal given Nation's (2007, 2009)

contention that balanced attention to all of the strands constitutes a fundamental prerequisite for effective language learning. The existing literature shows inconsistencies in textbook quality and alignment with theoretically sound pedagogical principles, underscoring the need for a systematic evaluation which directly examines such a balance. Furthermore, the successful implementation of any balanced pedagogical framework is heavily contingent upon educators' professional literacy, including their capacity to critically evaluate instructional materials and align classroom practices with learners' actual communicative requirements (Soodmand Afshar, Tofighi, Asoudeh, & Ranjbar, 2018). This study seeks to assess the above-mentioned gap by evaluating Iranian EFL textbooks employing the Four Strands framework, resulting in providing empirical evidence regarding their capacity to create balanced language learning opportunities and informing future curriculum development and textbook design.

Objectives and Research Questions

The present study aims to investigate Iranian high school English textbooks through the theoretical lens of Nation's Four Strands framework. Contemporary EFL textbooks incorporate the four core language skills including speaking, listening, writing, and reading. However, significant imbalances often exist in the method of engaging with each of these skills within the curriculum, underscoring the necessity for a more integrated and theoretically grounded language curriculum which ensures equitable attention to all of the critical learning dimensions. This study seeks to examine the widely used Vision and Prospect textbook series, addressing their structural and pedagogical alignment with Nation's Four Strand framework. In addition, the present study aims to recognize strengths and deficiencies within these materials to provide actionable insights for educators to enhance their instructional practices and course materials, resulting in fostering an effective and more balanced language learning experience for Iranian high school students.

The overarching objectives of this study necessitate a focused investigation guided by the following questions.

Research Question One: How closely does the design of Iranian high school English textbooks correspond to Nation's Four Strands of language learning principles (Nation, 2007)?

Research Question Two: To what extent have Iranian high school English textbooks integrated approaches which facilitate effective "meaning-focused input" learning opportunities?

Research Question Three: How systematically have Iranian high school English textbooks incorporated approaches which facilitate effective "meaning-focused output" learning opportunities?

Research Question Four: To what degree have approaches in Iranian high school English textbooks been designed to facilitate effective "language-focused learning" opportunities?

Research Question Five: How systematically have Iranian high school English textbooks integrated approaches which facilitate effective "fluency development" learning opportunities?

The aforementioned questions seek to analyze each strand representation within the textbooks, offering a systematic evaluation of the degree to which these materials align with the theoretically grounded, balanced approach to language learning advocated by Nation's framework.

Methodology

Research Design

Here, a mixed-methods evaluation design was used to examine the alignment of Iranian high school English textbooks with Nation's Four Strands framework. Integrating qualitative and quantitative approaches enabled a robust investigation, allowing for statistical analysis of teacher perceptions and detailed textual examination of the textbook content. The methodology was structured in two distinct yet complementary phases. The first phase adopted a quantitative approach applying a researcher-made

evaluation checklist to collect numerical data regarding educators' perceptions of the textbooks alignment with the Four Strands principles. This phase aimed to capture the subjective experiences and judgments of practicing teachers concerning the effectiveness and balance of the materials. The second phase utilized a qualitative content analysis methodology to assess the actual content of the textbooks by categorizing the tasks, activities, and instructional focus within the selected lessons based on Nation's four distinct learning strands. This phase utilized a systematic content analysis to examine the instantiation of the theoretical framework within the curriculum materials. A convergent mixed-methods approach was employed to compare the textual analysis with teacher perceptions.

Participants

The quantitative phase encompassed a cohort of 22 Iranian high school EFL teachers recruited from diverse geographical locations across the country including major cities, towns, and rural regions. This purposive sampling strategy aimed to capture a broad spectrum of teaching experiences and perspectives. The participant cohort comprised male and female educators aged between 25-55 years. Pedagogical expertise varied significantly, from early-career teachers with 2-5 years of experience to seasoned educators with over 20 years in the field. Educational qualifications of the participants ranged from Bachelor's to Doctoral degrees, ensuring a diverse academic background among the respondents. All of the participants maintained active high school English teaching responsibilities, along with direct experience employing the textbooks under investigation. A purposive convenience sampling strategy was employed to recruit participants. This approach was selected due to the specific requirement that participants must have direct, current teaching experience with the exact Prospect and Vision textbook series under investigation. Given the logistical constraints of accessing geographically dispersed schools and the voluntary nature of survey participation, convenience sampling within a purposively defined pool ensured both practical feasibility and contextual relevance (Etikan et al., 2016; Palinkas et al., 2015). While non-probability sampling limits statistical generalizability, it remains widely accepted in applied linguistic research when studying specific pedagogical contexts or instrument validation (Dörnyei, 2020; Mackey & Gass, 2022). The recruited cohort of 22 teachers was deliberately diversified across urban, suburban, and rural regions to capture varied instructional realities, thereby mitigating some limitations inherent to convenience-based recruitment.

Instruments

Textbook Corpus

The primary materials contained six widely used Iranian high school English textbooks including the Prospect (Prospects 1, 2, and 3) and Vision series (Visions 1, 2, and 3). The Prospect series, developed for junior high school students, featured structured lessons with components such as "Conversation," "Practices," "Listening and Reading," as well as "Speaking and Writing." The Vision series, designed for senior high school students, involved units organized around themes with sections covering grammar, vocabulary, listening, reading, writing, and speaking skills recorded by non-native speakers, which may impact input quality although both series incorporate audio materials.

Content Analysis Framework

A systematic content analysis protocol was developed to study the textbooks against Nation's Four Strands framework. The process involved generating detailed coding schemes to categorize learning activities within the textbooks. To this aim, two lessons from each of the six textbooks were selected for analysis. These units were purposively chosen to represent the standard instructional structure of the series. While a larger sample would offer broader coverage, these lessons were selected based on the consistent structural pattern observed across the textbooks. Consequently, the findings presented in the Results and Discussion reflect the pedagogical balance within these sampled units.. Then, the activities were classified into one of the four strands of meaning-focused input, meaning-focused output,

language-focused learning, or fluency development. Additionally, the analysis documented the quantity and duration of activities allocated to each strand by focusing on overlapping activities which could simultaneously analyze multiple strands.

Researcher-Made English Language Textbook Evaluation Checklist

A comprehensive evaluation instrument was developed based on Nation's Four Strands theoretical framework. The checklist development followed established validity procedures based on frameworks. To ensure content validity, the initial draft was reviewed by three senior EFL instructors, who evaluated item clarity and relevance. A pilot study was conducted with 10 teachers to refine ambiguous items. The final instrument was structured into three main sections including demographic information, general attributes of the textbooks, and specific evaluations related to the Four Strands. The Four Strands section contained 42 Likert-scale items distributed for meaning-focused input (14 items), meaning-focused output (10 items), language-focused learning (8 items), fluency development (13 items), and overall strand balance (7 items). Items were rated on a 5-point scale ranging from 'very dissatisfied' (1) to 'very satisfied' (5). Reliability analysis revealed a Cronbach's alpha of ($r=0.89$) for the overall scale. The finalized checklist was administered electronically using the 'Digi Survey' platform during the final two weeks of September 2024.

Data Collection Procedure

Data were collected systematically across both phases to ensure reliability and validity. During the first phase, the researcher-made evaluation checklist was distributed electronically among the 22 participating teachers. Participants, who were provided with detailed instructions, were given two weeks to complete the survey. Reminder notifications were sent midway through the data collection period to maximize response rates. During the second phase, the content analysis was conducted by the primary researcher, who independently coded all of the selected textbook lessons based on a detailed coding protocol. This protocol established clear decision rules for strand classification, particularly for overlapping activities; a primary strand was assigned based on the dominant skill focus, while time was pro-rated between secondary strands where necessary. To ensure coding reliability, a sample of lessons was independently coded by a second trained researcher. Inter-rater reliability was calculated applying Cohen's Kappa, yielding a coefficient of ($r=0.85$), indicating substantial agreement. Discrepancies were resolved through discussion and consultation with the theoretical framework.

Data Analysis

Quantitative data from the evaluation checklist were analyzed employing descriptive and inferential statistical methods. Mean scores and standard deviations (SDs) were calculated for each strand and overall balance measures. Inferential statistics including one-way ANOVA were used to assess potential differences based on participant demographics such as teaching experience and educational level. Qualitative data from the content analysis were evaluated through systematic coding and categorization. Frequency counts of activities within each strand were calculated, followed by determining the percentages relative to total instructional time or activity counts. Cross-textbook comparisons were performed to identify patterns and variations in strand implementation across different series and grade levels.

Integrating conclusions from both phases followed a convergent approach, where quantitative and qualitative results were contrasted to provide a comprehensive evaluation of textbook alignment with Nation's Four Strands framework. This triangulation of methods strengthened the conclusions, resulting in providing a more nuanced understanding of the strengths and limitations of the examined textbooks.

Results

Here, six English textbooks widely used in Iranian high schools were investigated by focusing on their alignment with Nation's Four Strands framework for language learning. The framework posited that effective language instruction should provide balanced opportunities across key areas of meaning-focused input, meaning-focused output, language-focused learning, and fluency development, each ideally obtaining approximately 25% of instructional time. To address the research questions, the researcher conducted a detailed content analysis of two lessons from each textbook, as well as collecting evaluative data from 22 English teachers utilizing a researcher-made checklist. The following sections present the results derived from the content analysis of the textbooks and teacher evaluation survey.

Content Analysis Results

Here, two lessons from the Prospect and Vision series were studied applying Nation's Four Strand framework. The researcher analyzed the activities within each lesson to address the research questions and assess their alignment with the criteria of balance across the four strands as outlined by Nation.

First question: Extent of "Meaning-focused Input" Implementation

Table 1 indicates a detailed comparative analysis of the activities and time allocated to the meaning-focused input strand (listening and reading for comprehension) within the selected textbook lessons.

Table 1

Activities and Time Allocation to the Meaning-Focused Input Strand in the Prospect and Vision Series

Name of the textbook	No. of lesson	No. of activities	Percentage	Time	Percentage
Prospect 1	Lesson 3	5	33.33%	60 min	33.33%
Prospect 1	Lesson 8	5	33.33%	65 min	36.11%
Prospect 2	Lesson 1	6	28.57%	60 min	31.58%
Prospect 2	Lesson 5	6	28.57%	60 min	31.58%
Prospect 3	Lesson 2	7	25%	75 min	23.44%
Prospect 3	Lesson 4	7	25%	75 min	23.44%
Vision 1	Lesson 1	21	29.58%	225.5 min	29.48%
Vision 1	Lesson 4	20	28.98%	204 min	26.67%
Vision 2	Lesson 2	21	28.38%	283 min	27.34%
Vision 2	Lesson 3	21	29.58%	293 min	28.58%
Vision 3	Lesson 1	18	25.35%	285 min	28.22%
Vision 3	Lesson 2	21	29.17%	345 min	35.57%
Average	12 lessons	13.7	28.74%	169.21 min	29.61%

Note: Percentages represent the proportion of activities or time dedicated to each strand relative to the total activities or time within that specific lesson. The 'Average' row reports the mean percentage across all twelve analyzed lessons.

The content analysis revealed that the Prospect series devoted an average instructional time (29.91%) to meaning-focused input, indicating proportional contributions from Prospect 1 (34.72%), Prospect 2 (31.58%), and Prospect 3 (23.44%). The Vision series exhibited a lower mean time allocation of 29.31% to this strand. Vision 3 allocated the highest percentage (31.9%), followed by Vision 1 (28.08%), and Vision 2 (27.96%) (Figs. 2 & 3). The analysis provided a comparative overview of the time allocated to the meaning-focused input strand across all of the high school English textbooks.

Figure 2

Time Devoted to Meaning-focused Input Strand

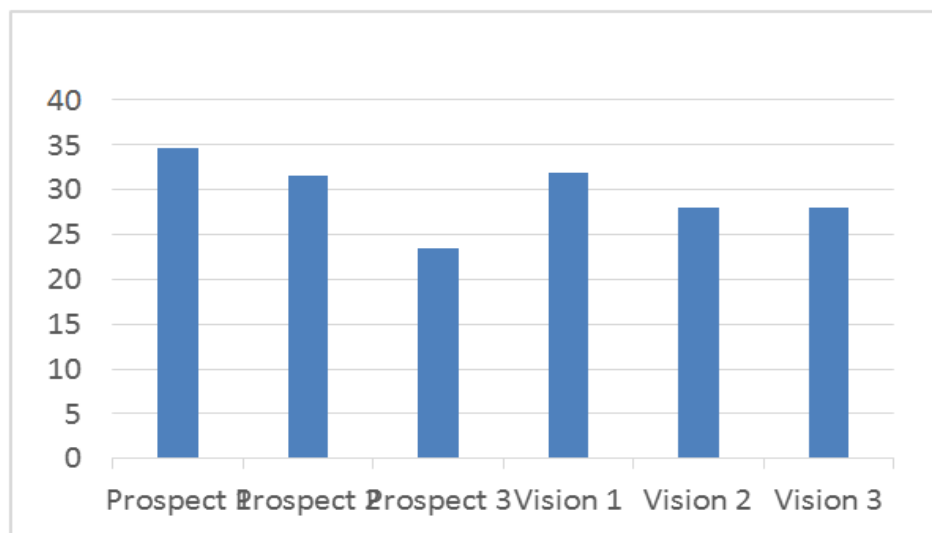
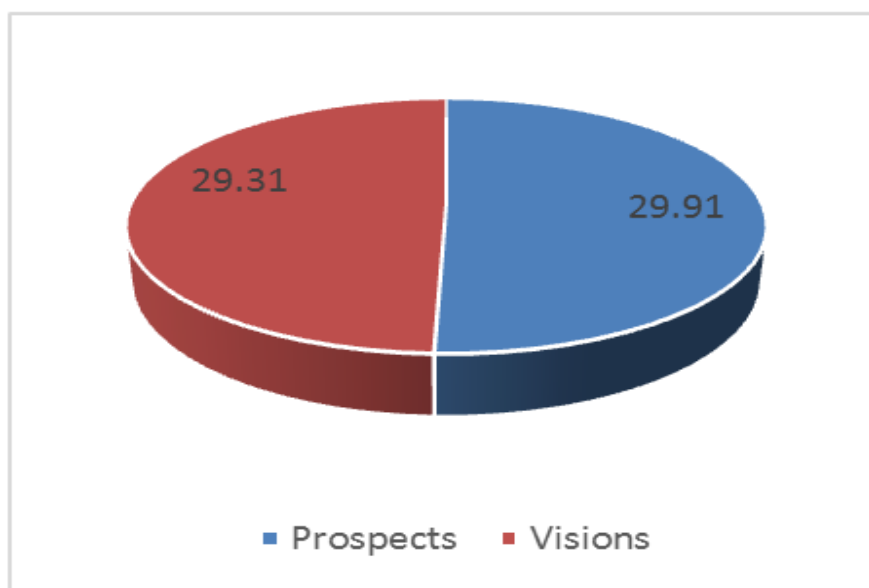


Figure 3

Comparing the Time Allocation to the Meaning-focused Input Strand



The content analysis demonstrated that the Prospect series dedicated an average instructional time (29.91%) to meaning-focused input, representing proportional contributions from Prospect 1 (34.72%), Prospect 2 (31.58%), and Prospect 3 (23.44%). The Vision series exhibited a lower average of 29.31%, with Vision 3 allocating the highest percentage (31.9%), followed by Vision 1 (28.08%), and Vision 2 (27.96%).

Second Question: Extent of "Meaning-focused Output" Implementation

Table 2 shows the percentage distribution of activities and time dedicated to the meaning-focused output strand (speaking and writing for communication) across the analyzed lessons.

Table 2

Activities and Time Allocation to the Meaning-Focused Output Strand in the Prospect and Vision Series

Name of the textbook	No. of lesson	No. of activities	Percentage	Time	Percentage
Prospect 1	Lesson 3	4	26.67%	35 min	19.44%
Prospect 1	Lesson 8	4	26.67%	30 min	16.67%
Prospect 2	Lesson 1	6	28.57%	45 min	23.68%
Prospect 2	Lesson 5	6	28.57%	40 min	21.05%
Prospect 3	Lesson 2	8	28.57%	72.5 min	22.66%
Prospect 3	Lesson 4	8	28.57%	72.5 min	22.66%
Vision 1	Lesson 1	15	21.13%	166.5 min	21.76%
Vision 1	Lesson 4	14	20.29%	160 min	20.92%
Vision 2	Lesson 2	11	14.86%	150 min	14.49%
Vision 2	Lesson 3	11	15.49%	145 min	14.15%
Vision 3	Lesson 1	16	22.54%	200 min	19.80%
Vision 3	Lesson 2	13	18.06%	160 min	16.49%
Average	12 lessons	116	19.48%	1376.5 min	287.09%

Note: Percentages represent the proportion of activities or time dedicated to each strand relative to the total activities or time within that specific lesson. The 'Average' row reports the mean percentage across all twelve analyzed lessons.

The analysis revealed that the Prospect series allocated a higher average percentage of time to meaning-focused output (21.03%) compared to the Vision series (17.93%), indicating proportional contributions from Prospect 1 (18.06%), Prospect 2 (22.36%), and Prospect 3 (22.66%). However, Vision 1 allocated the highest percentage (21.34%), followed by Vision 2 (14.32%), and Vision 3 (18.14%) (Figs. 4 & 5). The analysis offered a detailed comparative overview of the time allocated to the meaning-focused output strand across all of the high school English textbooks.

Figure 4

Time Devoted to Meaning-focused Output Strand

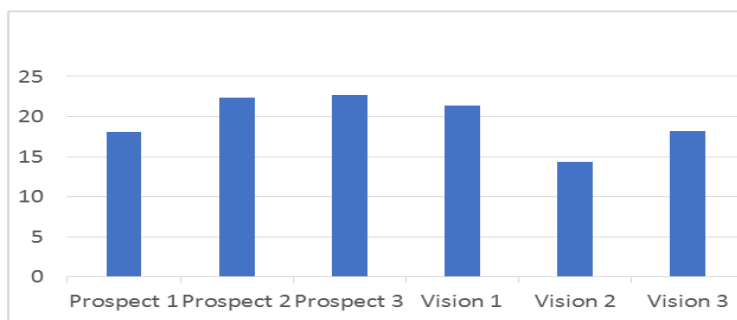
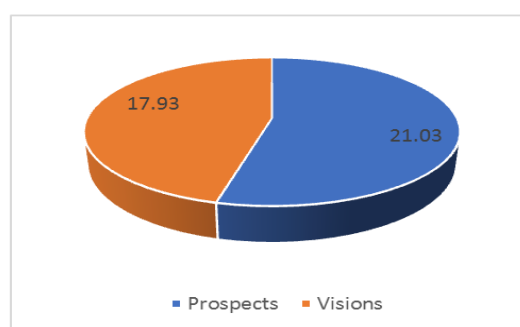


Figure 5

Comparing the Time Allocation to the Meaning-focused Output Strand



The content analysis showed that the Prospect series dedicated more time to meaning-focused output than the Vision series. The Prospect series achieved an overall mean of 21.03 % with percentages of 18.06, 22.36, and 22.66% for Prospects 1, 2, and 3, respectively. However, the Vision series demonstrated average percentages of 21.34, 14.32, and 18.14% for Visions 1, 2, and 3, respectively, leading to a total average of 17.93%.

Third Question: Extent of "Language-focused Learning" Implementation

Table 3 represents the comparative analysis of activities and time devoted to the language-focused learning strand (direct attention to vocabulary, grammar, pronunciation, and the like).

Table 3

Activities and Time Allocation to the Language-focused Learning Strand in the Prospect Series

Name of the textbook	No. of lesson	No. of activities	Percentage	Time	Percentage
Prospect 1	Lesson 3	3	20%	70 min	38.89%
Prospect 1	Lesson 8	3	20%	70 min	38.89%
Prospect 2	Lesson 1	5	23.81%	60 min	31.58%
Prospect 2	Lesson 5	5	23.81%	55 min	28.95%
Prospect 3	Lesson 2	8	28.57%	120 min	37.5%
Prospect 3	Lesson 4	8	28.57%	120 min	37.5%
Vision 1	Lesson 1	25	35.21%	307 min	40.13%
Vision 1	Lesson 4	25	36.23%	313 min	40.92%
Vision 2	Lesson 2	27	36.49%	457 min	44.15%
Vision 2	Lesson 3	26	36.62%	442 min	43.12%
Vision 3	Lesson 1	25	35.21%	425 min	42.08%
Vision 3	Lesson 2	26	36.11%	365 min	37.63%
Average	12 lessons	186	38.44%	2804 min	461.34%

Note: Percentages represent the proportion of activities or time dedicated to each strand relative to the total activities or time within that specific lesson. The 'Average' row reports the mean percentage across all twelve analyzed lessons.

The content analysis indicated that the language-focused learning strand obtained the highest average time allocation. The Prospect series dedicated an average time (35.55%) to this strand with percentages of 38.89, 30.26, and 37.5% for Prospects 1, 2, and 3, respectively. The Vision series showed even higher percentages, averaging 41.34%. Vision 1 allocated the highest percentage (40.52%), followed by Vision 2 (43.64%), and Vision 3 (39.86%) (Figs, 6 & 7). The analysis delivered a comprehensive comparative overview of the time allocated to the language-focused learning strand across all of the high school English textbooks.

Figure 6

Time Devoted to Language-focused Learning Strand

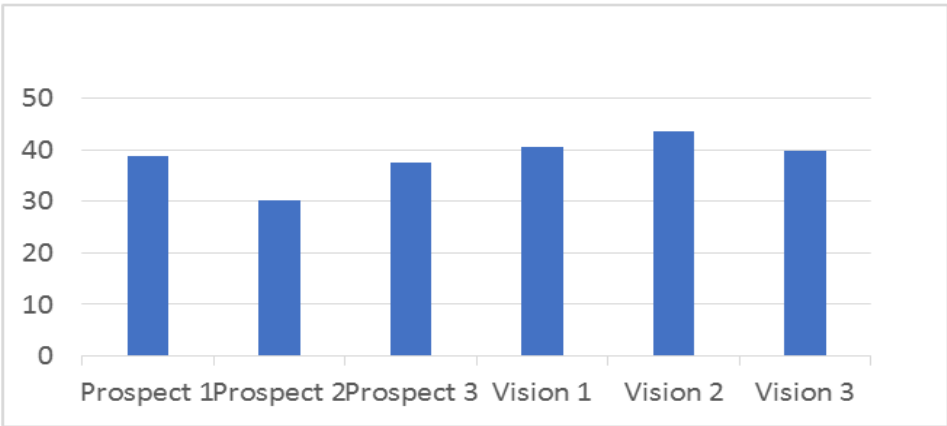
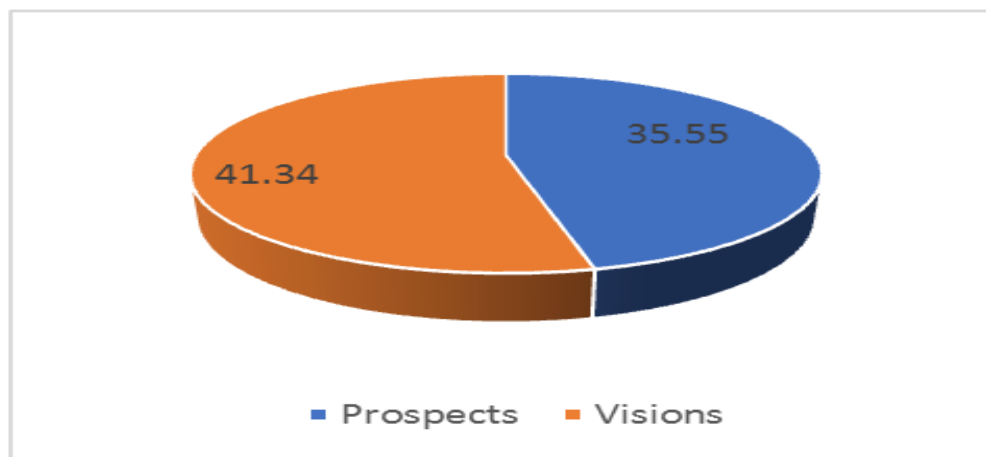


Figure 7

Comparing the Time Allocated to the Language-focused Learning Strand



The text discusses the time allocated to language-focused learning across different educational materials. In the Prospect series, Prospect 1 demonstrated the highest percentage (38.89%), followed by Prospect 3 (37.5%), and Prospect 2 (30.26%), with an average of 35.55%. However, the Vision series showed higher percentages overall, with Visions 1, 2, and 3 at 40.52, 43.64, and 39.86%, respectively, leading to an average of 41.34%.

Fourth Question: Extent of "Fluency Development" Implementation

Table 4 presents the activities and time allocated to the fluency development strand (practicing known language quickly and accurately across all of the skills).

Table 4

Activities and Time Allocation to the Fluency Development Strand in the Prospect and Vision Series

Name of the textbook	No. of lesson	No. of activities	Percentage	Time	Percentage
Prospect 1	Lesson 3	3	20%	15 min	8.33%
Prospect 1	Lesson 8	3	20%	15 min	8.33%
Prospect 2	Lesson 1	4	19.05%	25 min	13.16%
Prospect 2	Lesson 5	4	19.05%	35 min	18.42%
Prospect 3	Lesson 2	5	17.86%	52.5 min	16.41%
Prospect 3	Lesson 4	5	17.86%	52.5 min	16.41%
Vision 1	Lesson 1	10	14.08%	66 min	8.63%
Vision 1	Lesson 4	10	14.49%	88 min	11.50%
Vision 2	Lesson 2	15	20.27%	145 min	14.01%
Vision 2	Lesson 3	13	18.31%	145 min	14.15%
Vision 3	Lesson 1	12	16.90%	100 min	9.90%
Vision 3	Lesson 2	12	16.67%	100 min	10.31%
Average	12 lessons	178	12.46%	2504 min	329.09%

Note: Percentages represent the proportion of activities or time dedicated to each strand relative to the total activities or time within that specific lesson. The 'Average' row reports the mean percentage across all twelve analyzed lessons.

The fluency development strand achieved the least attention across all of the textbooks. The Prospect series allocated an average time (13.51%) to this strand, comprising Prospect 1 (8.33%), Prospect 2 (15.79%), and Prospect 3 (16.41%). However, the Vision series allocated a slightly lower average of 11.41%, containing Vision 1 (10.06%), Vision 2 (14.08%), and Vision 3 (10.1%) (Figs. 8 & 9).

The data provided an in-depth comparative analysis of the time dedicated to the fluency development strand across all of the high school English textbooks.

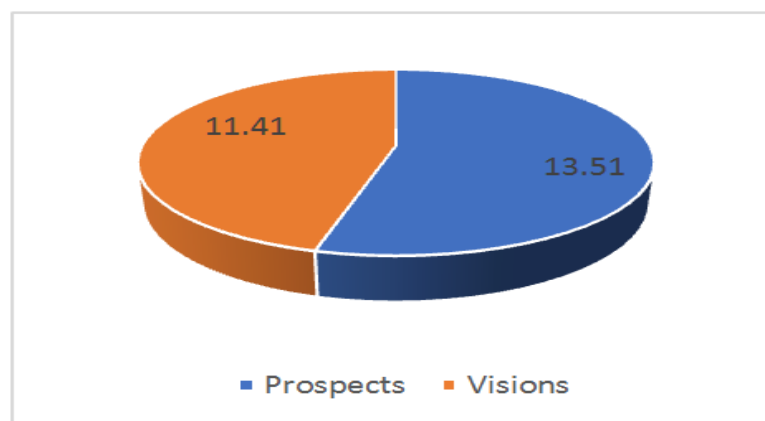
Figure 8

Time Devoted to Fluency Development Strand



Figure 9

Comparing the Time Allocation to the Fluency Development Strand



The fluency development strand attracted more attention in the Prospect series compared to the Vision series. The average time spent on this strand was higher in Prospect 2 (15.79%) and Prospect 3 (16.41%), leading to a total average of 13.51% for the Prospect series. However, the Vision series exhibited lower averages, comprising Vision 1 (10.06%), Vision 2 (14.08%), and Vision 3 (10.1%), leading to a total average of 11.41%.

Summary of Content Analysis: Overall Balance

Combining the results from all of the strands, the content analysis represented a significant imbalance in the allocation of instructional time across the textbooks. As illustrated in Figs. 10 and 11, the overall average time allocation is approximately 29.61, 19.48, 38.44, and 12.46% for meaning-focused input, meaning-focused output, language-focused learning, and fluency development, respectively. This distribution deviates considerably from the Nation's recommended equal allocation of 25% for each strand.

Figure 10

Rough Balance of Four Strands in Prospect and Vision Series

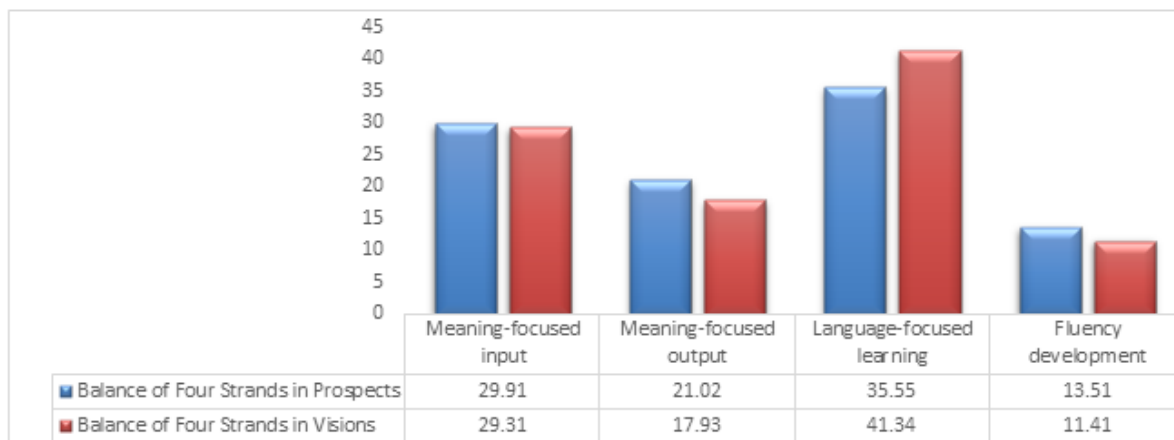
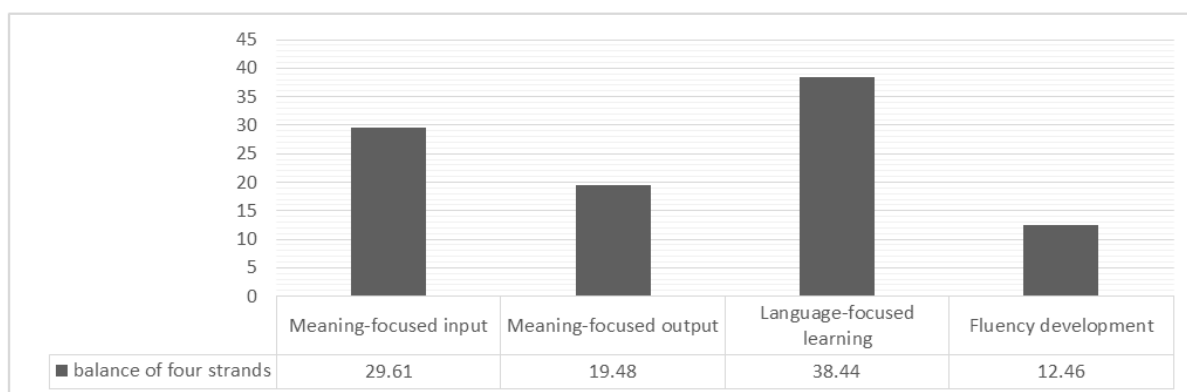


Figure 11

Average Balance of Four Strands in All of the Textbooks



Teacher Evaluation Checklist Results

Data collected from the 22 teachers through the researcher-made checklist provided insights into educators' perceptions regarding the alignment of textbooks with the Four Strands framework. To this aim, one-sample t-test was conducted to compare the mean scores for each strand and overall balance against a neutral midpoint of 3.0 on the 5-point Likert scale (where 1=very dissatisfied, 5=very satisfied).

Perceptions of Individual Strands

Table 5

Results of One-Sample T-Test for the Meaning-Focused Input

Test value = 3.0						
	t	df	Sig. (2-tailed)	Mean difference	95% Confidence interval of the difference	
					Lower	Upper
Average	-3.979	21	.001	-.4825	-.735	-.230

Meaning-focused input: As indicated in Table 5, the result of one-sample t-test ($t(21)=-3.98$, $p=.001$) is statistically significant. The negative t-value and mean difference (-0.4825) represents that teachers were significantly dissatisfied with the extent to which the textbooks facilitated meaning-focused input.

Table 6*Results of One-Sample T-Test for the Meaning-Focused Output*

Test value = 3.0						
	t	df	Sig. (2-tailed)	Mean difference	95% Confidence interval of the difference	
					Lower	Upper
Average	-5.013	21	.000	-.7455	-1.055	-.436

Meaning-focused output: Table 6 shows the results for meaning-focused output. The test result ($t(21)=-5.01$, $p<.001$) was highly significant, with a negative mean difference (-0.7455), indicating strong teacher dissatisfaction with the provision for meaning-focused output activities.

Table 7*Results of One-Sample T-Test for the Language-Focused Learning*

Test value = 3.0						
	t	df	Sig. (2-tailed)	Mean difference	95% Confidence interval of the difference	
					Lower	Upper
Average	-1.355	21	.190	-.2330	-.591	.125

Language-focused learning: Table 7 represents the results for language-focused learning. The result ($t(21)=-1.36$, $p=.190$) was not statistically significant. The mean difference (-0.2330) suggests a trend toward dissatisfaction although teachers' perceptions remained neutral overall for this strand.

Table 8*Results of One-Sample T-Test for the Fluency Development*

Test value = 3.0						
	t	df	Sig. (2-tailed)	Mean difference	95% Confidence interval of the difference	
					Lower	Upper
Average	-5.018	21	.000	-.7622	-1.078	-.446

Fluency development: Table 8 presents the results for fluency development. The test result ($t(21)=-5.02$, $p<.001$) was highly significant, with a substantial negative mean difference (-0.7622), demonstrating strong teacher dissatisfaction with the emphasis placed on fluency development in the textbooks.

Perceptions of Overall Balance

Balancing the four strands: Table 9 indicates the results for the overall balance among the four strands. The result ($t(21)=-6.49$, $p<.001$) was highly significant, with the largest negative mean difference (-1.0000), confirming teachers' strong dissatisfaction regarding the overall balance of the four strands in the textbooks.

Table 9*Results of One-Sample T-Test for Balancing the Four Strands*

Test value = 3.0						
	t	df	Sig. (2-tailed)	Mean difference	95% Confidence interval of the difference	
					Lower	Upper
Average	-6.487	21	.000	-1.0000	-1.321	-.679

Rough Balance of Four Strands

The content analysis of the Prospect and Vision series revealed that the meaning-focused input strand in both series is allocated an average of 29.91 and 29.31% of the total time, respectively. In addition, the meaning-focused output strand is allocated 21.02 and 17.93% of the time, respectively. Further, the language-focused learning strand is given 35.55 and 41.34% of the total time in the Prospect and Vision series, respectively. Finally, the fluency development strand is allocated 13.5 and 11.41% of the total time in the Prospect and Vision series, respectively. Based on the results, high school English textbooks allocated an average of 29.61 and 19.48% of their time to the meaning-focused input and meaning-focused output, respectively. Further, the language-focused learning obtained 38.44% of the time, while the fluency development accounted for only 12.46%.

Qualitative feedback from teachers corroborated the quantitative results. Participants overemphasized grammar and vocabulary (language-focused learning) and a lack of activities promoting balanced skill development, especially in speaking and writing (meaning-focused output). Additionally, they highlighted the insufficient time allocated for English instruction (two hours per week) as a contributing factor to the imbalance. Teachers critiqued the authenticity of language use and limited cultural perspectives presented in the textbooks, while some participants appreciated the inclusion of teacher guides. Overall, participants held a negative view regarding the general attributes of the textbooks, as confirmed by a significant one-sample t-test ($t(21)=-3.18, p=.005$).

In summary, both the objective content analysis and subjective teacher evaluations revealed a significant imbalance in the implementation of Nation's Four Strands within the analyzed Iranian high school English textbooks. The textbooks disproportionately emphasized language-focused learning while underrepresenting meaning-focused output and fluency development, failing to provide the balanced approach advocated by Nation's framework.

Discussion

The results reveal a systemic imbalance in the practical application of Nation's (2007) Four Strands framework within the analyzed Iranian high school English textbooks, particularly within the Vision series. The findings indicate a clear departure from the theoretically established principle of equal time allocation. (approximately 25% each) across meaning-focused input, meaning-focused output, language-focused learning, and fluency development (Nation, 2007). The empirical data, obtained through systematic content analysis and teacher evaluations, reveal a pedagogical environment in which structural language instruction takes precedence over the development of communicative and fluency-based competencies. While these interpretations should be regarded in light of the study's methodological constraints, they establish a clear pattern that requires contextualization against divergent findings in the literature.

Based on the content analysis, language-focused learning activities consume a disproportionately high percentage of instructional time, averaging 35.55% and 41.34% in the Prospect and Vision series, respectively. This dominant focus on structural components—namely grammar and vocabulary—reflects the structural bias pervasive in Iranian EFL textbooks (Kaffash et al., 2018; Jamshidian et al., 2024). However, this finding contrasts with earlier empirical work documenting a more equitable distribution of CLT-based activities. For instance, Ebrahimi and Sahragard (2017) documented positive teacher attitudes toward CLT-designed textbooks in Isfahan, attributing success to adaptive planning and teacher agency. Similarly, Arab (2020) found that a more balanced strand distribution emerged when teachers closely followed the Prospect 3 manual. These studies demonstrate that high levels of implementation fidelity and strategic teacher mediation can serve to ameliorate structural imbalances. Collectively, the literature indicates that textbook design does not necessarily prescribe classroom practice; instead, teacher beliefs and pedagogical autonomy serve as critical moderating variables. In addition, the analysis reveals a clear underrepresentation of meaning-focused output activities with the Prospect series allocating an average of 21.03% and the Vision series only 17.93% of instructional time to speaking and writing tasks designed for genuine communication. This

deficiency warrants pedagogical attention, particularly in light of the theoretical importance of output in promoting syntactic processing and hypothesis testing (Swain, 1995; Nation, 2009). The scarcity of production opportunities aligns with findings by Goodarzi (2020) and Sadeghzadeh (2021), who identified weaknesses in activity design and methodology within the textbook series. Teacher evaluation data reinforces this concern, with participants expressing significant dissatisfaction ($t(21) = -5.01$, $p < .001$) with the provision for meaning-focused output. However, these results should be contextualized within the broader challenges of CLT implementation in Iran. Yaghoubinejad et al. (2017) indicated that Iranian junior high school teachers frequently experience demotivation when forced to implement CLT principles without resources, adequate training, or institutional support. This “CLT trap” elucidates the tendency for output-focused activities to be marginalized or pedagogicalized into more constrained, grammar-translation formats, irrespective of the intended textbook design.

The fluency development is significantly underrepresented, averaging only 13.51% and 11.41% in the Prospect and Vision series, respectively, contradicting Nation's (2009) assertion that fluency activities play a critical role in building confidence and proficiency. Teachers' strong negative perceptions regarding fluency development opportunities ($t(21) = -5.02$, $p < .001$) validate the quantitative results, confirming that educators recognize the limitations imposed on developing learners' ability to employ language effortlessly. However, recent studies advise against a reductive interpretation of low fluency allocation as being solely indicative of pedagogical failure. Ahmadi and Derakhshan (2022) note that in exam-driven EFL contexts like Iran, fluency development is often deprioritized due to systemic assessment pressures rather than textbook design flaws. Thus, the observed imbalance reflects broader curricular constraints rather than being attributable solely to isolated deficiencies in the materials.

The overemphasis on language-focused learning, combined with insufficient attention to meaning-focused output and fluency development, contributes to persistent challenges in developing communicative competence among Iranian EFL learners. These results align with a portion of the existing literature yet stand in contrast to other studies, illustrating the inherent complexity of assessing instructional materials in contexts undergoing a transition toward CLT. The implications of such an imbalance present significant challenges for Iranian EFL education; however, they must be considered in light of specific contextual realities. The overemphasis on language-focused learning, combined with insufficient attention to meaning-focused output and fluency development, contributes to persistent challenges in developing communicative competence among learners. However, as Yaghoubinejad et al. (2017) and Kheradmehr (2022) argue, successful CLT implementation requires systemic alignment including assessment reform, teacher training, and institutional support must accompany material redesign. Lacking these essential components, theoretically sound, balanced instructional materials will likely fail to catalyze a meaningful shift in actual classroom pedagogy.

However, there are several limitations which warrant consideration. The focus on only two lessons per textbook may not fully represent the entire corpus of activities within each series. Additionally, the teacher evaluation sample was relatively small, failing to capture the full spectrum of perspectives across Iran's varied educational landscape. Future studies should utilize larger samples and longitudinal designs to examine the impact of such imbalances on actual student learning outcomes.

The present study indicates that the analyzed Iranian high school English textbooks, do not particularly the Vision series, may not fully implement Nation's Four Strands framework as intended, resulting in an imbalanced approach that overemphasizes structural learning while neglecting essential communicative and fluency development opportunities. These results point to clear shortcomings and indicate a need for curriculum reform and textbook revision which prioritizes a more balanced approach to language instruction, as well as valuing comprehension, production, structural knowledge, and fluent communication. Such reforms can play a crucial role in preparing Iranian students to meet the communicative demands of the 21st century, provided they are accompanied by teacher support and assessment alignment.

Conclusion and Implications

The comprehensive evaluation of Iranian high school English textbooks, specifically the Prospect and Vision series, through the lens of Nation's Four Strands framework suggests a potential pedagogical imbalance that may impede the efficacy of language learning. While these textbooks incorporate activities ostensibly covering all of the language skills, the analysis indicates a potential deficiency in the provision of sufficient opportunities for meaning-focused output and fluency development within the sampled lessons. The content analysis provided a quantitative measure of this imbalance, revealing that language-focused learning occupies a disproportionate share of instructional time (averaging 38.44% across all of the textbooks), exceeding the recommended 25%, while fluency development emerges as the least represented strand (averaging only 12.46%). In addition, meaning-focused input (29.61%) and meaning-focused output (19.48%) misalign with the ideal of a balanced pedagogical distribution, which suggests a possible limitation in the textbooks' capacity to foster communicative fluency as the core tenet of CLT. Thus, the main objective of enabling students to apply English effectively and confidently in real-world contexts may not be fully supported by these materials.

The results suggest the following insights into the current state of Iranian high school English textbooks:

- Textbooks incorporate meaning-focused input activities, primarily through reading and listening passages. However, these could potentially be enhanced by incorporating more authentic and engaging materials which represent real-world language use effectively.
- Opportunities for meaningful output, especially speaking and writing tasks focused on genuine communication, appear to be constrained, suggesting a need for incorporating more interactive, communicative, and production-oriented activities.
- Statistical analysis reveals significant emphasis on language forms and structures (language-focused learning). However, this focus could be better balanced with more opportunities for authentic language use and application within meaningful contexts.
- The critical strand of fluency development is underrepresented in the analyzed samples, suggesting that students may not be provided with sufficient practice to employ known language elements quickly, accurately, and effortlessly across all of the skill areas.

The findings of this study carry substantial implications for a broad spectrum of stakeholders within the Iranian ELT landscape.. English teachers at junior and senior high school levels, textbook writers, researchers, curriculum designers can derive valuable insights from this study. The study underscores significant pedagogical gaps in contemporary high school English textbooks, specifically regarding the marginalization of meaningful communication, which plays a critical role in learning a foreign language effectively. The present study highlights the necessity for an equitable emphasis on all of the strands in language learning including meaning-focused input, meaning-focused output, language-focused learning, and fluency development. The marginalization of any single strand may compromise overall language proficiency, thereby impeding learners' capacity to attain academic and professional goals within a globalized socioeconomic landscape.. Additionally, the study pinpoints the significance of integrating authentic, culturally relevant materials and incorporating interactive, task-based activities to enhance comprehension and productive communication skills. Revising English textbooks to align more closely with current, evidence-based pedagogical principles, especially Nation's balanced framework, could potentially improve student learning outcomes and motivation. Overall, these results offer guidance for textbook developers and educators, underscoring the value of adopting balanced, holistic approaches and innovative methodologies in ELT to foster a more robust and pedagogically sound learning environment.

Limitations

This study encountered several methodological limitations which should be acknowledged. Prior studies evaluated textbooks using Nation's Four Strands framework within the Iranian EFL context, making the present one of the pioneering comprehensive assessments of its kind. In addition, the content analysis relied primarily on investigating textbook materials and teacher guides, despite being

systematic. Direct classroom observation could have provided richer, contextualized data on the method of implementing and obtaining the materials. Further, the teacher evaluation component utilized a relatively small sample size of only 22 participants, predominantly female, which restricts the generalizability of the perceptual results to the broader population of Iranian EFL educators. Furthermore, the study scope was delimited to an analysis of specific textbook series (Prospect and Vision), focusing exclusively on the Four Strand evaluation framework, resulting in ignoring other potential extraneous variables influencing curriculum effectiveness. Senior high school English courses typically offer more weekly instructional hours, potentially allowing for more effective implementation of the four strands. However, this study failed to address the impact of varying instructional time allocations. Finally, the content analysis was confined to only two lessons from each textbook, hypothesizing that the structured approach across units results in creating similar activity designs, differing mainly in thematic content.

Suggestions for Future Studies

The present study is intended to provide a robust empirical baseline for future inquiries into English language textbook evaluation and material development within EFL contexts. The findings underscore the imperative for large-scale, longitudinal research to evaluate the distribution of the four strands across a more diverse corpus of high school ELT materials, including comparative analyses with global and L1-targeted textbooks. Future studies should prioritize collecting comprehensive data on user needs and expectations from diverse stakeholder groups including students to ensure the development of more effective and learner-centered English textbooks. This study was grounded in the analysis of textbook content and the perspectives of English teachers. However, future studies should incorporate learners' direct perspectives and experiences to enhance the validity and applicability of the results. Finally, longitudinal studies can evaluate the effect of strand-balanced versus imbalanced textbook approaches on actual student learning outcomes and proficiency gains.

Acknowledgements

I extend my heartfelt appreciation to all of the participants involved in the present study for their invaluable cooperation and willingness to share their insights. Their contributions played a key role in the success of this study. I am also deeply grateful to Farhangian University for providing access to essential facilities and supporting this research.

Declaration of Conflicting interests

The authors confirm the absence of any financial or non-financial conflicts of interest potentially influencing the research, authorship, or publication of this manuscript, indicating the integrity and objectivity of the work presented.

Funding Details

According to the authors, the study was conducted without external funding sources.

References

- Ahmadi, A., & Derakhshan, A. (2016). EFL teachers' perceptions towards textbook evaluation. *Theory and Practice in Language Studies*, 6(2), 260-267. <https://doi.org/10.17507/tpls.0602.06>
- Arab, H. (2020). *A critical analysis of Prospect 3 based on Nation and Macalister's four-strand framework* [Master's thesis, Islamic Azad University, Shahrood Branch]. <https://ganj.irandoc.ac.ir>

- Baqeri, M. (2024). *Textbook evaluation of Prospect One and Family & Friends One: A multi-instrumental study* (Publication No. 1061631) [Master's thesis, Shahid Bahonar University of Kerman]. IranDoc. <https://ganj.irandoc.ac.ir>
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H. & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: putting the linguistic approach to the test. *Language Teaching Research*, 10(3), 245-261. <https://doi.org/10.1191/1362168806lr195oa>
- Cunningsworth, A. (1995). *Choosing your coursebook*. Oxford: Heinemann.
- Derakhshan, A. (2021). 'Should textbook images be merely decorative?' Cultural representations in the Iranian EFL national textbook from a semiotic approach perspective. *Language Teaching Research*, 28(1), 79–113. <https://doi.org/10.1177/1362168821992264>
- Ebrahimi, F., & Sahragard, R. (2017). Teachers' perceptions of the New English textbooks in Iranian junior high schools. *The Journal of AsiaTEFL*, 14(2), 355-363. <https://doi.org/10.18823/asiatefl.2017.14.2.12.355>
- Ellis, N. C. (2012). What can we count in language, and what counts in language acquisition, cognition, and use?. *Frequency Effects in Language Learning and Processing*, 1, 7–34. <https://doi.org/10.1515/9783110274059.7>
- Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Basil Blackwell.
- Ellis, R. (1997). The empirical evaluation of language teaching materials. *ELT Journal*, 51, 36–42. <https://doi.org/10.1093/elt/51.1.36>
- Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, 40(1), 83–107. <https://doi.org/10.2307/40264512>
- Goodarzi, A. (2020). *A comprehensive evaluation of Iranian junior high school English course books* (Publication No. 626334) [Doctoral dissertation, Razi University]. <https://ganj.irandoc.ac.ir>
- Harwood, N. (2005). What do we want EAP teaching materials for? *Journal of English for Academic Purposes*, 4(2), 149–161. <https://doi.org/10.1016/j.jeap.2004.07.008>
- Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly* 40(1), 109–131. <https://doi.org/10.2307/40264513>
- Iswandari, Y., & Paradita, L. (2019). Extensive reading in an EFL setting: A Special interview with Professor Paul Nation. *TEFLIN Journal- a Publication on the Teaching and Learning of English*, 30(2), 187-196. <https://doi.org/10.15639/teflinjournal.v30i2/187-196>
- Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition*, 24(4), 541-577. <https://doi.org/10.1017/s0272263102004023>
- Jamshidian, F., Tabatabaei, O., Salehi, H., & Vahid Dastjerdi, H. (2024). Iran's high school English textbooks (vision series) and organizational knowledge as a component of the CLT model. *Journal of Critical Studies in Language and Literature*, 5(4), 31-44. <https://doi.org/10.46809/jcsll.v5i4.279>
- Kaffash, F., Yazdanmehr, E., & Ghanizadeh, A. (2018). EFL teachers' and students' evaluation of prospect 3 in the light of Cisar's communicative model. *International Journal of Applied Linguistics & English Literature*, 7(3), 143-150. <https://doi.org/10.7575/aiac.ijalel.v.7n.3p.143>
- Kalfut, T. (2024). Approaches and limitations of L2 textbook evaluation research. *International Journal of Language and Literary Studies*, 6(1), 216–225. <https://doi.org/10.36892/ijlls.v6i1.1617>
- Karbalaei, A. (2025). We believe in CLT, but...': Iranian EFL teachers' perceptions and implementation challenges. *GIST – Education and Learning Research Journal*, 31, 111-129. <https://doi.org/10.26817/16925777.2058>.

- Kheradmehr, V. (2022). *A study of curricular components in high school English textbooks in Iran on the basis of ACTFL guidelines* [Master's thesis, Chabahar Maritime University]. <https://ganj.irandoc.ac.ir>
- Long, M. M. (2017). Instructed second language acquisition (ISLA). *Instructed Second Language Acquisition*, 1(1), 7-44. <https://doi.org/10.1558/isla.33314>
- Nation, I.S.P. (1996). The Four strands of a language course. *TESOL in Context*, 2(6). <https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/paul-nations-publications/publications/documents/1996-Four-strands.pdf>
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/s0008413100018260>
- Nation, I.S.P. (2007). The four strands. *Innovation in Language Learning and Teaching*, 1(1), 1-17. <https://doi.org/10.2167/illt039.0>
- Nation, I.S.P. (2008). *Teaching ESL / EFL reading and writing*. Routledge. <https://doi.org/10.4324/9780203891643>
- Nation, P. (2017). Keeping it practical and keeping it simple. *Language Teaching*, 51(1), 138-146. <https://doi.org/10.1017/s0261444817000349>
- Nation, I.S.P., & Hu, M. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430. <https://doi.org/10.64152/10125/66973>
- Nation, I.S.P., & Macalister, J. (2010). *Language curriculum design*. Routledge.
- Nation, I.S.P., & Newton, J. (2009). *Teaching ESL / EFL listening and speaking*. Routledge.
- Nation, P., & Yamamoto, A. (2020). Applying the four strands. <https://doi.org/10.26686/wgtn.12552020>
- Ranjbar, N., & Soodmand Afshar, H. (2021). A survey of EAP needs in Iran from the viewpoints of teachers and students. *Iranian Journal of English for Academic Purposes*, 10(1), 54-74. [20.1001.1.24763187.2021.10.1.4.8](https://doi.org/10.24763/187.2021.10.1.4.8)
- Renandya, W. A. (2013). Essential factors affecting EFL learning outcomes. *English Teaching*, 68(4), 23-41. <https://doi.org/10.15858/engtea.68.4.201312.23>
- Richards, J. C. (1998). *Beyond training*. Cambridge: Cambridge University Press.
- Saarinen, O. (2024). *The significance of language input in the production of output skills* [Doctoral dissertation, University of Turku]. <https://www.utupub.fi/server/api/core/bitstreams/8581c96f-9da4-4644-bb76-8d6241f681d4/content>
- Sadeghzadeh, E. (2021). *An evaluation of English textbooks used in Iranian high schools based on teachers' perceptions* [Master's thesis, University of Tabriz]. <https://ganj.irandoc.ac.ir>
- Sadrabadi Haqiqi, D. (2023). *Islamic Cultural Capital in ELT Textbooks: A Content Analysis* [Master's thesis, Ferdowsi University of Mashhad]. <https://ganj.irandoc.ac.ir>
- Salaberry, M. R. (1997). The role of input and output practice in second language acquisition. *The Canadian Modern Language Review*, 53(2), 422-451. <https://doi.org/10.3138/cmlr.53.2.422>
- Salehi, H. (2024). *An Analytical Examination of Visual Representations in High School English Textbooks* [Unpublished master's thesis]. Shahrood University of Technology. <https://ganj.irandoc.ac.ir>
- Snow, C. E., Perlmann, R., & Nathan, D. (2021). Why routines are different: Toward a multiple-factors model of the relation between input and language acquisition. In *Children's language* (pp. 65-97). Psychology Press.

- Soodmand Afshar, H., Tofighi, S., Asoudeh, & Ranjbar, N. (2018). The impact of alternative assessment knowledge, teaching experience, gender, and academic degree on EAP teachers' assessment literacy. *Iranian Journal of English for Academic Purposes*, 7(2), 1–15. [20.1001.1.24763187.2018.7.2.1.6](https://doi.org/10.1001.1.24763187.2018.7.2.1.6)
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass and C. Madden (eds). *Input in Second Language Acquisition*, (pp. 235–253). Rowley, MA: Newbury House.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook and B. Seidelhofer (Eds). *Principle and Practice in Applied Linguistics: Studies in Honour of H.G. Widdowson* (pp. 125–144). Oxford University Press.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, 82 (3), 320–337. <https://doi.org/10.1111/j.1540-4781.1998.tb01209.x>
- Tomlinson, B. (2011). *Materials development in language teaching*. Cambridge University Press. <https://doi.org/10.1017/9781139042789.001>
- Toth, P. D. (2006). Processing instruction and a role for output in second language acquisition. *Language Learning*, 56(2), 319-385. <https://doi.org/10.1111/j.0023-8333.2006.00349.x>
- Widiatmoko, P. (2014). *Developing classroom material to encourage integrated skill teaching*. 3rd UAD TEFL International Conference. Proceeding ISBN978-602-18907-1-4.
- Williams, J. (2005). Form-focused instruction. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning*, 671–691. <https://doi.org/10.4324/9781410612700-51>
- Yaghoubinejad, H., Zarrinabadi, N., & Nejadansari, D. (2016). Culture-specificity of teacher demotivation: Iranian junior high school teachers caught in the newly-introduced CLT trap! *Teachers and Teaching*, 23(2), 127-140. <https://doi.org/10.1080/13540602.2016.1205015>