Korean Journal of English Language and Linguistics, **Vol 26, August 2026, pp. 988-1010**
DOI: 10.15738/kjell.26..202608.988

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Developing a Diagnostic Questionnaire for Learner Characteristics in Korean Secondary EFL Classrooms

Eunjung Kim · Mun-Hong Choe (Chonnam National University)

Received: October 16, 2025
Revised: May 3, 2026
Accepted: July 15, 2026

Kim, Eunjung (First author)
Lecturer, Department of Education
Chonnam National University
Email: eunjung1202@jnu.ac.kr

Choe, Mun-Hong (Corresponding author)
Professor, Department of English Education
Chonnam National University
Yongbong-ro 77 Buk-gu
Gwangju, Korea
Tel: +82-62-530-2435
Email: munhong@jnu.ac.kr

ABSTRACT

Kim, Eunjung and Mun-Hong Choe. 2026. Developing a diagnostic questionnaire for learner characteristics in Korean secondary EFL classrooms. *Korean Journal of English Language and Linguistics 26*, 988-1010.

This study aims to develop a diagnostic questionnaire to investigate individual differences in language learning dispositions among Korean secondary school students learning English as a foreign language, with particular interest in how balanced or imbalanced skill development may result from learners' variability in language-related preferences or competences as they interact with the given learning environment. Drawing on current theories in second language acquisition, learner characteristics were operationalized in terms of three macro-constructs: sensory-mode and functional preferences in L2 learning, sensitivities and capacities in L2 processing, and adaptive-strategic efforts for L2 development. An initial pool of items was developed through a review and adaptation of previous scales, and then interviews with students, expert panel reviews, and pilot testing were conducted to refine the questionnaire, resulting in an operative version consisting of 25 items. A pilot study was conducted with 236 middle school students. The Rasch analysis results indicated a person reliability of 0.91 and an item reliability of 0.92, demonstrating a high degree of measurement consistency. The results of confirmatory factor analysis indicated that a six-factor model comprising Listening-Pronunciation, Vocabulary-Grammar, Discourse-Pragmatics, Verbal Memory, Metalinguistic Awareness, and Developmental Effort demonstrated the best fit when treated separately, $\chi^2(233) = 435.102$, $p < .001$. The discussion suggests that when learner characteristics interact linearly with rigid learning environments, they can restrict input processing, consequently giving rise to imbalanced development.

KEYWORDS

individual differences, learner characteristics, scale development, learner traits, EFL

1. Introduction

While research in language learning has primarily focused on inter-individual differences, an individual's ability to process and produce a language, whether it is their first language (L1) or second (L2), may vary considerably across different aspects of linguistic knowledge and skill (Castro and Peck 2005, Dörnyei 2005, Joe et al. 2017, Li et al. 2022). Such uneven developments across language skills are pronounced in English-as-a-foreign-language (EFL) contexts, where most students receive input and instruction in controlled classroom settings, often skewed toward lexicogrammatical knowledge and receptive skills. In interaction with these environmental conditions, a learner's specific traits related to language learning play a role in reinforcing or inhibiting the development of certain language skills (Andreou et al. 2008, Banaruee et al. 2022, Pawlak 2022). In fact, little is currently known about the degrees and patterns of uneven development in L2 acquisition, the distinct and collective roles of input, instruction, and learner traits in the emergence of imbalances, the extent to which they affect the learner's persistent development, and its implications for the distinct nature of first and second language acquisition.

Given that language learning involves dynamic interactions between learner internal and external factors, investigating the sources of developmental disparities within a learner can contribute to our understanding of how the interaction of three components of language learning—input properties, instructional settings, and learner characteristics—arises in the process of learning (Larsen-Freeman and Cameron 2008, Long 2013, Ortega 2019). Depending on the roles they serve and how they interact in combination, a learner's overall language development may be balanced or uneven across skill/knowledge domains. In a simplified scenario, when a learner's characteristic traits accord with the focus of the learning environment, they are likely to benefit from the environment, developing skills that match this trait-environment alignment. However, this alignment may also limit the learner's exposure to other language skills that neither their inherent traits nor the learning environment particularly support, leading to uneven development. Conversely, when a learner's characteristic traits and the learning environment do not match, they may struggle to engage with the learning process or feel compelled to suppress their natural tendencies to conform to the environment. This mismatch may then result in the underdevelopment of both the language skills favored by the learner's predispositions and those emphasized by the environment.

Assessing balancedness in an L2 learner's language skills is therefore important not only for improving learning environments but for uncovering the factors and mechanisms that contribute to individual differences in language learning (DeKeyser 2012, Wen et al. 2023). This seems particularly relevant in EFL contexts where school-age students are placed in externally controlled environments, and so their individual learning traits have minimal influence on the way they are taught. This raises questions about whether, and to what extent, students exhibit balanced or skewed development and to what degree students' underachievement in certain areas can be attributed to a mismatch between their learning traits and the environments in which they are taught.

This issue is of course not new. One may argue that the notion of learner traits is merely a rearticulation of well-established constructs such as language aptitude, learner background, and learning styles. Language aptitude, which has recently attracted renewed attention (Doughty 2014, 2019, Jackson 2020, Li and DeKeyser 2021, Wen et al. 2023), has traditionally been conceived as an inter-individual factor to predict successful versus unsuccessful learners. Aptitude models and measurement tools, while emphasizing cognitive abilities or domain-general intelligence, give less attention to individual variability in specific component skills. Learner background questionnaires may also serve similar purposes, but their primary function was to categorize learners into groups based on cohort variables such as age of onset, length of exposure, and manner of acquisition. The questionnaires are often used to validate participant groupings and to exclude ineligible participants who do not meet required

conditions such as L2 proficiency level. Learning styles are also a well-documented construct used to describe how language-related abilities vary within individuals. Prior research has focused on several style-preference aspects such as field dependency and tolerance of ambiguity, but other important aspects such as sensory-modality, receptive-productive, and conceptual leveling-sharpening tendencies have not been the focus of detailed research (Ehrman and Leaver 2003, Ehrman et al. 2003, VanPatten et al. 2020). In addition to these preference-aptitude dimensions of the language learner profile, another critical variable is an individual's willingness or ability to engage with and consciously adapt to any given learning environment (Dörnyei 2005, Ushioda 2011). Learner traits of this kind have been explored in L2 research through constructs such as motivation, engagement, learning strategies, and L2 selves. Profiling a language learner and accounting for learner variability requires an examination of at least three macro-constructs: preferences in language learning, sensitivities and capacities in language processing, and efforts for language development. The selection of learning preferences, processing sensitivities, and developmental efforts provides a parsimonious yet comprehensive framework that overcomes the limitations of traditional univariate aptitude and background metrics. Integrating these dimensions can make the interplay between an individual's perceptual-cognitive predispositions and their conscious self-regulation more observable, offering a practical lens through which to explore the sources of skill imbalances in formal classroom settings.

The non-uniform nature of language skill development invites divergent interpretations from competing theories of second language acquisition. From the perspective of Skill Acquisition Theory (DeKeyser 2012), uneven development is a function of individual differences in processing sensitivity, memory capacity, and the amount of deliberate practice. It would suggest that learners possessing superior memory and perceptual-cognitive sensitivity in specific modalities are better able to transition input from declarative to procedural knowledge. In contrast, sociocultural and sociocognitive traditions emphasize the social and affective dimensions of learning (Lantolf and Beckett 2009). In this view, attentional resources are not merely static cognitive capacities but are shaped by identity alignment and environmental affordances. Learners often prioritize skills perceived as having high social or functional utility. For instance, a learner with high self-efficacy in literacy may be guided by a personal preference for written language or by social experiences that reward development in that modality, leaving other skills to lag accordingly. Synthesizing these perspectives can constitute a foundation for examining the interplay between learner traits and instructional settings, which will ultimately reveal how learner predispositions and social contexts converge to produce both inter-learner variations and intra-learner balance in language development.

From a practical perspective, assessing intra-learner variability is an integral part of curriculum adaptation and personalized intervention, offering more accurate insights into learners' potential and its limiting factors. Recent research has begun to notice this gap. For instance, Kautto et al. (2024) put forward the Intra-Individual Variability Hypothesis, suggesting that inconsistencies in processing times can lead to disparities in learning outcomes across various linguistic areas. Their findings indicate that increased variability in response times may degrade learning in different aspects of language, resulting in individual differences in language abilities. This emerging line of research underscores the need to explore the extent and characteristics of intra-individual differences in L2 acquisition. Against this backdrop, the present study aims to develop an operational, classroom-based questionnaire to diagnose variability in Korean EFL students' learning traits, with the goal of identifying potential causes and problems related to balanced development. The proposed tool is intended to lay the groundwork for a more comprehensive and in-depth investigation into the effects of inter- and intra-individual differences on L2 development in EFL contexts.

2. Research Review

2.1 Imbalanced Development in L2 Skills

As mentioned, an individual's ability to learn and use a second language may vary considerably across different areas of linguistic knowledge or skill. This unbalanced development has been a critical concern in EFL pedagogy, though it has received relatively little attention in L2 research (Gohar and Sadeghi 2015, Koizumi et al. 2022, Shen and Chiu 2019). Rigid institutional constraints, controlled curricula, and high-stakes testing, *inter alia*, may restrict opportunities for balanced language development. In interaction with the monolithic setting, students' characteristic dispositions, sensitivities to linguistic input, and affective attributes mediate how they perceive and approach L2 learning in the classroom. Thus, addressing disparities in EFL students' skill development hinges on an understanding of the interplay between the learning environment and individual learner traits (Kersten et al. 2021). The learning environment broadly refers to the type and amount of input and instruction that students typically receive in a given setting (Truscott and Sharwood Smith 2019). In the Korean EFL context, input is often disproportionately centered on written academic language and lexicogrammatical knowledge, with less emphasis on spoken language and productive-creative skills applicable to real-life situations. Input qualities, such as frequency, complexity, and relevance to the learner's interests, can further constrain the development of language skills (Kersten et al. 2021, VanPatten et al. 2020). Teachers' awareness of these environmental constraints, along with their ability to provide tailored interventions, can support more balanced language development. In the absence of such instructional strategies, suboptimal imbalances are likely to occur.

The likelihood for a learner to process and retain linguistic input is determined by how they engage with the component and contextual features of that input, whether consciously or implicitly (DeKeyser 2012). Learner traits can thus reinforce or suppress the development of specific language skills depending on their congruence or incongruence with the given environment. For example, a learner with strong preferences for spoken and interpersonal language use may struggle in environments where academic reading and grammatical accuracy are heavily emphasized. Accordingly, the importance of integrated-skills approaches in EFL contexts has long been acknowledged, but the inherent complexity and elusive nature of individual differences make it difficult to examine how learner traits and learning environments interact in the course of L2 development (Unsworth et al. 2014). This calls for exploratory research into how variations in learner traits relate to imbalanced skill development, the consequences of such imbalances on overall proficiency attainment, and strategies to mitigate them. Given the largely controlled homogeneity of input and instructional variables in the Korean EFL context, it provides a setting well-suited for studying the relationship between learner traits and L2 development.

2.2 Prior Approaches to Learner Variability in Foreign Language Learning

2.2.1 Preferences in language learning and use

A learning style refers to a relatively consistent way in which a person perceives, conceptualizes, organizes, and recalls information. Styles in L2 learning encompass various attributes, including field dependency, sensory modality preference, tolerance for ambiguity, constricted versus flexible control, inclination for automatization, convergent versus divergent thinking, willingness to imitate models, and the need for immediate closure (Castro and Peck 2005, Dewaele and Wei 2013, Hatami 2013). Though traditionally conceived as independent of

contextual variables, learning styles can be reinforced or inhibited by learners' experiences and learning environments (Ehrman and Leaver 2003). Since language ability involves a range of knowledge and skill components, no definitive advantage of one style over its counterpart (e.g., field-independent vs. field-dependent) has been established in terms of overall proficiency development (Andreou et al. 2008). Yet, certain preference aspects appear to be more closely linked to language development than others.

Above all, sensory modality preferences play a direct role in language processing and use. While English L1 speakers strongly prefer kinesthetic learning, and ESL learners of most nationalities are primarily tactile or kinesthetic, Korean EFL students are known to be the most visual and least auditory group (Grey et al. 2015). Breadth of categorization and the conceptual leveling-sharpening distinction are often cited as style dimensions related to L2 acquisition. Research findings are not convergent. Ehrman et al. (2003) found that sharpeners performed better on language memory tasks, but Larsen-Freeman and Long (2014) found no significant effect of categorization style on L2 achievement. Among the widely recognized models for learner styles are those by Dunn (1984), Ehrman and Leaver (2003), Fleming (2006), Kolb (1999), and Reid (1995). These models differ considerably in their conceptualizations and categorizations of learning styles. Dunn (1984) emphasizes emotional, environmental, sociological, physiological, and psychological factors, making it a broad and multifaceted approach. Ehrman and Leaver (2003) focused on domain-general cognitive styles, specifically contrasting field-dependent and field-independent learners. Kolb's (1999) model, rooted in experiential learning, classifies learners as convergers, divergers, assimilators, or accommodators, while Fleming (2006) categorizes learners based on sensory preferences: visual, auditory, read/write, and kinesthetic. These complementary perspectives demonstrate that learning styles are multidimensional, and that effective learning environments should account for both intrinsic preferences and situational adaptability.

Rather than treating language as a singular competence, this study adopts a quadrant-based approach to learner style preferences, distinguishing between spoken/written modalities and receptive/productive directions. This division acknowledges that sensory modality preferences directly influence initial language processing, while directionality accounts for the distinct cognitive demands of recognition versus retrieval. This conceptualization is uniquely positioned to uncover intra-learner imbalances that a generalized measure of overall proficiency would likely obscure.

2.2.2 Sensitivities and capacities in language processing

Language learners' varying predispositions toward noticing, comprehending, and applying subtle contrasts in the target language can be a factor in imbalanced development within individuals. This dimension of learner traits, referred to as *linguistic sensitivity*, encompasses phonological, morphosyntactic, and pragmadiscoursal intuitions that allow learners to process input in a more elaborate way. In fact, the term *sensitivity* is not frequently used to refer to language-related competences. Perhaps the debate over sensitive periods for language acquisition is the only context in which the term is legitimately used, though it serves as a synonym for non-defective input processing, rather than referring to an individual's varying abilities to use language (Granena and Long 2013). The sensitive periods hypothesis for language acquisition actually refers to a family of hypotheses, with researchers positing different time frames and criteria for defining nativeness (DeKeyser et al. 2010, Long 2013). For example, Granena and Long (2013) examined the hypothesis that the relationship between age of onset and ultimate proficiency is discontinuous before and after the sensitive periods in three domains: pronunciation, vocabulary, and grammar. They found that there were multiple sensitive periods: an early closure for pronunciation, followed by vocabulary, and a relatively delayed closure for grammar, implying that individuals' linguistic sensitivities

become more divergent after the sensitive period for a given area of perception or knowledge. Research into linguistic sensitivity, however, has focused on its role in the early stages of language acquisition and language processing, paying little attention to how sensitivities in different skill-knowledge domains interact and diverge over time within an individual.

By far the most extensively discussed topic concerning individual differences is language learning aptitude. The past decade has seen a revitalized attention to it, which gives rise to a number of competing theories and assessment tools (Doughty 2014, 2019, Granena and Long 2013, Jackson 2020, Meara 2005, Robinson 2013). As illustrated by the cursory comparisons in Table 1, researchers agree that language learning is essentially implicit and associative, and that learners with higher aptitude have more acute perceptual sensitivities and superior memory capacities for processing linguistic input (Li and DeKeyser 2021).

Table 1. Constructs of Language Learning Aptitude in Comparison

Carroll and Sapon (1959) MLAT	Meara (2005) LLAMA	Doughty (2014, 2019) Hi-LAB
. Phonetic coding . Grammatical sensitivity . Inductive learning ability . Associative memory ability	. Phonetic recognition . Vocabulary pairing . Implicit learning . Grammatical inference	· Auditory perceptual acuity · Language-related memory - working memory - associative memory - long-term memory retrieval · Implicit learning · Processing speed

As the most recent and comprehensive measurement tool, Doughty's (2014, 2019) High-level Language Aptitude Battery (Hi-LAB) comprises 13 sections that assess nine language-related abilities. Linck et al. (2013), one of its validation studies, found that three subtests in particular, paired associates (an associative learning measure), serial reaction time (an implicit learning measure), and letter span (a phonological short-term memory measure) had significant degrees of discriminative power. This corroborates the conclusion of Li's (2015) meta-analysis, which found that Carroll and Sapon's (1959) Modern Language Aptitude Test (MLAT) remains the most predictive measure of successful L2 learning (r = 0.34). From a theoretical perspective, Robinson (2013, 2019) proposes that language aptitude is a multifaceted construct that accounts for characteristic phenomena in L2 acquisition such as implicit learning, fossilization, and metalinguistic awareness. His model of language aptitude consists of four layered structures: basic cognitive abilities, aptitude complexes, task aptitudes, pragmatic-interactional traits. The innermost layer contains cognitive abilities such as pattern recognition, phonological working memory, processing automatization, etc. These abilities are then combined to form complexes, such as noticing the gap and implicit learning from contingent speech. The third circle represents where aptitude complexes interact with the features and conditions of specific tasks such as planning time and topic familiarity, referred to as task aptitudes. The outermost circle connects learners' task aptitudes with their personality, which modulates their actual performance in the given situation. This model offers an illustrative sketch of language aptitude not as a single disposition scale but as composites of interacting abilities, further underscoring the importance of a systematic approach to assessing intra-learner variability.

2.2.3 Adaptive and strategic efforts for language development

Another line of research concerning individual difference factors has focused on learners' efforts to cultivate or compensate for their (lack of) abilities in order to make further progress in a given learning environment—an

aspect considered particularly critical in EFL settings. While the style-preference and sensitivity-capacity aspects of learner traits are more or less endowed tendencies, engaged and strategic efforts involve self-management of affective, cognitive, and behavioral dynamics to enhance language development. Prior research shows that cognitive factors work in tandem with affective factors such as motivation and anxiety, influencing the language learning process and outcome, and that linguistic sensitivities, as discussed above, are closely related to individuals' formation of attitudes toward and motivation for learning an L2 (DeKeyser 2012, Garrett 2010). Moreover, learners' linguistic identity and self-awareness, i.e., their *L2 selves* in Dörnyei's (2005) terms, facilitate, or hinder, the development of the target language, and a strong link exists between strategic learning efforts and L2 achievement (Oxford 2017, Wong and Nunan 2011). Therefore, the nature and degree of self-regulated efforts that the learner puts into L2 acquisition constitute a relevant aspect of intra-learner variability.

The role of learner effort in L2 development is generally categorized into two distinct, though not mutually exclusive, dimensions: engaged effort and strategic effort. Engaged effort refers to the active, sustained affective investment a learner brings to the process, manifested through an intrinsic interest in authentic language, voluntary production, and a proactive drive to enhance proficiency. Complementing this is strategic effort—the conscious endeavor to cultivate strengths or compensate for deficiencies (Oxford 2017). While style preferences and cognitive sensitivities are largely endowed tendencies, strategic effort involves the deliberate self-management of affective, cognitive, and behavioral dynamics. Prevailing research suggests that cognitive resources, such as attention and memory, interact closely with the formation of learner attitudes and motivation, establishing a robust link between strategic self-regulation and L2 achievement (Dörnyei and Ushioda 2021). It is important to note, however, that foundational metrics in this field such as the Strategy Inventory for Language Learning and the Attitude-Motivation Test Battery were primarily validated in ESL contexts. Consequently, they may not fully capture the unique pedagogical and environmental nuances of EFL settings. This discrepancy underscores the need for context-sensitive scales tailored to the specific characteristics of the EFL classroom.

Finally, it is worth noting that the established scales, such as the Strategy Inventory for Language Learning (SILL), the Attitude-Motivation Test Battery (AMTB), and the Perceptual Learning Styles Preference Questionnaire (PSPQ) were all developed in ESL settings. EFL students' learning goals and experiences differ markedly from those of ESL students; therefore, the scales may fail to represent the learning processes and contextual factors pertinent to their experiences. Using these tools without accounting for the cultural and educational specifics of EFL settings may result in findings that do not accurately reflect the learners' characteristic dispositions related to language learning. Isemonger and Sheppard (2007), for example, examined the factor structure of a Korean version of the PSPQ through a confirmatory factor analysis, which led to the rejection of the hypothesized six-factor model. The researchers concluded that the instrument was flawed and that the results of previous research based on it were questionable. This again calls for a diagnostic tool specifically tailored for EFL students and implementable by practitioners in the classroom. Not only does the tool require conceptual validation, as outlined in the foregoing, but its practicality must also be ensured.

3. Methodology

3.1 Item Development

The construct structure in Table 2 below was employed to develop a preliminary pool of items. The initial pool of items was built through an adaptation of Choe's (2022) EFL learner profiling questionnaire (see Appendix).

This instrument maps out learners' perceived language processing and sensory preferences, domain-specific competence, metalinguistic awareness, and strategic efforts. The first section examined students' sensory-mode and functional preferences in L2 learning, with the processing modes distinguished in terms of receptive (perceptive-cognitive) and productive (productive-creative), and the modality preferences for spoken versus written language (see Table 3). This section included items to assess a learner's functional inclinations in language use, broadly divided into two domains: ideational-intellectual and interpersonal-practical (Halliday and Matthiessen 2014). This does not imply that the two are independent from each other; rather, they constitute a gradient with individuals inclined toward either pole to varying degrees. Some of the initial items are presented below.

Table 2. A Model Construct Structure of EFL Learner Traits

	Dimension	Trait inclination
Style-Preference Aspects	Sensory-mode preferences	· Spoken ↔ Written · Receptive ↔ Productive
	Functional preferences in language use	· Ideational-intellectual · Interpersonal-practical
Sensitivity-Capacity Aspects	Sensitivities in language processing	· Auditory-articulatory · Lexicogrammatical · Pragmadiscoursal
	Processing capacities	· Language-related memory
Adaptive-Strategic Efforts	Developmental orientation	· L2 selves · Metacognitive awareness
	Behavioral adaptation	· Strategic learning

Table 3. Example Items for Sensory-Mode and Functional Preferences

Category	Initial Sample Items
Spoken language	· I feel more anxious when speaking English than when writing it. · I can concentrate better when listening to English than when reading it.
Written language	· I am more confident in writing English than in speaking it.
Receptive processing of language	· If there is no major issue in conveying meaning, I don't think differences in expression are very important · I believe that accurately conveying meaning is more important than using figurative or implicit expressions
Productive-creative use of language	· I try to memorize many commonly used English expressions. · I like finding or creating new English expressions.
Ideational-intellectual function	· I pay close attention when I see sentences with complex structures.
Interpersonal-practical function	· I like practicing real-life conversations more than studying English vocabulary or grammar. · I can learn English better in real-life situations than in the classroom

The second section, which aimed to tap into students' sensitivities and capacities in L2 processing, was based on the notions of linguistic competence or aptitude (see Table 4). Unlike the view of language aptitude as a unified construct, it was hypothesized that a learner's sensitivities in each component ability might differ considerably. It assumed three modules of linguistic sensitivity: auditory-articulatory, lexicogrammatical, and pragmadiscoursal, and two crucial aspects of language-related memory: phonological working memory and form-meaning-use associations in long-term memory.

Table 4. Example Items for Sensitivities and Capacities in Language Processing

Category	Initial Sample Items
Aural-articulatory sensitivity	· I can detect differences in speech sounds that others often fail to distinguish. · I pay attention to weakly pronounced words as much as to clearly spoken ones. · There are often times when English speakers don't understand my pronunciation.
Lexicogrammatical sensitivity	· Sometimes, I feel that there are too many words with similar meanings in English. · When using an English word, I also think of its antonyms and synonyms. · As long as the meaning is conveyed well, grammatical mistakes are not a big problem. · Regardless of who I am speaking to, I pay close attention to avoid making mistakes.
Pragmadiscoursal sensitivity	· I tend to think in advance about how to start and continue a conversation in English. · I sometimes feel that the flow of a text is rough and not carefully interwoven.
Language-related memory	· I am confident in accurately repeating an English sentence that consists of ten or more words. · When speaking English, I find it difficult to stop Korean from coming to mind.

The third section, adaptive-strategic efforts for language development, probed learners' self-regulated and strategic engagement in developing their language ability (see Table 5 for sample items). It is assumed that learners who can recognize and adapt their inherent tendencies are better able to enhance their strengths or compensate for any lack of competence. They are also presumably better able to notice gaps in their learning processes and try to apply more effective strategies to address them (Oxford 2017). This section thus included self-awareness and the behavioral modification following from it.

Table 5. Example Items for Adaptive-Strategic Efforts for L2 Development

Category	Initial Sample Items
Strategic efforts	· I believe that knowing a lot of words will naturally improve my English skills. · I know quite a lot about the learning methods used by people who are good at English. · I don't want to try learning methods that don't suit me. · I prefer fun learning methods, even if they are slightly less effective, over boring ones.
Metalinguistic awareness	· When speaking English, I tend to notice my grammar mistakes right away. · Sometimes, I feel embarrassed when I look at English writing I did in the past. · I can easily spot English sentences that can be interpreted in two different ways. · I am aware of my weaknesses in learning English and focus on improving them.

The scale employed a 5-point Likert format, ranging from 'strongly disagree' to 'strongly agree,' with participants instructed to respond based on the degree to which each statement reflected their learning characteristics.

3.2 Interviews and Expert Panel Reviews

With the initial pool of items, pilot testing and interviews were carried out with 32 Korean EFL students. Each question attempted to stimulate interviewees' recognition of and reactions to challenges that they encountered in the course of learning English. Interviewees were also asked to check the question items for clarity in meaning, applicability of response options, and time taken to complete them. The researchers then worked with two experts to analyze the interviewees' responses. The final operative version of the instrument was derived through the following three stages.

First, we conducted a pilot testing and cognitive interviews. The initial pool of 60 items was subjected to pilot testing and interviews with Korean EFL students ($N = 32$). The primary objective was to evaluate the face validity, comprehension, and psychological resonance of the items. Students were asked to evaluate each item based on

three explicit criteria: semantic clarity (i.e. absence of ambiguity or double-barreled phrasing), applicability of response options, and practical feasibility (e.g. time taken to complete). During this phase, items were not eliminated but substantially revised to better surface the students' recognition of learning challenges. For instance, idioms or complex structures that caused cognitive fatigue or misinterpretation among lower-proficiency students were simplified to ensure uniform comprehension across all proficiency levels.

Second, following the qualitative refinement of the 60 items, a panel of five experts—comprising professors in TESOL/Applied Linguistics and specialists in psychometric scale development—evaluated the item pool. The experts rated each item on a 5-point Likert scale across three core parameters: Content Relevance, Diagnostic Appropriateness, and Technical Flawlessness. To ensure a transparent and systematic item-reduction process, two statistical thresholds were employed for item retention. First, items were required to achieve a mean rating of 3.5 or above on a 5-point scale. This ensured that every retained item possessed strong content relevance to the constructs of EFL learner characteristics. Second, to ensure consensus among judges, the inter-rater agreement index was calculated. Items with an IRA lower than 0.50 were eliminated due to a lack of conceptual consensus among the experts.

Finally, the application of these elimination criteria resulted in the removal of 35 items from the original 60-item pool. Specifically, 17 items were eliminated due to low mean ratings (< 3.5), 11 items were discarded due to low inter-rater agreement (< 0.50), and 7 items were further removed due to conceptual redundancy during the expert consensus meetings. At this stage, the macro-level tripartite constructs were elaborated into six meso-level constructs. This filtering process yielded a final version consisting of 25 items, as shown in Table 6.

Table 6. A Classroom Version of the Questionnaire

Category	Item
Listening-pronunciation	1. I can easily recognize and remember the names of English speakers. 2. I can detect differences in speech sounds that others often fail to distinguish. 3. I pay attention to weakly and quickly pronounced words as much as to clearly spoken ones. 4. I can accurately mimic the intonation and pitch variations of English speakers.
Lexico-grammar	1. I have hesitated for a long time while writing to find the perfect word I like the most. 2. When using an English word, I also think of its antonyms and synonyms. 3. I pay close attention to make sure there are no errors in my speech or writing.
Pragma-discourse	1. I tend to think in advance about how to start and continue a conversation in English. 2. When reading English, I sometimes feel that the flow of the text is rough and not smooth. 3. When listening to or reading English, I sometimes vividly feel the author's emotions and train of thought. 4. I have felt upset during an English conversation because of the other person's way of expressing thoughts.
Verbal memory	1. When listening to English, I try to remember the parts I didn't catch well and go back to them later. 2. I am confident in accurately repeating an English sentence that consists of 10 or more words. 3. When reading English aloud, I can interpret the meaning at the same time. 4. Among the English expressions I remember, some consist of more than 10 words. 5. For the English words I know, I can name at least three synonyms or antonyms. 6. I tend to remember spoken words better than written text.
Metalinguistic awareness	1. I am aware of how my pronunciation differs from that of fluent speakers. 2. When speaking English, I tend to notice my grammar mistakes right away. 3. Sometimes, I feel embarrassed when I look at English writing I did in the past. 4. I often spot mistakes in the speech and writing of English speakers. 5. I can easily identify English sentences that can be interpreted in two different ways.
Developmental efforts	1. I am aware of my weaknesses in learning English and focus on improving them. 2. I have tried using my own English learning methods instead of following those suggested by others. 3. I know quite a lot about the learning methods used by people who are good at English.

While the preliminary conceptualization framed learner characteristics around three broad domains, the empirically derived six-factor structure reflects a transition from macro-level heuristic categories to more specific and skill-sensitive competencies. Constructs such as Metalinguistic Awareness and Developmental Efforts are assumed to bridge language processing capacities and self-regulatory mechanisms, suggesting that individual variation operates across an integrated continuum rather than static categories.

3.3 Quantitative Validation

A total of 236 students were recruited from public middle schools; 63 were male (26.7%) and 173 were female (73.3%). The data from those who were unresponsive or provided insincere answers were excluded from analysis ($N = 20$). The questionnaire was presented in Korean. The dimensionality of the measures and the functionality of individual items were examined with the following questions: (1) To what extent does the proposed instrument fit the Rasch model for measurement? (2) Does the instrument demonstrate validity when it undergoes a confirmatory factor analysis?

Accordingly, two statistical methods were employed: First, subfactors of language learning traits were distinguished, and a rating scale analysis was conducted for each subfactor. The analysis was conducted using the Winstep 5.7.1. After the unidimensionality assumption was tested, a principal component analysis of the standardized residuals followed with respect to each subfactor. The unidimensionality assumption is considered satisfied if the eigenvalue of the first contrast is less than 3. To verify the degree to which the data deviate from the Rasch model's expected unidimensionality, outfit and infit statistics were examined. The acceptable range for fit mean square (*MNSQ*) was set based on the guidelines proposed by Linacre (2017), which suggests a range of 0.5 to 1.5 for rating scales. A point-biserial correlation coefficient threshold of 0.4 or higher was applied. If the MNSQ values fall within this range, the items are considered fitted to the model.

Following the rating scale analysis for each subfactor, a confirmatory factor analysis (CFA) was performed using AMOS 23. Model fit indices were evaluated using the following criteria: the root mean square error of approximation (*RMSEA*) was required to be less than or equal to 0.08, the comparative fit index (*CFI*) was required to be greater than or equal to 0.90, and the standardized root mean square residual (*SRMR*) was required to be less than or equal to 0.08. Finally, construct reliability (*CR*) and average variance extracted (*AVE*) were calculated to assess the alignment between subfactors and their respective measurement variables. Values of *CR* of 0.7 or higher and *AVE* of 0.5 or higher were considered indicative of an acceptable level of construct validity.

While Rasch analysis and CFA stem from different measurement traditions, their combined use provides a more rigorous validation than either method alone. We employed Rasch analysis to ensure the item-level requirements of fundamental measurement, specifically looking at item fit, category functioning, and the invariance of item difficulties. Subsequently, CFA was utilized to verify the higher-order structural validity of the latent constructs. This complementary approach is supported by contemporary psychometricians (e.g. Khine 2013) as it ensures that the instrument is both internally consistent at the item level and structurally representative of the theoretical framework at the construct level.

4. Results

4.1 Rasch Analysis

To examine the psychometric properties of the questionnaire, the reliability and separation indices for both respondents and items were analyzed using the Rasch model. These indices are critical indicators for evaluating the validity of a measurement tool, providing insights into how well the instrument discriminates varying levels of ability. The results indicated that the person reliability was 0.91, suggesting that the respondents' dispositions were measured with a high degree of consistency. A reliability coefficient of 0.80 or above is generally considered indicative of a well-functioning instrument. The person separation index was 3.12, which implies that the instrument can effectively stratify respondents into approximately three distinct levels.

The item reliability was found to be 0.92, reflecting a high degree of consistency in how the items assess the respondents' dispositions. It indicates that the items are reliably measuring the intended construct. The item separation index was 3.29, suggesting that the items were well distinguished by their varying levels. A separation index of 2 or more is considered adequate for differentiating item difficulty. To further confirm the unidimensionality of the Rasch model, a principal component analysis of standardized residuals was conducted. The eigenvalue of the first contrast was 2.31, which is below the critical threshold of 3.0, thereby supporting the assumption of unidimensionality for the model. This finding supports that the Rasch model employed in this study is valid and that the instrument appropriately captures a single underlying latent trait. Thus, the high reliability and separation indices for both persons and items, alongside the confirmation of unidimensionality, demonstrate that the tool exhibits robust psychometric qualities, ascertaining its validity in assessing the intended construct.

Subsequently, item fit was analyzed using the Rasch model, employing indicators such as Measure, Standard Error (*SE*), Infit and Outfit Mean Square (*MNSQ*), and Pearson-Measure Correlation (*PMC*) values. These indicators were used to assess item difficulty, measurement precision, response pattern consistency, and the relationship between item functioning and respondent ability. The Measure values provide an indication of where each item is positioned along the respondent ability continuum, and in this analysis, they were distributed evenly between -0.59 and 0.69. This distribution suggests that the items were appropriately calibrated to reflect a range of respondent abilities. For instance, item 1 had a Measure value of -0.12, while item 24 had a Measure value of 0.46, illustrating that item difficulty was adequately distributed across the ability spectrum. The SE values ranged from 0.09 to 0.10, indicating that the measurement precision for each item was generally high.

Infit and Outfit *MNSQ* values, which are critical for assessing item fit, were also examined. The Infit *MNSQ* reflects the degree to which observed responses align with expected responses based on respondent ability, while the Outfit *MNSQ* is more sensitive to unexpected or extreme response patterns. Most items in the analysis fell within the acceptable *MNSQ* range; however, item 1 displayed an Infit *MNSQ* value of 1.64 and an Outfit *MNSQ* value of 1.59, both of which exceed the ideal range, suggesting that this item may have functioned in a manner inconsistent with the model's expectations. The *PMC* values, which indicate the correlation between item performance and respondent ability, were also evaluated. A *PMC* value of 0.4 or higher is generally considered indicative of adequate item discrimination. Most items in this study exceeded this threshold, demonstrating that they adequately discriminated between respondents of different ability levels. However, item 1 recorded a *PMC* value of 0.41, which, while meeting the minimum threshold, suggests a weaker correlation between item performance and respondent ability.

An analysis of the subfactors revealed that listening/pronunciation-related items, particularly items 1 and 11, generally exhibited higher Infit and Outfit *MNSQ* values, indicating greater variability in response patterns and

less stable functioning. Item 1, in particular, warrants further investigation due to its elevated *MNSQ* values. However, item 11, despite showing somewhat elevated *MNSQ* values, satisfies the recommended criterion that Infit and Outfit *MNSQ* values should be between 0.5 and 1.5. Therefore, while item 11's fit indices are higher than ideal, they do not indicate any significant issues, and the item can be considered to perform adequately. In contrast, item 1 exceeds these thresholds and may require further revision to improve its fit. Table 7 presents the results for each item.

Item 23 exhibited marginal fit, with Infit (1.43) and Outfit (1.37) statistics reaching the upper threshold of the acceptable range. While these values suggest a degree of measurement noise, the item remained productive for the construct. Furthermore, it demonstrated the strongest discrimination in the set, as evidenced by a *PMC* of 0.55, indicating that it effectively distinguishes between high- and low-ability respondents despite the observed variance.

Table 7. Rasch Item Statistics for the Proposed Questionnaire

Subfactors	Item no.	Measure	*SE*	Infit *MNSQ*	Outfit *MNSQ*	*PMC*
Listening Pronunciation	1	-0.12	0.1	1.64	1.59	0.41
	11	0.69	0.09	1.2	1.46	0.43
	23	-0.29	0.1	1.43	1.37	0.55
	25	-0.42	0.1	1.18	1.37	0.49
Lexicogrammar	2	-0.09	0.1	1.32	1.3	0.57
	12	-0.05	0.1	1.29	1.28	0.53
	20	0.51	0.09	1.14	1.29	0.52
Pragmadiscourse	3	-0.16	0.1	1.19	1.23	0.54
	22	-0.59	0.1	1.2	1.14	0.53
	19	0.26	0.09	1.18	1.18	0.64
	4	-0.59	0.1	1.12	1.08	0.49
Verbal Memory	5	0.22	0.09	0.98	0.92	0.67
	13	-0.28	0.1	0.89	0.85	0.64
	18	0.33	0.09	0.87	0.88	0.67
	6	-0.1	0.1	0.87	0.85	0.66
	17	-0.36	0.1	0.86	0.82	0.68
	16	-0.05	0.1	0.85	0.82	0.67
Metalinguistic Awareness	7	0.15	0.09	0.83	0.8	0.66
	14	0.36	0.09	0.77	0.8	0.64
	8	-0.07	0.1	0.75	0.73	0.66
	21	0.34	0.09	0.74	0.74	0.69
	24	0.46	0.09	0.72	0.7	0.70
Developmental Effort	9	-0.27	0.1	0.65	0.63	0.70
	15	0.39	0.09	0.62	0.61	0.74
	10	-0.26	0.1	0.6	0.57	0.71

MNSQ: mean square, PMC: point measure correlation

Table 8. Analyses of the Rating Scale Response Categories

Category	%	Observed Average	Expected Average	Infit *MNSQ*	Outfit *MNSQ*	Threshold
1	8%	-1.67	-1.63	1.01	1.07	NONE
2	17%	-0.74	-0.73	0.98	0.99	-1.97
3	53%	-0.18	-0.18	0.77	0.79	-1.59
4	17%	0.56	0.5	0.95	0.91	1.26
5	5%	1.57	1.72	1.3	1.26	2.29

In order to evaluate whether the 5-point Likert categories were functioning as intended, diagnostic analyses of the rating scale response categories were conducted. The results of the diagnostic analyses are presented in Table 8. Both the average measures and threshold estimates increased progressively across the categories, indicating that the categories were operating appropriately. The Infit and Outfit *MNSQ* statistics for each category were all within the range of ±0.7 to 1.3, suggesting that the response patterns were consistent with model expectations and that the categories were sufficiently distinct from one another in terms of how they captured respondents' varying dispositions, as shown in Table 8.

Figure 1 illustrates the category probability curve, which visually represents the likelihood of selecting a particular response category based on the relative difference between person trait scores and item difficulties.

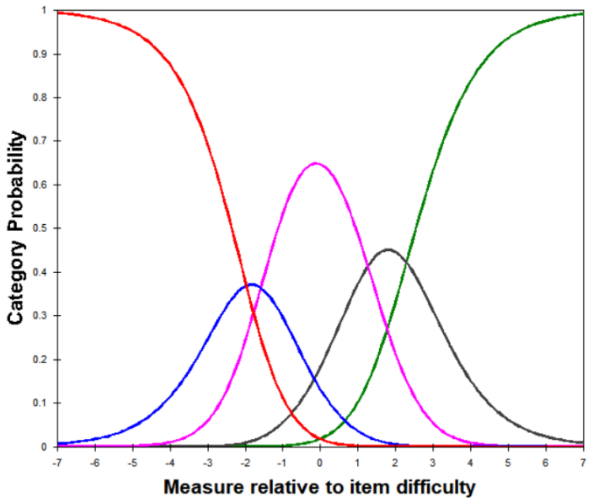

Figure 1. Category Probability Curves for the Rating Scale

The threshold estimates correspond to the intersection points where one category's probability curve meets the curve for the adjacent category. This point marks the transition between the likelihoods of selecting one category over the next. The hill-shaped curves for each category indicate that all categories were utilized as expected, with distinct peaks, further affirming that respondents were able to distinguish between the levels of agreement.

Importantly, in an instrument that incorporates multiple psychological dimensions, the latent continuum evaluated by the category probability curves does not represent a monolithic, performance-based ability score. Instead, it signifies the overall level of manifestation of the learner's self-perceived profile of inclination for and sensitivity to L2 learning. An endorsement of "(Strongly) Agree" therefore indicates the learner's cognizance of these language-related abilities and strategic learning behaviors, whereas "(Strongly) Disagree" reflects their

perceived deficiency or absence. The category probability curves thus demonstrated that all five categories transitioned monotonically, with distinct peaks for each category and ordered threshold calibrations. Rather than suggesting that these categories partition students into absolute proficiency brackets, the results add empirical support for the psychometric usability and validity of the 5-point configuration. The ordered thresholds confirm that the target population of Korean secondary EFL learners could meaningfully distinguish between adjacent response choices. That is, although the scale aggregates distinct psychological domains, the category functioning analysis suggests that the 5-point Likert scale provides a psychometrically sound, orderly continuum suitable for capturing the multi-faceted profile of the learners.

4.2 Confirmatory Factor Analysis

Subsequently, a confirmatory factor analysis was conducted to examine three hypothetical factor models, with 24 items after removing item 1 due to low item fit in the Rasch analysis. The three models aimed to explore different factor structures for categorizing language-related sensitivity and capacity factors. The first model was a two-factor model, where the items were grouped into two primary factors: Sensitivity and Cognition. Under this structure, the Sensitivity factor comprised three subfactors: Listening-Pronunciation, Lexicogrammar, and Pragmadiscourse, while the Cognition factor included Verbal Memory, Metalinguistic Awareness, and Developmental Effort. The second model tested was a four-factor model, where the three subfactors of Listening-Pronunciation, Lexicogrammar, and Pragmadiscourse were grouped as a single Sensitivity factor, while the remaining cognitive subfactors were retained as separate factors. Finally, the third model, a full six-factor model, explored the most detailed factor structure. In this configuration, each of the six factors, Listening-Pronunciation, Lexicogrammar, Pragmadiscourse, Verbal Memory, Metalinguistic Awareness, and Developmental Effort, was treated as separate factors.

The decision to test three competing models was based on a theory-driven, nested-model comparison strategy. This approach is recommended in instrument validation to determine the most parsimonious and representative factor structure. More specifically, the two-factor model tested the hypothesis of broad domain-general traits; the six-factor model tested the hypothesis of maximum granularity; and the four-factor model explored a hybrid structure. By comparing the fit indices across these three models, we were able to provide empirical evidence for the specific level of multidimensionality that best characterizes the Korean EFL learner profile, rather than simply assuming a structure *a priori*.

Upon conducting the CFA, the model fit indices revealed variations in fit across the three models. The two-factor model yielded a chi-square statistic of $\chi^2(256) = 558.128$, $p < .001$, with additional fit indices of $CFI = .864$, $RMSEA = .075$, and $SRMR = .057$. For the four-factor model, the chi-square statistic was $\chi^2(245) = 538.129$, $p < .001$, with fit indices showing an improvement: $CFI = .870$, $RMSEA = .074$, and $SRMR = .0547$. Lastly, the six-factor model demonstrated the best fit among the three, with a chi-square value of $\chi^2(233) = 435.102$, $p < .001$, alongside $CFI = .910$, $RMSEA = .064$, and $SRMR = .053$. Comparing the fit indices of the three models, it was evident that the six-factor model outperformed the other two models, exhibiting a higher *CFI* and relatively lower values for both the *RMSEA* and the *SRMR*. Consequently, the six-factor model was chosen as the preferred model in this study, with further analyses conducted on its coefficients, validity, and reliability to explore its robustness in capturing the intra-learner difference factors effectively. The results are presented in Table 9.

The unstandardized coefficients (*B*) for all items within each factor were statistically significant at the $p < .001$ level, indicating that each factor demonstrated a substantial degree of explanatory power in relation to the measured variables. Analyzing the construct reliability (*CR*) for each factor, the Listening-Pronunciation factor had a *CR*

value of .73, meeting the recommended threshold of .70. The Lexicogrammar factor showed a *CR* value of .65, which did not meet the threshold but approached it, indicating an acceptable level of reliability. The Pragmadiscourse factor had a *CR* of .59, which also fell below the threshold, though it demonstrated a moderate level of reliability. In contrast, the Verbal Memory factor displayed a high level of reliability with a *CR* value of .79, meeting the threshold of .70, while the Metalinguistic Awareness factor also exceeded the threshold with a *CR* of .74. Finally, the Developmental Effort factor showed a *CR* of .73, indicating adequate reliability.

Regarding the Average Variance Extracted (*AVE*), the Lexicogrammar, Pragmadiscourse, and Metalinguistic Awareness factors had *AVE* values of .39, .33, and .38, respectively, which are below the recommended threshold of .50, indicating a lower level of convergent validity. The Listening/Pronunciation, Verbal Memory, and Developmental Effort factors had *AVE* values of .47, .40, and .47, respectively, which also did not reach the threshold but were within an acceptable range, suggesting a satisfactory level of validity. These results indicate that while each factor generally met the threshold for construct reliability, some factors did not meet the *AVE* threshold, indicating the need for improvement in convergent validity.

Table 9. Validity Indices for the Six-Factor Model of Linguistic Sensitivities

Subfactor	Item	*B*	*SE*	*β*	*t*	Validity
Listening Pronunciation	2	1.00	0.10	0.70	9.87	CR=.73 AVE=.47 *α*=.72
	3	0.99	0.10	0.67	9.50	
	4	1.00		0.69		
Lexicogrammar	1	1.19	0.16	0.66	7.53	CR=.65 AVE=.39 *α*=.59
	2	0.99	0.16	0.51	6.29	
	3	1.00		0.55		
Pragmadiscourse	1	1.47	0.25	0.63	5.84	CR=.59 AVE=.33 *α*=.58
	2	1.02	0.21	0.42	4.82	
	3	1.78	0.30	0.65	5.89	
	4	1.00		0.41		
Verbal Memory	1	0.56	0.09	0.44	6.15	CR=.79 AVE=.40 *α*=.79
	2	0.87	0.11	0.55	7.68	
	3	1.03	0.10	0.72	10.10	
	4	0.94	0.09	0.76	10.53	
	5	0.87	0.11	0.55	7.77	
	6	1.00		0.71		
Metalinguistic Awareness	1	0.73	0.09	0.55	7.94	CR=.74 AVE=.38 *α*=.73
	2	1.00	0.09	0.74	11.20	
	3	0.48	0.10	0.35	4.87	
	4	0.72	0.09	0.54	7.84	
	5	1.00		0.79		
Developmental Effort	1	1.05	0.11	0.72	9.37	CR=.73 AVE=.47 *α*=.72
	2	1.253	0.139	0.68	8.989	
	3	1		0.66		

To evaluate the structural validity of the questionnaire, a confirmatory factor analysis was conducted on the 24 retained items (following the exclusion of Item 1 due to poor Rasch fit). The testing of competing models was

fundamentally guided by the study's initial theoretical framework, which conceptualizes learner traits into three overarching macro-constructs: Language Learning Preferences, Sensitivities/Capacities, and Adaptive-Strategic Efforts. To empirically test whether these macro-constructs should be operationalized as broad, aggregated dimensions or as highly differentiated, granular factors, three nested models were compared.

The first model (2-factor) tested a broad, amalgamated structure where items were collapsed into two macro-level traits: Sensitivity (Listening-Pronunciation, Lexicogrammar, and Pragmadiscourse) and Cognition (encompassing Verbal Memory, Metalinguistic Awareness, and Developmental Effort). The second model (4-factor) explored a hybrid structure, maintaining Sensitivity as a single macro-construct while treating the remaining cognitive and strategic dimensions as separate factors. Finally, the third model (6-factor) treated all six underlying subfactors as distinct dimensions. This 6-factor structure most closely aligns with the original conceptual framework by positing Sensitivities (Listening-Pronunciation, Lexicogrammar, Pragmadiscourse), Capacities (Verbal Memory, Metalinguistic Awareness), and Efforts (Developmental Effort) as discrete components within those macro-domains. Upon conducting the CFA, the fit indices revealed variations across the three models.

The 2-factor model yielded a poor fit, $\chi^2(256) = 558.128$, $p < .001$, $CFI = .864$, $RMSEA = .075$, $SRMR = .057$. The 4-factor model demonstrated minor improvements, $\chi^2(245) = 538.129$, $p < .001$, $CFI = .870$, $RMSEA = .074$, $SRMR = .055$. The 6-factor model demonstrated the most optimal fit among the three configurations, $\chi^2(233) = 435.102$, $p < .001$, $CFI = .910$, $RMSEA = .064$, $SRMR = .053$. To determine if the descriptive superiority of the 6-factor model was statistically meaningful, a series of chi-square difference tests for nested models was conducted. The 6-factor model demonstrated a statistically significant improvement in fit over both the 2-factor model ($\Delta\chi^2(23) = 123.03$, $p < .001$) and the 4-factor model ($\Delta\chi^2(12) = 103.03$, $p < .001$). Therefore, the 6-factor structure was selected as the most empirically supportable framework among the tested alternatives.

As shown in Table 9, all unstandardized factor loadings were statistically significant. However, the psychometric properties of the 6-factor model suggest localized areas of instability that warrant cautious interpretation. Regarding Construct Reliability, while Verbal Memory (.79), Metalinguistic Awareness (.74), Listening-Pronunciation (.73), and Developmental Effort (.73) surpassed the standard recommended threshold of .70, the Lexicogrammar (.65) and Pragmadiscourse (.59) factors fell short, indicating lower internal consistency within these domains. Furthermore, the AVE values across all six factors (.33 to .47) fell below the standard ideal threshold of .50. This indicates that a substantial portion of the variance is captured by measurement error rather than the underlying constructs, revealing a limitation in convergent validity. Rather than demonstrating a robust validation of the 6-factor structure, these data indicate that while the 6-factor model is the most statistically optimal and parsimonious configuration among the tested alternatives, the construct dimensions retain some instability. These psychometric limitations should be taken into consideration when interpreting diagnostic profiles, and they reveal critical targets for future item refinement.

Finally, a potential conceptual tension may exist between the unidimensionality established via the Rasch analysis and the multi-dimensional structure identified in the six-factor CFA. However, these findings may not be mutually exclusive. The unidimensionality confirmed in the Rasch analysis (via PCA of residuals) demonstrates that all 24 items fundamentally map onto a single underlying latent trait: a global propensity toward EFL learning. This suggests that the questionnaire measures a coherent global trait without confounding statistical noise or irrelevant secondary dimensions. Every learner can thus be located on this macro-continuum based on their overall linguistic disposition. On the other hand, the six-factor structure confirmed by the CFA demonstrates how this overarching L2 learning-related trait differentiates into more specialized, observable dimensions in actual learning contexts. Statistically, this relationship operates as a higher-order hierarchical structure, wherein the macro-

construct manifests through six meso-constructs or subfactors (Listening-Pronunciation, Lexicogrammar, Pragmadiscourse, Verbal Memory, Metalinguistic Awareness, and Developmental Effort). It can thus be said that the latent construct of this study is neither a fragmented collection of unrelated traits nor a completely undifferentiated flat structure. The Rasch analysis supports that the scale measures a unified latent continuum while the CFA provides the psychometric justification for segmenting that continuum into six more specialized sub-constructs. This dual-analytical approach may be better suited for a diagnostic questionnaire, as it secures macro-level construct validity while simultaneously providing the meso-level granularity required to produce characteristic profiles for individual EFL learners.

To sum up, the factors in this study generally satisfied the construct reliability criteria, with the Listening-Pronunciation, Verbal Memory, and Developmental Effort factors exhibiting high levels of reliability. The Lexicogrammar, Pragmadiscourse, and Metalinguistic Awareness factors had relatively lower *AVE* values. This notwithstanding, they are still considered to be within an acceptable range for the goal of this study.

5. Conclusion

This study developed a diagnostic questionnaire to assess Korean EFL students' varying dispositions toward English learning. The psychometric properties of the questionnaire were evaluated to establish its construct validity using the Rasch rating scale analysis and the confirmatory factor analysis. The Rasch analysis confirmed the reliability and validity of the instrument in differentiating respondent types and item functions. However, certain subfactors, such as Lexicogrammar and Pragmadiscourse, exhibited lower *AVE* values, indicating that refining it to enhance the convergent validity of subfactors remains a critical task.

This study constitutes the first step towards a systematic investigation into the interaction of within- and between-individual factors in EFL contexts. It provides an initial model to explore the extent to which students' linguistic sensitivities and self-regulated efforts correlate with their L2 proficiency achievement, which facet of learner differences has more to do with imbalanced L2 development, styles and preferences, linguistic sensitivities and capacities, or adaptive-strategic efforts—or their close interactions—and what combinations of learner internal and external factors predict their balanced development in L2 proficiency. The study thus can serve as a foundation for diagnosing Korean EFL students' needs, raising their awareness, and tailoring interventions. Another contribution of this study lies in proposing a practical framework to assess Korean students' language learning traits, not simply relying on the established scales that may not adequately capture EFL students' learning experiences. The proposed questionnaire can be implemented as a classroom diagnostic tool at the beginning of a semester to map students' profiles in language sensitivity and self-regulated learning. Teachers can then design and tailor tasks, assisting students with varying linguistic sensitivities while providing scaffolding for those who require adaptive-strategic support.

To conclude, as the conception of language as a universal competence, rather than as a collection of specific abilities, dispositions, and experiences that vary within, as well as across, individuals, has prevailed in contemporary linguistics and its associated fields, individuals' imbalanced development in L2 skills (i.e. to what extent and why learners show different abilities in different areas of the language that they are learning) has received limited attention in mainstream research. Imbalanced development may be a precursor to, or a partial representation of, fossilization. This state of premature cessation occurs as learners become habituated to processing restricted input types, effectively narrowing the scope of their further L2 acquisition. Further questions arise of whether learner maturation is related to the degree of variation in learner traits within and across learners,

and whether learner maturation interacts with learner traits in influencing the likelihood of imbalanced L2 development. The current practice of English language teaching in Korea does not emphasize individual learners' relative strengths or preferred ways of using an L2 for their personal goals. This learning environment may further restrain non-matching students' willingness to learn and communicate with the target language, a perspective that can advance discussions on how imbalanced learning environments interact with learner motivation. Future research therefore should explore the longitudinal change in factors of imbalanced development and its impact on fossilization and achievement over time. If this perspective is on the right track, the characteristic phenomena of imbalanced development in EFL contexts can be attributed not just to individual differences in psychological dispositions, but to their interaction with the learning context, which may, in fact, constrain students' development.

References

Andreou, E., G. Andreou and F. Vlachos. 2008. Learning styles and performance in second language tasks. *TESOL Quarterly* 42(4), 665-674.

Banaruee, H., D. Farsani and O. Khatin-Zadeh. 2022. EFL learners' learning styles and their reading performance. *Discover Psychology* 2, 45.

Carroll, J. B. and S. M. Sapon. 1959. *The Modern Language Aptitude Test* (MLAT). New York: The Psychological Corporation.

Castro, O. and V. Peck. 2005. Learning styles and foreign language learning difficulties. *Foreign Language Annals* 38(3), 401-409.

Choe, M. H. 2022. Profiling second language learners' linguistic sensitivity through a self-reporting questionnaire. *Lingua Humanitatis* 24(2), 469-497.

DeKeyser, R. 2012. Interactions between individual differences, treatments, and structures in SLA. *Language Learning* 62(Supplement 2), 189-200.

DeKeyser, R., I. Alfi-Shabtay and D. Ravid. 2010. Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics* 31(3), 413-438.

Dewaele, J. M. and L. Wei. 2013. Is multilingualism linked to a higher tolerance of ambiguity? *Bilingualism: Language and Cognition* 16(1), 231-240.

Dörnyei, Z. 2005. *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Dörnyei, Z. and E. Ushioda. 2021. *Teaching and Researching Motivation*. New York: Routledge.

Doughty, C. J. 2014. Assessing aptitude. In A. Kunnan, ed., *The Companion to Language Assessment*, 25-46. Hoboken, NJ: Wiley-Blackwell.

Doughty, C. J. 2019. Cognitive language aptitude. *Language Learning* 69(S1), 101-126.

Dunn, R. 1984. Learning style: State of the science. *Theory into Practice* 23(1), 10-19.

Ehrman, M. and B. L. Leaver. 2003. Cognitive styles in the service of language learning. *System* 31(3), 393-415.

Ehrman, M. E., B. L. Leaver and R. L. Oxford. 2003. A brief overview of individual differences in second language learning. *System* 31(3), 313-330.

Fleming, N. D. 2006. *V.A.R.K Visual, Aural/Auditory, Read/Write, Kinesthetic*. Christchurch, New Zealand: Bonwell Green Mountain Falls.

Garrett, P. 2010. *Attitudes to Language*. Cambridge: Cambridge University Press.

Gohar, M. J. and N. Sadeghi. 2015. The impact of learning style preferences on foreign language achievement: A case study of Iranian EFL students. *Procedia-Social and Behavioral Sciences* 171, 754-764.

Granena, G. and M. H. Long. 2013. Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research* 29(2), 311-343.

Grey, S., J. N. Williams and P. Rebuschat. 2015. Individual differences in incidental language learning: Phonological working memory, learning styles, and personality. *Learning and Individual Differences* 38, 44-53.

Halliday, M. A. K. and C. M. I. M. Matthiessen. 2014. *Halliday's Introduction to Functional Grammar*. New York: Routledge.

Hatami, S. 2013. Learning styles. *ELT Journal* 67(4), 488-490.

Isemonger, I. and C. Sheppard. 2007. A construct-related validity study on a Korean version of the perceptual learning styles preference questionnaire. *Educational and Psychological Measurement* 67(2), 357-368.

Jackson, D. O. 2020. Working memory and second language development: A complex, dynamic future? *Studies in Second Language Learning and Teaching* 10(1), 89-109.

Joe, H. K., P. Hiver and A. H. Al-Hoorie. 2017. Classroom social climate, self-determined motivation, willingness to communicate, and achievement: A study of structural relationships in instructed second language settings. *Learning and Individual Differences* 53, 133-144.

Kautto, A., H. Railo and E. Mainela-Arnold. 2024. Introducing the intra-individual variability hypothesis in explaining individual differences in language development. *Journal of Speech, Language, and Hearing Research* 67(8), 2698-2707.

Kersten, K., C. Schelletter, A. C. Bruhn and K. Ponto. 2021. Quality of L2 input and cognitive skills predict L2 grammar comprehension in instructed SLA independently. *Languages* 6(3), 1-19.

Khine, M. S. 2013. *Application of Structural Equation Modeling in Educational Research and Practice*. Leiden, The Netherlands: Brill.

Koizumi, R., T. Agawa, K. Asano and Y. In'nami. 2022. Skill profiles of Japanese English learners and reasons for uneven patterns. *Language Testing in Asia* 12(1), 1-34.

Kolb, D. A. 1999. *The Kolb's Learning Styles Inventory*. Boston, MA: Hay Group.

Lantolf, J. P. and T. G. Beckett. 2009. Sociocultural theory and second language acquisition. *Language Teaching* 42(4), 459-475.

Larsen-Freeman, D. and L. Cameron. 2008. *Complex Systems and Applied Linguistics*. Oxford: Oxford University Press.

Larsen-Freeman, D. and M. H. Long. 2014. *An Introduction to Second Language Acquisition Research*. New York: Routledge.

Li, S. 2015. The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics* 36(3), 385-408.

Li, S. and R. DeKeyser. 2021. Implicit language aptitude: Conceptualizing the construct, validating the measures, and examining the evidence. *Studies in Second Language Acquisition* 43(3), 473-497.

Li, S., P. Hiver and M. Papi. 2022. *The Routledge Handbook of Second Language Acquisition and Individual Differences*. New York: Routledge.

Linacre, J. M. 2017. *A User's Guide to WINSTEPS Rasch-Model Computer Programs*. Chicago, IL: Winsteps.com.

Linck, J. A., M. M. Hughes, S. G. Campbell, N. H. Silbert, M. Tare, S. R. Jackson and C. J. Doughty. 2013. Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning* 63(3), 530-566.

Long, M. H. 2013. Maturational constraints on child and adult SLA. In G. Granena and M. H. Long, eds., *Sensitive*

Periods, Language Aptitude, and Ultimate L2 Attainment, 3-41. Amsterdam: John Benjamins.

Meara, P. 2005. *LLAMA Language Aptitude Tests*. Swansea, UK: Lognostics.

Ortega, L. 2019. SLA and the study of equitable multilingualism. *The Modern Language Journal* 103(1), 23-38.

Oxford, R. L. 2017. *Teaching and Researching Language Learning Strategies: Self-Regulation in Context*. New York: Routledge.

Pawlak, M. 2022. Research into individual differences in SLA and CALL: Looking for intersections. *Language Teaching Research Quarterly* 31, 200-233.

Reid, J. M. 1995. *Learning Styles in the ESL/EFL Classroom*. Boston, MA: Heinle & Heinle Publishers.

Robinson, P. 2013. Individual differences, aptitude complexes, SLA processes, and aptitude test development. In M. Pawlak, ed., *New Perspectives on Individual Differences in Language Learning and Teaching*, 57-75. Berlin: Springer-Verlag.

Robinson, P. 2019. Aptitude and second language acquisition. In C. Chapelle, ed., *The Concise Encyclopedia of Applied Linguistics*, 129-133. Oxford: Blackwell.

Shen, M. Y. and T. Y. Chiu. 2019. EFL learners' English speaking difficulties and strategy use. *Education and Linguistics Research* 5(2), 88-102.

Truscott, J. and M. Sharwood Smith. 2019. *The Internal Context of Bilingual Processing*. Amsterdam: John Benjamins.

Unsworth, S., L. Persson, T. Prins and K. de Bot. 2014. An investigation of factors affecting early foreign language learning in the Netherlands. *Applied Linguistics* 36(5), 527-548.

Ushioda, E. 2011. Language learning motivation, self and identity: Current theoretical perspectives. *Computer Assisted Language Learning* 24(3), 199-210.

VanPatten, B., M. Smith and A. G. Benati. 2020. *Key Questions in Second Language Acquisition: An Introduction*. Cambridge: Cambridge University Press.

Wen, Z. E., P. Skehan and R. L. Sparks. 2023. *Language Aptitude Theory and Practice*. Cambridge: Cambridge University Press.

Wong, L. L. and D. Nunan. 2011. The learning styles and strategies of effective language learners. *System* 39(2), 144-163.

Appendix. The Original Pool of Items

[Styles and Preferences in Language Use]
1. I feel more anxious when speaking English than when listening or reading.
2. I find listening to English easier and more confident than reading.
3. I can concentrate better when listening to English than when reading.
4. I feel less bored when listening to English than when reading.
5. I can concentrate better when writing in English than when listening.
6. I can focus better when speaking English than when writing.
7. I feel more bored when writing in English than when speaking.
8. I am more confident in writing English than in speaking.
9. I try to memorize many commonly used English expressions.
10. I enjoy finding or creating new English expressions.
11. I think imitating is the best way to practice speaking and writing in English.
12. There are times when I am impressed by creative and meaningful English expressions.
13. I believe that as long as the meaning is conveyed correctly, differences in expression are not very important.
14. I believe there can be multiple expressions that best convey a particular meaning.
15. I think delivering an accurate meaning is more important than using figurative or implicit expressions.
16. I find practicing conversation more enjoyable than studying English vocabulary or grammar.
17. I can focus better when having a conversation in English than when reading or writing alone.
18. I can learn English better in real-life situations than in the classroom.
19. I become interested when I see sentences with complex structures.
20. There are times when I want to know the origins and history of words that are not commonly used in daily life.
21. I sometimes think about how the principles of expression in Korean and English differ.

[Sensitivities and Capacities in Language Processing]
1. I can easily recognize and remember the names of English speakers.
2. I can detect subtle differences in speech sounds that others may not notice.
3. I pay attention to weakly and quickly pronounced words as much as to clearly spoken ones.
4. I often rely on context rather than sound to distinguish similar words in English.
5. I can accurately mimic the intonation and pitch variations of English speakers.
6. There are times when English speakers do not understand my pronunciation.
7. Sometimes, I feel that there are too many words with similar meanings in English.
8. When I use an English word, its antonyms and synonyms also come to mind.
9. I have hesitated for a long time while writing to find the perfect word.
10. I believe that as long as the meaning is clear, grammatical mistakes are not a big issue.
11. Regardless of who I am speaking or writing to, I make a great effort to avoid errors.
12. I think grammatical knowledge and communication skills are unrelated.
13. I tend to think in advance about how to start and continue a conversation in English.
14. When reading English, I sometimes feel that the flow of the text is awkward and not smooth.
15. While reading English, I sometimes vividly sense the author’s emotions and thought process.
16. In similar situations, I usually use the one English expression I feel most confident in.
17. I believe that direct expressions are always more effective than indirect ones when speaking English.
18. I have felt upset during an English conversation because of how the other person expressed something.
19. I tend to remember spoken words better than written text.
20. When I hear a new English word, I immediately check its spelling.
21. When listening to English, I try to remember unclear parts and review them later.

22. When I hear an English word and recall its meaning, its spelling also comes to mind.
23. I am confident that I can accurately repeat an English sentence of about 10 words.
24. When reading English aloud, I can interpret the meaning at the same time.
25. Among the English expressions I remember, some contain more than 10 words.
26. For the English words I know, I can name at least three synonyms or antonyms.
27. When speaking English, I find it difficult to stop my mind from switching back to Korean.

[Adaptive and Strategic Efforts for Language Development]
1. I believe that knowing many words will automatically improve my English ability.
2. I am aware of how my pronunciation differs from that of fluent English speakers.
3. When speaking English, I usually notice my own grammar mistakes right away.
4. I sometimes feel embarrassed when I look at my past English writing.
5. I often notice mistakes in the speech and writing of English speakers.
6. I can easily identify English sentences that have two possible interpretations.
7. I know which aspects of English I struggle with and focus on improving them.
8. I have tried using my own learning methods rather than relying on those suggested by others.
9. I am familiar with many learning techniques used by fluent English speakers.
10. I believe that the only way to improve listening and speaking is to practice them a lot.
11. I don’t want to try learning methods that don’t suit me.
12. I prefer fun learning methods, even if they are less effective than boring ones.

Examples in: English
Applicable Languages: English
Applicable Level: Secondary

