

Information-seeking behavior in second language acquisition: Conceptualization and operationalization

Mostafa Papi^{a,*}, Yuichi Suzuki^b, Claudia J. Kim^a, Miseong Kim^a

^a Florida State University, United States

^b Waseda University, Japan

ARTICLE INFO

Keywords:

Proactive language learning
Second language learning
Explicit L2 knowledge
Implicit L2 knowledge

ABSTRACT

Drawing on Proactive Language Learning Theory (Papi & Hiver, 2025), this study is the first to formally operationalize the construct of information-seeking behavior, which is defined as proactive engagement in seeking and using explicit linguistic and metalinguistic information to enhance one's L2 knowledge and skills. Questionnaire scales were developed and validated in two studies: Study 1 administered the scales to 207 Spanish learners in the US, and Study 2 administered them to 294 English learners in Iran. Exploratory factor analysis supported a two-factor solution: Information Gathering, referring to learners' efforts to gather explicit L2 information, and Information Use, representing learners' incorporation of the information in their L2 use. Confirmatory factor analysis, correlational analyses, and reliability analyses confirmed the discriminant and convergent validity of the scales. Hierarchical regression analyses supported the predictive validity of the scales: Information Gathering significantly predicted proficiency when entered before Information Use, while Information Use made a significant unique contribution to proficiency in the final models. Together, these findings support Information-Seeking Behavior as a multidimensional construct and provide a theoretically grounded instrument for investigating learners' proactive engagement with explicit L2 information.

1. Introduction

Intentional and explicit learning remain two of the most important areas of inquiry and controversy in second language acquisition (SLA) research and theory. Explicit learning refers to "a process during which participants acquire conscious (explicit) knowledge; this is generally associated with intentional learning conditions" (Rebuschat, 2015, p. xiii). The main argument in support of this kind of learning is that, given the inherent differences between L1 and adult L2 acquisition, L2 learners tend to naturally use their explicit learning capacities to acquire different aspects of the target language (DeKeyser, 2000). Research on intentional and explicit learning, however, has mainly been conducted from a cognitive-instructional perspective, with researchers examining the effects of explicit or implicit instructional techniques on various learning outcomes. In a seminal meta-analysis of 77 studies published between 1980 and 1998, Norris and Ortega (2000) showed that in both meaning-focused and form-focused instructional contexts, explicit instruction led to larger immediate and delayed effects on both implicit and explicit L2 outcomes than did implicit instruction (see also Ellis, 2002). Although understanding the role of explicit instruction in L2 learning has been valuable, prior research has tended to view explicit information as merely an instructional resource, without considering how learners proactively use it to advance their acquisition

* Corresponding author. School of Teacher Education, 1114 W. Call St, G129 Stone Building, Tallahassee, FL, 32306, United States.

E-mail addresses: mpapi@fsu.edu (M. Papi), yszkw@waseda.jp (Y. Suzuki), jk14v@fsu.edu (C.J. Kim), mkim10@fsu.edu (M. Kim).

process.

The learner plays a central role in the SLA process, particularly in the explicit and intentional dimensions of learning. This is because L2 learners have a strong tendency for intentional learning, which is defined as “a deliberate attempt to commit factual information to memory” (Hulstijn, 2013, p. 2632). That is, they tend to use their analytic skills and capacities to improve their L2 knowledge, typically in an explicit manner. It has been gradually recognized that intentional and explicit learning plays an important role in developing L2 knowledge and skills (DeKeyser & Suzuki, 2025; Hulstijn & Laufer, 2001; Leow, 2015; Pawlak, 2025; Saito & Plonsky, 2019). According to Leow and Zamora (2017), learners tend to use their explicit learning skills regardless of the experimental condition they are assigned. This observation is also supported by recent research on individual learning differences (Papi et al., 2023; Teimouri et al., 2022). Yet there is a paucity of research investigating how learners utilize explicit learning. Therefore, it makes logical sense to explore learners’ patterns of engagement in the intentional learning of explicit L2 information.

To address this gap, Papi and Hiver (2025) recently proposed the Proactive Language Learning Theory (PLLT), offering a learner-centered perspective that focuses on understanding learners’ behavioral patterns in L2 acquisition. From their point of view, L2 learning is considered “a pursuit wherein the learner is the proactive agent who does the learning rather than a passive recipient and processor of linguistic input to whom L2 learning happens” (Papi & Hiver, 2025, p. 298). Input, interaction, explicit (meta)linguistic information, and corrective feedback are thus not only instructional resources but, more importantly, learning resources and opportunities that learners use to enhance their L2 knowledge and skills. Building on decades of research in SLA, PLLT outlines learners’ input-seeking, interaction-seeking, feedback-seeking, and information-seeking behavior (InfoSB) as the major proactive behaviors they engage in to learn a new language. The inclusion of InfoSB as one of the four pillars of PLLT acknowledges the importance of explicit learning, given its prevalence in L2 learning and teaching and the theoretical and empirical support for its contribution to the development of L2 knowledge and skills.

Building on Papi and Hiver’s (2025) PLLT, the present study is an initial attempt to conceptualize and operationalize the notion of InfoSB in the field of SLA. Defined as “learners’ agentic and strategic efforts to seek, engage with, and use explicit L2 information (e.g., lexical, grammatical, phonological, pragmatic, cultural) for the purpose of improving their knowledge of the target language” (Papi & Hiver, 2025, p. 307), the notion of InfoSB provides a theoretical lens for understanding how learners proactively seek out, process, and apply explicit L2 information in pursuit of their learning goals. Such a construct can enhance our understanding of differential learning outcomes across instructional settings, and contribute to SLA theory by providing a conceptual and empirical tool for examining how learners engage in intentional learning of explicit L2 information. It can also help teachers and learners create conditions that support effective intentional learning of such information.

2. Background

2.1. Explicit learning in SLA research

The role of explicit knowledge—defined as conscious awareness of linguistic forms and the ability to verbalize linguistic rules—remains a central locus of inquiry in SLA. While early non-interface positions, most famously Krashen’s (1982) Monitor Model, argued that explicitly learned knowledge could never convert into implicit competence, modern cognitive-SLA perspectives increasingly emphasize the facilitative, and often essential, role of explicit knowledge as a precursor for adult L2 development. From the perspective of Skill Acquisition Theory, for instance, L2 acquisition is conceptualized as a developmental progression from initial declarative understanding to proceduralization, and eventually, to automatization (DeKeyser & Suzuki, 2025).

In this framework, explicit information serves as declarative representations, which are “objects of thought” that learners must consciously hold in working memory (Gade et al., 2014). Through deliberate and repeated practice, these declarative representations are used to create new procedural rules that directly guide performance such as using the L2 for communication. This “representational change” is critical for instructed L2 learners, particularly those with limited exposure to the target language, as explicit knowledge provides the scaffolding needed to engage in communicative tasks before high levels of fluency are reached (Suzuki and DeKeyser, 2026). Without this explicit foundation, adult learners often struggle to move beyond basic communicative strategies to achieve high levels of grammatical accuracy. This cognitive architecture is further detailed in Ullman, 2025 Declarative-Procedural model, which posits that explicit knowledge is housed within the declarative memory system—a system responsible for “knowledge that” (e.g., semantic facts and idiosyncratic linguistic rules). In contrast, the procedural memory system handles “knowledge how,” enabling the fluid, unconscious execution of linguistic skills. While these systems are neuroanatomically distinct, they are often functionally cooperative. Adult learners, who may have diminished implicit learning capacities compared to children, rely on declarative memory to intentionally encode complex linguistic patterns, which are then refined into procedural skills through usage.

From an interactionist perspective, when learners attempt to apply their explicit knowledge in real-time interaction, they engage in hypothesis testing. In doing so, they may receive corrective feedback—whether delivered explicitly through metalinguistic clues or implicitly through conversational recasts (Gass et al., 2025). Such feedback can draw learner’s attention to problematic aspects of their L2 use, triggering the “noticing” process (Schmidt, 1990). This, in turn, pushes the learner toward producing “modified output,” thereby creating opportunities for subsequent L2 development (Sato, 2023). In essence, interaction serves as the behavioral context in which declarative knowledge is tested, challenged, and eventually restructured into more stable procedural forms.

The importance of explicit learning has been empirically documented across various linguistic domains. For instance, in vocabulary acquisition, the role of intentional learning has been particularly emphasized. As demonstrated in a meta-analysis by Webb et al. (2020), intentional vocabulary-learning activities—such as flashcards, word lists, writing, and fill-in-the-blanks—contribute to relatively large learning gains. This underscores the potential for learners to proactively expand their linguistic repertoire through

deliberate, word-focused engagement. However, Webb et al. also observed that gains dropped significantly on delayed tests, suggesting that while information gathering is effective for initial acquisition, long-term retention depends on the learner's subsequent encounters and continued use of those items. Similar theoretical developments have recently gained traction in the domain of L2 pronunciation, where the distinction between explicit and implicit (automatized) knowledge is increasingly recognized as a determinant of L2 development (Pawlak, 2025). As Crowther and Loewen (2025) observe, while Instructed SLA (ISLA) research has historically prioritized grammar and vocabulary, a burgeoning body of work now emphasizes the systematic manipulation of learning conditions to facilitate phonological development. In this context, explicit phonetic instruction—which often involves providing metalinguistic information such as articulatory demonstrations—is essential for facilitating initial “noticing” of features that may be perceptually non-salient to the learner (Crowther & Loewen, 2025). From a Skill Acquisition Theory perspective, explicit pronunciation knowledge serves as a conscious “crutch” that must be proceduralized through practice and subsequently automatized to be functionally available under the time pressures of unplanned conversation (Saito & Plonsky, 2019).

2.2. From explicit instruction to proactivity

The role of explicit learning has traditionally been operationalized within the domain of explicit instruction in ISLA research. A massive body of work has provided robust empirical support for the efficacy of directing learners' attention to language forms through rule explanation or inductive guidance. Meta-analyses spanning several decades have consistently demonstrated that explicit instruction yields larger immediate and delayed effect sizes than implicit instruction across nearly all linguistic domains, including grammar, pragmatics, and pronunciation (Goo et al., 2015; Norris & Ortega, 2000; Spada & Tomita, 2010). Furthermore, domain-specific syntheses in pronunciation (e.g., Saito & Plonsky, 2019) corroborate that explicit instruction does more than facilitate initial noticing; it actively supports the proceduralization of complex features by enabling “monitored production” during communicative activities. This suggests that when learners have access to explicit rules, they can consciously monitor their output to resolve linguistic ambiguity, gradually fine-tuning their performance until the target features become automatized.

However, despite substantial evidence in favor of explicit instruction, it is vital to recognize that the mere acquisition of explicit, declarative knowledge is not a sufficient condition for L2 development. While directing attention to linguistic features—particularly those that are non-salient or differ significantly from the learner's L1 system—clearly benefits acquisition, an overemphasis on abstract rules carries potential drawbacks. It is often the case that some learners become excessively rule-conscious, creating a cognitive bottleneck in which the retrieval of explicit facts interferes with communicative fluency. Therefore, it is critical to move beyond the study of mere “rule-knowing” and examine how learners strategically utilize explicit information for actual L2 usage (Ellis & Wulff, 2025).

A key mechanism in this transition from knowledge to use is the concept of languaging—the process of using language to mediate thinking and problem-solving (Swain, 2006). In this view, producing language (orally or in writing) is not just a result of learning, but a cognitive tool that allows learners to externalize their thoughts, making them available for reflection and revision. Whether through metatalk in small groups or written languaging (Ishikawa & Suzuki, 2023), this process functions as a behavioral bridge where explicit knowledge is transformed into a learning resource.

Despite these insights into the value of languaging and output processing, the current body of research has paid relatively little attention to how learners process and utilize explicit information proactively (Papi & Hiver, 2025). In other words, learners are often treated as passive recipients and processors of information. By viewing explicit information primarily as an instructional resource provided by the teacher through feedback or grammar explanations, the field has inadvertently marginalized the learner's role as the strategic director of their own acquisition process.

Because learners' own agentic capacity to apply metalinguistic information to solve linguistic issues during language use is critical (Toth, 2011), research must examine their strategic tendencies as they unfold across different learning contexts and over time. This requires a new theoretical model that redefines explicit information not merely as an external treatment provided by an instructor, but as a dynamic resource that learners proactively seek out, process, and apply to shape their own language learning. By operationalizing these behaviors through an information-seeking model, we can better understand the agentic processes that allow learners to successfully navigate the path from declarative awareness to functional skills.

2.3. Proactive Language Learning Theory

In response to the predominantly psycholinguistic approach to theorizing and researching SLA, Papi and Hiver (2025) proposed the Proactive Language Learning Theory (PLLT), which shifts attention from mental processes and abstract linguistic representation in the mind to what learners do to learn a new language. The shift from learners' cognition to their behavior places learners themselves, rather than their mental representations, at the center of L2 acquisition. From a PLLT perspective, “it is not mere availability of linguistic data, provision of instruction, opportunities for interaction, or exposure to the context that leads to L2 learning. Rather, just like every other human endeavor, L2 learning is a task that is deliberately and consciously accomplished by a human being who thinks, makes decisions, acts, reflects on experiences, and reaches new insights” (p. 298).

Based on that reasoning, the authors proposed PLLT to represent learners' proactive and strategic involvement in learning an additional language along four main dimensions. These include input-seeking, interaction-seeking, feedback-seeking, and information-seeking behavior, which refer to learners' efforts to seek and use input, interaction, feedback, and information to enhance their L2 knowledge and skills. Of the four dimensions, feedback-seeking behavior has received increasing scholarly attention (e.g., Mahbodi et al., 2025; Papi et al., 2019, 2020, 2024), as it was introduced before the PLLT proposal. More recently, qualitative studies have

explored other proactive learning behaviors. For instance, using semi-structured interviews with 10 international graduate students in the United States, Farhang and Papi (2025) classified interaction-seeking behaviors into seeking, engaging, and using interaction opportunities for L2 learning (see also Qian & Papi, 2026). In another study, Chiocca (2025) found that 10 Chinese learners of French enrolled in a structured e-Tandem program employed all four proactive L2 learning behaviors in their e-Tandem learning experience. In addition, the researcher used an intervention involving conversation, transcription, and guided reflection to enhance students' cognitive engagement in their learning process.

More relevant to the present study, Qian et al. (2026) explored 11 English learners' InfoSB at an English language center at a major university in the United States. Based on several hours of semi-structured interviews on how learners seek and use linguistic and meta-linguistic information to further their English knowledge and skills, the researchers found that the behaviors could be classified into two major themes: *information gathering* and *information use*. Information-gathering refers to learners' purposeful efforts to obtain and study explicit information from various resources, and information use represents learners' efforts to put that knowledge into action during L2 use. Here, the term 'information' refers to explicit knowledge of the lexical, grammatical, phonological, pragmatic, and cultural aspects of the target language. Such information can be accessed through human, digital, or textual resources.

2.4. Distinguishing InfoSB from learning strategies and self-regulated learning

Although InfoSB may appear to overlap with traditional language learning strategies, there are important distinctions between the two in terms of theoretical grounding, scope, and function.

The first distinction concerns theoretical grounding. Traditional strategy frameworks (e.g., Cohen, 1998; O'Malley & Chamot, 1990; Oxford, 1990) were developed to classify and describe the techniques learners employ to facilitate language learning. By contrast, InfoSB is embedded within PLLT (Papi & Hiver, 2025), which conceptualizes language learning as emerging from learners' proactive engagement in input-, interaction-, feedback-, and information-seeking behaviors. Thus, whereas strategy research is primarily concerned with identifying and categorizing techniques that help enhance learning specific linguistic forms or items, PLLT seeks to explain how learners proactively seek, create, and exploit learning opportunities. InfoSB represents one manifestation of this broader theoretical framework, focusing specifically on learners' efforts to obtain and use explicit linguistic or metalinguistic information for language learning.

In terms of scope, while strategy taxonomies are defined by specific L2 learning techniques, InfoSB represents a higher-order pattern of proactive engagement with explicit information. In this sense, it reflects the extent to which a learner seeks and uses explicit information for language learning. Strategies are thus specific observable behaviors through which this behavioral tendency is enacted; they are situated beneath InfoSB in the conceptual hierarchy rather than alongside it. For instance, learners may enact InfoSB by consulting dictionaries, grammar references, teachers, peers, or AI tools. These activities do not constitute InfoSB itself but represent possible means through which the broader behavioral tendency can be expressed.

The final distinction concerns function. Strategy taxonomies describe techniques that facilitate language learning in general but remain largely agnostic about the cognitive mechanism by which knowledge becomes skill. InfoSB instead draws on SLA theories such as Skill Acquisition Theory to elucidate an acquisitional pathway: how learners accumulate declarative L2 knowledge by gathering explicit information, and how they proceduralize that knowledge by intentionally deploying it in communication. Crucially, this logic can be reflected in the structure of the construct itself. Info-gathering indexes the proactive accumulation of declarative L2 knowledge, whereas info-use indexes its deliberate proceduralization through L2 use (Qian et al., 2026). InfoSB thus provides a theoretically grounded framework for understanding how learners proactively gather and use explicit information in ways that may contribute to the development of linguistic skills.

A similar distinction can be made between InfoSB and self-regulated learning (Dörnyei, 2005; Oxford, 2017). Self-regulated learning refers broadly to learners' active management of their learning through processes such as goal setting, planning, monitoring, strategy selection, self-evaluation, and reflection (Zimmerman, 2002). InfoSB, by contrast, is narrower and more domain-specific: it concerns learners' proactive engagement with one particular type of SLA resource, namely explicit linguistic and metalinguistic information. Thus, InfoSB may involve self-regulatory processes, but it does not attempt to measure learners' broader capacity for self-regulation. Rather, it captures a specific proactive behavior through which self-regulation may be enacted in L2 learning. For example, when learners search for or use information about grammar, vocabulary, pronunciation, or cultural appropriateness, they may be acting in a self-regulated manner; however, the focus of info-gathering or info-use is not general planning or monitoring, but the proactive acquisition and use of explicit L2 information. In this way, InfoSB can be understood as a theoretically specified pattern of proactive, self-regulated learning through which learners seek out and put to use explicit information about L2 vocabulary, grammar, pronunciation, pragmatics, and cultural conventions.

In sum, InfoSB differs from both L2 learning strategies and self-regulated learning in scope and function. It is broader than isolated strategies because it captures recurring proactive behavioral tendencies rather than individual techniques, but narrower than self-regulated learning because it focuses specifically on seeking and using explicit L2 information. This positioning allows InfoSB to complement, rather than duplicate, existing strategy-based and self-regulation frameworks.

2.5. This study

Building on the theoretical framework of Proactive Language Learning Theory (Papi & Hiver, 2025) and informed by preliminary qualitative evidence on learners' InfoSB (Qian et al., 2026), the present study aims to formally conceptualize and operationalize the construct of InfoSB in SLA. Specifically, the goal is to develop a questionnaire measuring learners' proactive engagement in seeking

and using explicit information, to validate its factor structure using exploratory and confirmatory factor-analytic techniques, and to examine its predictive power in relation to L2 proficiency. To examine the robustness and potential applicability of the proposed model, two independent studies were conducted in distinct linguistic and sociocultural contexts.

In Study 1, the InfoSB scales were administered to 207 learners of Spanish as a foreign language at a university in the United States. Exploratory factor analysis was used to identify the underlying factor structure, followed by confirmatory factor analysis to test its fit and discriminant validity. Hierarchical multiple regression was then conducted to examine the predictive validity of the identified factors for self-assessed L2 proficiency.

In Study 2, parallel versions of the scales were administered to 294 learners of English as a foreign language at a private language institute in Iran. Parallel analyses were conducted to determine whether the factor structure identified in Study 1 would replicate in a different L2 and cultural context. This parallel two-sample design provides converging evidence for the construct validity of the InfoSB model and supports its potential applicability across distinct L2 learning contexts.

Across both studies, the following research questions guided the investigation:

- 1) What is the factor structure of information-seeking behavior in L2 acquisition?
- 2) What is the relationship between L2 information-seeking behavior and L2 proficiency?

3. Study 1

3.1. Participants

The participants consisted of 207 university students (77 males, 126 females, 3 others, 1 missing) in the United States who were learning Spanish as a foreign language in the state of Florida. Participants ranged in age from 17 to 58 years ($M = 21.62$, $SD = 3.31$). Almost all participants were native English speakers ($n = 206$), except for one individual who reported Bengali as their first language. Participants reported majors across various disciplines, including Physiology and Neuroscience (31.58%), Biology and life sciences (28.57%), Communication, Humanities, and Arts (21.80%), Computer Science (8.27%), Environmental Sciences (6.77%), and Physical Sciences (3.76%). Among the 166 participants who reported their current university language-course level, 87.3% of respondents were taking 1000- or 2000-level courses, indicating that the sample was predominantly composed of beginning- to intermediate-level language learners. Smaller proportions ($N = 21$) were enrolled at the 3000 level (7.8%) and 4000 level (4.8%), approximately corresponding to upper-intermediate and advanced proficiency, respectively.

3.2. Procedures

Following institutional review board approval, the research instruments were pilot-tested before the main study. An item pool was compiled from interviews with 14 L2 learners in the United States. Based on the item pool, scales were developed to represent learners' efforts to seek and use information for L2 learning. The developed questionnaire underwent further modification based on cognitive interviews with 10 participants. During the cognitive interviews, the students completed the questionnaire while thinking aloud and provided feedback on their comprehension of the items. After revisions were made, a trained research assistant visited the participating classes, invited students to participate voluntarily, informed them of their rights and research procedures, and administered the questionnaire, which took approximately 15 minutes to complete.

3.3. Instruments

A questionnaire was used to collect data. The first part of the questionnaire collected demographic information, including participants' gender, age, first language(s), university status (undergraduate or graduate), year of study, academic major, years of L2 study, and duration of residence in Spanish-speaking countries. The second part of the questionnaire included items measuring InfoSB through two subscales (see [Appendix](#)). The Info-Gathering subscale included six items that assessed the extent to which learners actively sought and obtained language-related information—including grammar, vocabulary, and pronunciation—by consulting various sources such as teachers, peers, native speakers, online tools, or AI-based applications. The Info-Use subscale also included six items measuring learners' tendencies to apply the information in L2 use, including the extent to which they incorporated newly learned forms or structures into their spoken or written production. Together, these subscales captured the behaviors learners engaged in to obtain and utilize (meta)linguistic information during language learning. All items used a six-point Likert scale, with values 1 through 6 corresponding to *Never/Rarely*, *Occasionally*, *Sometimes*, *Often*, *Frequently*, and *Constantly*, respectively.

The third part of the questionnaire assessed self-reported Spanish proficiency using a self-assessment scale ([Tigchelaar et al., 2017](#)) developed based on the American Council on the Teaching of Foreign Languages (ACTFL). The scale included five general descriptors, each representing Novice, Intermediate, Advanced, Superior, or Distinguished levels (see Supplementary File).

3.4. Data analysis

SPSS version 31 (IBM) was used to run exploratory factor analyses (EFA) with principal axis factoring method of extraction and direct Oblimin rotation to investigate the latent structure underlying InfoSB. Decisions regarding the number of factors to retain were informed by eigenvalues greater than 1.00, visual inspection of scree plots, and the theoretical foundations of the measurement scales.

To assess the suitability of the data for factor analysis, the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity were examined (Shrestha, 2021).

Confirmatory factor analyses (CFAs) were subsequently conducted in Mplus 8 to compare the one-factor and two-factor models. These analyses tested the fit of the hypothesized measurement structure and provided evidence for the distinctiveness and coherence of the two constructs. Because the datasets were non-normally distributed and contained some missing responses, CFAs employed the maximum likelihood robust (MLR) estimator, which yields more stable and accurate parameter estimates under violations of normality. Finally, correlational analyses and hierarchical multiple regression models were conducted to determine the extent to which Info-Gathering and Info-Use predicted self-assessed L2 proficiency. These analyses also evaluated the explanatory power of the revised conceptual model to determine which factor structure provided the best fit for the data.

3.5. Factor analytic results

The EFA analysis of L2-Spanish data showed that Bartlett's test of sphericity, $\chi^2(66) = 1557.30$, $p < .001$, and the Kaiser-Meyer-Olkin measure of sampling adequacy, $KMO = .892$, confirmed the suitability of the data for factor analysis ($N = 207$). The analysis proceeded iteratively to identify the most parsimonious and theoretically coherent factor structure. Initial analysis of the 12-item scale yielded a three-factor solution with eigenvalues exceeding 1.0. However, examination of the pattern matrix revealed substantial cross-loadings and theoretical ambiguities for four items, which were removed from the model. Following these deletions, the final 8-item solution yielded two factors with eigenvalues exceeding 1.0, accounting for 63.8% of the total variance. Factor 1 (Info-Use; 3 items) explained 54.5% of the variance, capturing behaviors related to applying learned linguistic information. Factor 2 (Info-Gathering; 5 items) explained an additional 9.3% of variance, representing behaviors related to actively seeking linguistic information from individuals or resources. The scree plot supported this two-factor solution, showing a clear elbow after the second factor. The two factors were moderately correlated ($r = .61$), indicating related but distinct dimensions. Examination of the pattern matrix (see Table 1) revealed strong factor loadings ranging from .78 to .86 for Info-Use items and from .57 to .84 for Info-Gathering items. All items demonstrated adequate communalities (range: .53 to .72), confirming that the extracted factors accounted for a substantial proportion of the variance in each item.

We then tested the two-factor structure identified in the EFA using a series of CFA models and compared it with an alternative single-factor model to assess the factors' discriminant validity. We selected MLR estimation because it provides accurate parameter estimates when data violate normality assumptions. This was appropriate for our data, given the Likert-type ordinal nature of responses and evidence of platykurtic distributions across all items (kurtosis values ranged from $-.71$ to -1.11). The initial two-factor model demonstrated marginal fit: $\chi^2(19) = 74.71$, $p < .001$; RMSEA = .115 [.088, .144]; CFI = .923; TLI = .893; SRMR = .060. Based on modification indices, three theoretically justified error covariances were added to the model.

The final modified model demonstrated an excellent fit across all indices: $\chi^2(16) = 20.93$, $p = .181$; RMSEA = .039 [.000, .080]; CFI = .993; TLI = .988; SRMR = .023. The nonsignificant chi-square indicated that the model adequately reproduced the observed covariance matrix. All fit indices exceeded the recommended thresholds (Hu & Bentler, 1999), with the RMSEA indicating a close fit, the CFI and TLI suggesting excellent incremental fit, and the SRMR demonstrating good absolute fit (Table 2). In addition, all standardized factor loadings were statistically significant ($p < .001$) and substantial in magnitude (see Table 3), ranging from .80 to .88 for Use ($M = .84$) and from .59 to .84 for Info-Gathering ($M = .74$). The latent correlation between Info-Use and Info-Gathering was significant ($r = .75$, $p < .001$), indicating the two dimensions share approximately 56% of their variance while remaining below the .85 correlation that is used as the conventional cutoff for suspecting discriminant validity (Brown, 2015). In addition, further evidence for the discriminant validity of the two scales was evidenced by the model-comparison CFA results, which showed that the two-factor model was far superior to an alternative one-factor model (see Table 2). The two-factor final model showed substantially better approximation ($\Delta RMSEA = -.135$, an improvement of 77%), incremental fit ($\Delta CFI = +.168$, $\Delta TLI = +.233$), and absolute fit ($\Delta SRMR = -.050$, a 68% improvement). Additionally, the Akaike Information Criterion strongly favored the two-factor final model ($AIC = 5036.34$) over both the two-factor baseline model ($AIC = 5095.92$) and the one-factor model ($AIC = 5193.98$), with lower AIC values indicating better model parsimony.

As Table 4 shows below, both factors demonstrated excellent construct reliability and validity. Composite reliability values (Info-

Table 1
Exploratory factor analysis: Pattern matrix loadings for Spanish InfoSB scale.

Item	L2-Spanish Results		
	Info-Use	Info-Gathering	h^2
Q1	.857	—	.718
Q2	.784	—	.659
Q3	.829	—	.699
Q10	—	.844	.643
Q11	—	.782	.527
Q5	—	.584	.616
Q6	—	.569	.549
Q7	.353	.567	.691

Note. Extraction: Principal axis factoring. Rotation: Direct oblimin. Cumulative variance = 63.8%. h^2 = communality. Loadings $< .30$ suppressed.

Table 2

Confirmatory factor analysis: Model fit comparison for Spanish InfoSB scale.

Model	χ^2 (df)	<i>p</i>	CFI	TLI	RMSEA [90% CI]	SRMR
Two-Factor Final	20.93 (16)	.181	.993	.988	.04 [.00, .08]	.02
One-Factor	144.56 (19)	< .001	.825	.755	.17 [.15, .20]	.07

Note. MLR = maximum likelihood estimation with robust standard errors.

Table 3

Standardized factor loadings for the final measurement model in the Spanish InfoSB scale.

Factor	Item	λ	SE	R^2
Info-Use	Q2	.80	.04	.63
	Q1	.88	.03	.77
	Q3	.83	.04	.69
Info-Gathering	Q6	.77	.04	.59
	Q5	.83	.04	.69
	Q7	.84	.04	.71
	Q10	.69	.05	.47
	Q11	.59	.05	.34

Note. λ = standardized factor loading; R^2 = proportion of variance explained. Factor correlation: $r = .75$, $p < .001$. All factor loadings are significant at $p < .001$.

Table 4

Descriptive statistics and reliability indices for final scales in the Spanish InfoSB scale.

Factor	<i>k</i>	<i>M</i>	<i>SD</i>	α	ω	CR	AVE
Info-Gathering	5	3.33	1.24	.87	.87	.86	.56
Info-Use	3	3.66	1.26	.88	.88	.87	.70

Note. *k* = number of items; CR = composite reliability; AVE = average variance extracted. All reliability indices exceed recommended thresholds (α , ω , CR > .70; AVE > .50).

Use: CR = .87; Info-Gathering: CR = .86) substantially exceeded the recommended threshold of .70. Average variance extracted values (Use: AVE = .70; Info-Gathering: AVE = .56) surpassed the recommended minimum of .50, supporting convergent validity. These psychometric properties indicate that both factors are well-defined and internally consistent.

3.6. Predictive validity

To examine whether InfoSB predicts L2 proficiency beyond the effects of learning duration, we conducted hierarchical multiple regression. Table 5 presents descriptive statistics and bivariate correlations for the L2-Spanish samples.

Descriptive statistics and correlations were calculated using complete cases, whereas missing values were replaced with variable means prior to the hierarchical regression analyses, resulting in a regression sample of 207 participants. In the hierarchical regression analysis for Spanish learners ($N = 207$), Length of Study was entered at Step 1, followed by Info-Gathering at Step 2 and Info-Use at Step 3 (Table 6). Self-reported Spanish proficiency (1 = Novice to 5 = Distinguished) served as the dependent variable.

In Step 1, Length of Study significantly predicted Spanish proficiency, $F(1, 205) = 47.50$, $p < .001$, accounting for 19% of variance ($R^2 = .19$, $\beta = .43$, $p < .001$). This substantial effect indicates that learners with longer Spanish study reported higher proficiency levels. Next, in Step 2, adding Info-Gathering significantly improved the model, $\Delta R^2 = .053$, $\Delta F(1, 204) = 14.31$, $p < .001$. The model now explained 24% of variance, $F(2, 204) = 32.45$, $p < .001$. In this model, both Length of Study ($\beta = .44$, $p < .001$) and Info-Gathering ($\beta = .23$, $p < .001$) emerged as significant predictors, suggesting that actively seeking linguistic information contributed to proficiency beyond the effect of learning duration. Adding Info-Use to the final step also significantly improved prediction, $\Delta R^2 = .014$, $\Delta F(1, 203) = 3.92$, $p < .05$. The final model explained approximately 26% of the variance, $F(3, 203) = 23.25$, $p < .001$. Length of Study remained significant ($\beta = .43$, $p < .001$) and Info-Use was also significant ($\beta = .16$, $p < .05$) whereas Info-Gathering was no longer significant ($\beta = .13$, $p = .12$), suggesting that its unique predictive contribution was reduced once the overlapping variance with Info-Use was taken into account. Multicollinearity diagnostics showed acceptable levels: tolerance values ranged from .56 to 1.00 and VIF values ranged from 1.01 to 1.78, suggesting no problematic multicollinearity. A Durbin-Watson statistic of 1.93 indicated no substantial autocorrelation.

Table 5

Descriptive Statistics and Correlations for the L2 Spanish Sample (N = 186).

Variable	<i>M</i>	<i>SD</i>	1	2	3	4
1. Proficiency	2.16	.79	—			
2. Length of Study	3.64	3.08	.46***	—		
3. Info-Gathering	3.33	1.24	.24***	-.02	—	
4. Info-Use	3.66	1.26	.29***	.07	.66***	—

Table 6

Hierarchical regression analysis predicting Spanish L2 proficiency (N = 207).

Model	Predictor	β	R^2	ΔF	<i>p</i>	Tolerance	VIF
1	Length of Study	.44***	.19	47.50***	< .001	1.00	1.00
2	Length of Study	.44***	.24	14.31***	< .001	1.00	1.00
	Info-Gathering	.23***			< .001	1.00	1.00
3	Length of Study	.43***	.26	3.92*	< .001	.99	1.01
	Info-Gathering	.13			.12	.57	1.77
	Info-Use	.16*			<.05	.56	1.78

Note. Dependent variable: Spanish proficiency (1 = Novice to 5 = Distinguished).

* $p < .05$, *** $p < .001$.

4. Study 2

4.1. Participants

The second study included 294 Iranian learners of English as a foreign language enrolled at different branches of a private language institute in Iran. This sample included 82 males and 212 females, aged 13 to 43 years ($M = 17.94$, $SD = 5.06$). Their length of English study ranged from 1 to 16 years ($M = 6.85$, $SD = 2.47$). Among the 289 participants who reported their first language, the majority were native Farsi speakers (98.27%, $n = 284$), with a small number reporting Azeri (1.04%, $n = 3$) or Kurdish (.69%, $n = 2$). Most participants reported Mathematics and Physics (31.58%) and Experimental Sciences (28.51%) as their academic majors. Others identified a variety of majors, including Arts and Humanities, Accounting and Finance, and Biology. Based on the institute's placement procedures, about 54% of participants were enrolled in high intermediate classes, and 46% were enrolled in advanced classes.

4.2. Instruments & procedures

An item pool was developed from interviews with 10 L2 learners in Iran, which closely aligned with the scales used in Study 1. The items were fine-tuned based on cognitive interviews with several participants. English proficiency was assessed using a 50-item ACTFL Self-Assessment (Tigchelaar et al., 2017). The instrument comprised five sets of 10 Can-Do statements spanning proficiency bands from Novice Low through Superior. Each statement was rated on a four-point scale: 1 ("Not Yet"), 2 ("With Much Help"), 3 ("With Little Help"), and 4 ("Yes, I Can Do This Well"). Responses were summed across the 50 items to produce a continuous self-assessed proficiency score, with a possible range of 50–200; higher scores indicated greater perceived L2 proficiency. In the present sample, scores ranged from 108 to 200. The questionnaire was administered online via Google Forms.

Table 7

Exploratory factor analysis: Pattern matrix loadings for Iranian InfoSB scale.

Item	Info-Gathering	Info-Use	h^2
Q8	.866	—	.625
Q6	.764	—	.534
Q9	.663	—	.588
Q5	.648	—	.642
Q7	.615	—	.579
Q1	—	.969	.825
Q3	—	.810	.696
Q4	—	.745	.620
Q2	—	.634	.671

Note. Extraction: Principal axis factoring. Rotation: Direct oblimin. Cumulative variance = 64.2%. h^2 = communality. Loadings < .10 suppressed.

4.3. Factor analytic results

The same procedures were followed to analyze the data. As with the first dataset, Bartlett's test of sphericity, $\chi^2(66) = 2709.67$, $p < .001$, and the Kaiser-Meyer-Olkin measure of sampling adequacy, $KMO = .918$, confirmed the data were suitable for factor analysis. Using principal axis factoring with direct oblimin rotation, the initial analysis of the 12-item scale yielded two factors with eigenvalues exceeding 1.0, accounting for 64.2% of the total variance. Factor 1 explained 57.6% of the variance, and Factor 2 explained 6.7% of the variance. The scree plot supported this two-factor solution. However, examination of the pattern matrix revealed substantial cross-loadings and theoretical ambiguities for three items, which were subsequently removed. Following these deletions, the final 9-item solution resulted in two factors with eigenvalues exceeding 1.0, accounting for 64.2% of the total variance, which was supported by the scree plot as well. Factor 1 (Info-Gathering; 5 items) explained 57.9% of the variance, and Factor 2 (Info-Use; 4 items) explained an additional 6.3% of variance (Table 7). All items showed adequate communalities (range: .53 to .83) and factor loadings ranged from .62 to .97. The two factors were strongly correlated yet sufficiently distinct ($r = .75$).

To test whether the two-factor structure identified in this independent sample best represents the data relative to a single-factor solution, a CFA using MLR estimation, appropriate for non-normal data, was conducted. The initial two-factor model demonstrated marginal fit: $\chi^2(26) = 122.57$, $p < .001$; RMSEA = .11 [.093, .133]; CFI = .896; TLI = .856; SRMR = .05. Two theoretically justified error covariances were added to the model based on modification indices. The final model exhibited excellent fit: $\chi^2(24) = 47.23$, $p = .003$; RMSEA = .06 [.033, .081]; CFI = .975; TLI = .962; SRMR = .03 (Table 8).

As shown in Table 9, all standardized factor loadings ranged from .68 to .85 and were statistically significant ($p < .001$), indicating strong indicator-construct relationships. Both factors demonstrated excellent construct reliability and validity (Table 10). Composite reliability values (Info-Gathering: CR = .87; Info-Use: CR = .84) exceeded recommended standards, and average variance extracted values (Info-Gathering: AVE = .58; Info-Use: AVE = .60) supported strong convergent validity. Inspection of normalized residuals revealed that all values fell within the acceptable range of ± 2.0 , indicating a good local fit (Brown, 2015).

The latent correlation between factors ($r = .86$, $p < .001$) was stronger than in the Spanish sample, indicating the two dimensions share approximately 74% of variance while remaining empirically distinct. To further explore the discriminant validity of the scales, as with the L2-Spanish sample, we tested whether a simpler one-factor model could adequately represent the data. The one-factor model (all nine items loading on a single information-seeking factor) demonstrated poor fit: $\chi^2(27) = 162.62$, $p < .001$; RMSEA = .131 [.112, .150]; CFI = .854; TLI = .805; SRMR = .062. All indices fell below acceptable standards. Comparison of fit indices strongly favored the two-factor solution, consistent with the results from the Spanish sample. The two-factor model showed substantially better approximation ($\Delta RMSEA = -.074$, an improvement of 56%), incremental fit ($\Delta CFI = +.121$, $\Delta TLI = +.157$), and absolute fit ($\Delta SRMR = -.028$). The AIC also favored the two-factor final model (AIC = 7042.25) over the one-factor model (AIC = 7253.74). These results provide converging evidence across two independent samples that Info-Use and Info-Gathering represent distinct, theoretically meaningful dimensions of InfoSB.

4.4. Predictive validity

Hierarchical multiple regression analyses were conducted to examine whether InfoSB components predict L2 proficiency beyond the effects of L2 study duration. As shown in Table 11, Info-Use and Info-Gathering demonstrated a strong intercorrelation ($r = .73$, $p < .001$), indicating substantial shared variance (53.3%).

In the hierarchical regression for Iranian learners ($N = 291$), the same predictor entry sequence as in Study 1 was used (See Table 12), with ACTFL proficiency scores serving as the dependent variable. In Step 1, Length of Study was a statistically significant but weak predictor, $F(1, 289) = 5.61$, $p = .019$, accounting for only 1.9% of variance ($R^2 = .019$, $\beta = .138$, $p = .019$). This minimal effect contrasts sharply with the L2-Spanish sample, suggesting that learning duration played a significantly smaller role in explaining proficiency differences among Iranian English learners. Next, in Step 2, adding Info-Gathering dramatically improved the model, $\Delta R^2 = .130$, $\Delta F(1, 288) = 43.94$, $p < .001$. The model now explained 14.9% of variance, $F(2, 288) = 25.20$, $p < .001$. This represented a substantial increase of 13% in explained variance. In this model, both Length of Study ($\beta = .114$, $p = .038$) and Info-Gathering ($\beta = .361$, $p < .001$) were significant predictors, with Info-Gathering showing a much larger effect. Lastly, adding Info-Use to the final model further improved the prediction significantly, $\Delta R^2 = .078$, $\Delta F(1, 287) = 28.84$, $p < .001$. The final model explained 22.7% of variance, $F(3, 287) = 28.03$, $p < .001$. Notably, in this final model, Info-Use emerged as the sole significant predictor ($\beta = .410$, $p < .001$), while Length of Study ($\beta = .085$, $p = .106$) and Info-Gathering ($\beta = .054$, $p = .400$) were no longer significant.

This pattern reveals a different story from the L2-Spanish sample. For Iranian English learners, InfoSB variables were the most dominant predictor of proficiency. The progression across steps suggests that, while both Info-Gathering and Info-Use contributed substantially to proficiency, Info-Use was the stronger predictor. The standardized coefficient for Info-Use ($\beta = .41$) in the final model

Table 8

Confirmatory factor analysis: Model fit comparison for Iranian InfoSB scale.

Model	χ^2 (df)	p	CFI	TLI	RMSEA [90% CI]	SRMR
Two-Factor Final	47.23 (24)	.003	.975	.962	.06 [.033, .081]	.03
One-Factor	162.62 (27)	< .001	.854	.805	.13 [.112, .150]	.06

Note. The Two-Factor Final Model includes two error covariances: Q6 with Q8 (grammar domain method effect) and Q1 with Q2 (resource use content overlap).

Table 9

Standardized factor loadings for the final measurement model in the Iranian InfoSB scale.

Factor	Item	λ	SE	R^2
Info-Gathering	Q6	.68	.05	.46
	Q8	.73	.04	.53
	Q5	.82	.03	.67
	Q9	.79	.03	.63
	Q7	.77	.03	.60
Info-Use	Q2	.69	.04	.48
	Q1	.70	.04	.49
	Q3	.85	.03	.72
	Q4	.84	.02	.71

Note. λ = standardized factor loading; R^2 = proportion of variance explained. All items measured on a 6-point Likert scale (1 = Never to 6 = Always). Factor correlation: $r = .86$, $p < .001$. All factor loadings are significant at $p < .001$.

Table 10

Descriptive statistics and reliability indices for final scales in Iranian InfoSB scale.

Factor	k	M	SD	α	ω	CR	AVE
Info-Gathering	5	4.72	.86	.88	.88	.87	.58
Info-Use	4	4.63	1.08	.90	.90	.84	.60

Note. k = number of items; CR = composite reliability; AVE = average variance extracted. All reliability indices exceed recommended thresholds (α , ω , CR > .70; AVE > .50).

Table 11Intercorrelations for the L2 English Sample ($N = 291$).

Variable	M	SD	1	2	3	4
1. Proficiency	179.63	21.06	—			
2. Length of Study	6.85	2.48	.14*	—		
3. Info-Gathering	4.72	.86	.37***	.07	—	
4. Info-Use	4.63	1.08	.47***	.12*	.73***	—

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 12

Hierarchical regression analysis predicting Iranian L2 proficiency.

Step	Predictor	B	$SE\ B$	β	R^2	ΔF	p	Tolerance	VIF
1	Length of study	1.174	.495	.138*	.019	5.61*	<.05	1.00	1.00
2	Length of study	.968	.454	.114*	.149	43.94***	<.05	.996	1.00
	Info-Gathering	9.398	1.198	.361***			<.001	.996	1.00
3	Length of study	.721	.445	.085	.227	28.84***	.106	.986	1.04
	Info-Gathering	1.408	1.469	.054			.400	.470	2.13
	Info-Use	7.873	1.485	.410***			<.001	.466	2.15

Note. Dependent variable: ACTFL-based proficiency self-assessment scores (range: 108–200).

* $p < .05$. *** $p < .001$.

was particularly substantial, indicating that learners who applied linguistic information more frequently demonstrated substantially higher proficiency.

5. General discussion

The present study aimed to conceptualize and operationalize InfoSB as a key dimension of proactive L2 learning (Papi & Hiver, 2025). While the role of explicit knowledge has historically been a point of inquiry in SLA, modern cognitive perspectives increasingly highlight its facilitative role as a precursor for adult L2 development. Despite the depth of research on explicit instruction, the field has lacked a theoretically grounded instrument to assess how learners habitually and proactively engage with explicit information as a learning resource.

Across two distinct L2 environments—Spanish learners in the US and English learners in Iran—our results consistently supported a two-factor structure. This model distinguishes between Info-Gathering, which represents proactive efforts to gather explicit L2 information, and Info-Use, which represents the incorporation of that information into communicative L2 production (see Fig. 1).

INFORMATION-SEEKING BEHAVIOR

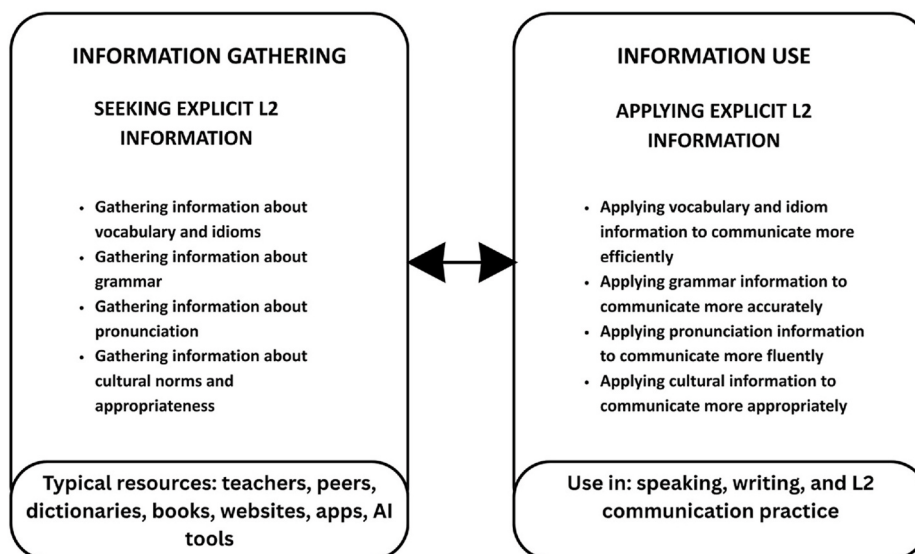


Fig. 1. Visual representation of the two dimensions of the information-seeking behavior (InfoSB) model identified in this study.

Overall, the Info-Gathering and Info-Use scales were largely comparable across the two contexts in their core coverage of vocabulary, grammar, and pronunciation, pragmatic and sociocultural aspects of L2 use, although the specific items retained within each factor varied somewhat across contexts. The hierarchical regression results in Study 2 provide compelling evidence for this distinction. While Info-Gathering was a significant predictor of L2 proficiency when entered alongside learning duration, it lost its statistical significance once Info-Use was added to the model. This suggests that the mere accumulation of explicit facts may not independently enhance proficiency; rather, it is the deployment of that knowledge in L2 communication that appears to explain the variance in proficiency outcomes (DeKeyser, 2007; Sato & McDonough, 2019; Suzuki, 2023).

The distinction between Info-Gathering and Info-Use finds robust theoretical grounding in Skill Acquisition Theory and the Declarative-Procedural model. Within these frameworks, Info-Gathering can be conceptualized as the proactive pursuit of declarative representations that serve as a critical foundation for adult L2 development. However, Skill Acquisition Theory posits that declarative knowledge remains functionally limited unless it undergoes a representational change into procedural knowledge (DeKeyser & Suzuki, 2025). This transition, known as proceduralization, occurs when learners repeatedly engage in tasks that require them to draw upon their declarative knowledge. Thus, Info-Use represents the behavioral manifestation of this cognitive shift. By distinguishing between these two behaviors, we can better account for the ‘knowledge-skill gap’ common in L2 classrooms: learners may gather vast amounts of declarative information (Info-Gathering) but fail to achieve proficiency because they lack the agentic engagement required to proceduralize that knowledge through active application (Info-Use).

Evidence for a similar distinction has also been offered by research on individual differences in how learners pursue goals. For instance, regulatory mode theory (Higgins et al., 2003) distinguishes between an assessment orientation—focused on cognitive analysis and accuracy—and a locomotion orientation, characterized by a focus on movement and taking action to reach a goal. Teimouri et al. (2022) found that when involved in L2 learning, assessors are concerned with cognitive analysis and the accuracy of their L2 use whereas locomotors use the L2 without much concern for the accuracy of their statements (see also Kadivar & Papi, 2026; Qian & Papi, 2026). A linguistic analysis of learners with different regulatory modes found that assessors showed higher syntactic and lexical complexity in their oral L2 production, whereas locomotors exhibited higher speed and more errors Papi et al. (2023). A similar distinction was made by Skehan and Foster (2013), who classified their research participants into “complexity-oriented learners, accuracy-oriented learners, and fluency-oriented learners, whose performance showed some stability whatever the task being done” (p. 252; see also Tavakoli & Skehan, 2005). Such individual differences in learners’ behavioral tendencies highlight the importance of understanding both the gathering and use dimensions of L2 learners’ InfoSB.

In sum, our findings validate a model of InfoSB, distinguishing between the proactive acquisition (Info-Gathering) and application (Info-Use) of explicit (meta)linguistic information. While Info-Gathering involves seeking formal, pragmatic, or cultural data through human or digital resources, Info-Use represents the active deployment of that knowledge to enhance L2 accuracy and fluency. Although CFA confirmed that these dimensions are psychometrically distinct and superior to a unitary model, the high inter-construct correlation suggests that, for many learners, gathering and using explicit information are phenomenologically intertwined. This implies that in foreign language learning, the act of seeking information may often be followed by an application in L2 use, making the

two behaviors practically inseparable in the learner's experience.

6. Practical implications

Although the present findings do not directly test instructional interventions, they suggest that information-seeking could be explicitly taught in the classroom and integrated into language classes as a teachable skill rather than treated as an informal and incidental activity. Teachers might design individual or collaborative tasks that require students to engage in information gathering and use. For example, students could be asked to identify a linguistic problem in their own L2 use, search for relevant information, and then revise or reproduce the language in a new communicative task (see [Chiocca, 2025](#)). Such activities may help students move beyond simply collecting explanations or rules and toward applying them in meaningful L2 production.

Teachers could also help their students recognize that learning explicit information is not limited to the instructional settings and can be accessed and used in a variety of digital and non-digital resources. They might design assignments that encourage learners to move beyond the classroom, for example, by consulting online forums, corpora, dictionaries, textbooks, or AI tools to gather and use relevant explicit information. Such activities may support learners' proactivity and autonomy by helping them view explicit information as a learning resource available across contexts. Students could keep a notebook in which they record the linguistic problems they encounter, describe how they search for lexical or grammatical information to address those problems, reflect on how they apply the new information in L2 communication, and assess how such use enhances their productive or receptive L2 skills.

Teachers may also help students develop information literacy for language learning. They can help students identify reliable sources for accessing explicit information. Training students to leverage digital or non-digital resources to access and use explicit information should be prioritized in today's world of global communication. Given the growing range of resources available to language learners, including dictionaries, corpora, websites, textbooks, online forums, and AI tools, students may benefit from guidance on evaluating the accuracy, appropriateness, and usefulness of the information they gather. For instance, teachers could train students to use AI tools such as ChatGPT responsibly and productively by formulating clear prompts, asking follow-up questions, requesting clarifications and examples, comparing responses with course materials or other reliable sources, and verifying the information before incorporating it into their L2 use.

7. Limitations & future directions

While this study provided preliminary evidence for the two-factor structure of InfoSB and its predictive power for self-assessed L2 proficiency across two distinct sociocultural contexts, several limitations should be considered in interpreting the results. First, the data collected in this study were cross-sectional, meaning they were collected at a single point in time. The relationship between InfoSB and proficiency could be dynamic and reciprocal, in that InfoSB can enhance L2 proficiency, which, in turn, increases learners' motivation and skills for further InfoSB. Therefore, causal inferences cannot be established until longitudinal or experimental studies are conducted to examine such relationships. Self-assessed proficiency in the first study was assessed using a single-item global rating with five different levels, which lacks the nuance and precision of a standardized test. In the second study, a comprehensive 50-item assessment scale was used, which can be considered a more detailed measure of self-perceived proficiency. However, self-assessments can deviate from the objective reality of students' proficiency. Future studies can employ standardized tests to assess L2 proficiency more objectively. The study was conducted among college students learning L2-Spanish in the United States and learners of English as a foreign language in Iran. Although the sociocultural and linguistic disparities between the two contexts lend support to the validity of the InfoSB model, the results cannot be easily generalized to other contexts, such as informal L2 learning settings, study abroad contexts, heritage language learners, or learners of less commonly taught languages. Further qualitative and quantitative studies across diverse settings can shed light on how InfoSB can take different shapes depending on the study context.

The current study identified Info-Gathering and Info-Use as two components of the InfoSB model. Future research can explore the demographic, personality, motivational, affective, and cognitive factors that shape these behaviors ([Li et al., 2022](#); [Papi & Hiver, 2025](#)). In addition, researchers can explore how contextual and relational factors at macro, meso, and micro contexts shape learners' InfoSB. Given the strong correlation between the two InfoSB dimensions, employing techniques such as interviews and think-alouds can help identify how learners distinguish between the two types of InfoSB. Future studies can also examine how learners' InfoSB mediates the relationship between explicit instruction and learning outcomes.

8. Conclusion

The present study represents an important step toward conceptualizing and operationalizing InfoSB as one of the four major dimensions of proactive L2 learning ([Papi & Hiver, 2025](#)). By establishing InfoSB as a multidimensional construct, this study bridges the gap between formal instruction of L2 forms and learners' self-directed engagement with L2 forms.

The findings provide a nuanced understanding of how learners navigate the diverse landscape of explicit L2 information, seeking it out and using it to address their communicative needs and to intentionally enhance their linguistic skills. By highlighting learners' InfoSB, this study takes a step toward broadening the focus of L2 research from an exclusive focus on explicit (meta)linguistic information as an instructional resource to how learners seek out and make use of such information as a learning resource. This shift offers a more comprehensive understanding of the relationship between explicit learning and L2 acquisition, not as something pre-determined but rather shaped by the learner as the architect of this learning process. By validating the two-dimensional factor structure of InfoSB in L2 acquisition, this study paves the way for researchers to move beyond the view of learners as passive subjects who

process linguistic information to an organic view of learners as self-directed agents who use available resources to build the declarative knowledge that they desire, and convert that knowledge into procedural skills through intentional use. From this perspective, how effectively and efficiently such knowledge conversion is accomplished depends more on the learner than on how teachers or researchers manipulate the target information. Finally, the developed questionnaire can serve both as an empirical tool for investigating InfoSB in SLA and a guiding tool for understanding and enhancing this L2 learning behavior.

CRediT authorship contribution statement

Mostafa Papi: Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yuichi Suzuki:** Writing – review & editing, Writing – original draft, Conceptualization. **Claudia J. Kim:** Writing – review & editing, Writing – original draft, Formal analysis. **Miseong Kim:** Writing – review & editing, Writing – original draft, Formal analysis.

Data availability statement

The data are not publicly available due to IRB restrictions, but are available from the corresponding author upon reasonable request.

We have no known conflict of interest to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.system.2026.104147>.

Appendix

Information-Seeking Questionnaire

Note: Original numbering was modified for consistency.

Information Gathering

Q5. I ask others (teachers, peers, etc.) or use resources (apps, websites, AI tools, dictionaries) to find information about new Spanish/English vocabulary and idioms.

Q6. I ask others (teachers, peers, etc.) or use resources (apps, websites, AI tools, books, etc.) to find information about Spanish/English grammar.

Q7. I ask others (e.g., teachers, peers) or use resources (e.g., apps, websites, AI tools) to find information about how something is pronounced in Spanish/English.

Q8. When I learn a new grammar rule of English, I search for more information (examples, explanations, etc.) on how to use it correctly. (English only).

Q9. When I learn a new English word, phrase, or idiom, I search for more information (examples, explanations, etc.) to understand how to use it correctly. (English only)

Q10. I ask others (e.g., teachers, peers) or use resources (e.g., apps, websites, AI tools) to find information about different varieties of the Spanish language, such as slang, informal, formal, accents, dialects, etc.

Q11. I ask others (e.g., teachers, peers) or use resources (e.g., apps, websites, AI tools) to find information about Spanish-speaking cultures and their norms of appropriateness (being nice, respectful, friendly, etc.). (Spanish only)

Information Use

Q1. In my use of English, I try to apply the information that I've learned about vocabulary and idioms to communicate more efficiently.

Q2. In my use of English, I try to apply the new information that I've learned about new grammar rules to communicate more accurately in this language.

Q3. In my use of English, I try to apply the new information that I've learned about English pronunciation to communicate more fluently.

Q4. I try to apply the information that I've learned about English-speaking cultures and norms (such as being nice, respectful, friendly, etc.) to communicate more appropriately. (English only)

References

- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford Publications.
- Chiocia, E. S. (2025). The REFLECT model: Cultivating beginning language learner agency through eTandem conversations, transcriptions, and reflections. *Journal of China Computer-Assisted Language Learning*, (0)https://doi.org/10.1515/jccall-2025-0009
- Cohen, A. D. (1998). *Strategies in learning and using a second language*. Longman.
- Crowther, D., & Loewen, S. (2025). Instructed second language acquisition and second language pronunciation. *Language Learning*, 75(S1), 30–54. https://doi.org/10.1111/lang.12716
- DeKeyser, R. (2007). Skill acquisition theory. In B. Van-Patten, & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 94–113). Routledge.
- DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499–533. https://doi.org/10.1017/S0272263100004022
- DeKeyser, R. M., & Suzuki, Y. (2025). Skill acquisition theory. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition: An introduction* (4th ed., pp. 157–182). Routledge.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum.
- Ellis, N. C., & Wulff, S. (2025). Usage-based approaches to second language acquisition. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition: An introduction* (4th ed., pp. 42–63). Routledge.
- Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge?: A review of the research. *Studies in Second Language Acquisition*, 24(2), 223–236. http://www.jstor.org/stable/44486614
- Farhang, A., & Papi, M. (2025). Interaction-seeking behavior in second language learning: A preliminary exploration in the United States. *Innovation in Language Learning and Teaching*. https://doi.org/10.1080/17501229.2025.2584372
- Gade, M., Druey, M. D., Souza, A. S., & Oberauer, K. (2014). Interference within and between declarative and procedural representations in working memory. *Journal of Memory and Language*, 76, 174–194. https://doi.org/10.1016/j.jml.2014.07.002
- Gass, S. M., Mackey, A., & Montee, M. (2025). Input, interaction and output in L2 acquisition. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition: An introduction* (4th ed., pp. 183–216). Routledge.
- Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning: Norris & Ortega (2000) revisited and updated. In *Implicit and explicit learning of languages* (pp. 443–482). John Benjamins.
- Higgins, E. T., Kruglanski, A. W., & Pierro, A. (2003). Regulatory mode: Locomotion and assessment as distinct orientations. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, 35 pp. 293–344. Elsevier Academic Press. https://doi.org/10.1016/S0065-2601(03)01005-0
- Hu, Li-tze, & Bentler, Peter M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. https://doi.org/10.1080/10705519909540118. http://www.tandfonline.com/doi/abs/10.1080/10705519909540118
- Hulstijn, J. H. (2013). Incidental learning in second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*, 5 pp. 2632–2640). Wiley-Blackwell. https://doi.org/10.1002/9781405198431.wbeal0530
- Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), 539–558. https://doi.org/10.1111/0023-8333.00164
- Ishikawa, M., & Suzuki, W. (2023). Effects of written languaging on second language learning: Mediating roles of aptitude. *The Modern Language Journal*, 107(S1), 95–112. https://doi.org/10.1111/modl.12817
- Kadivar, N., & Papi, M. (2026). Motivational dispositions shape feedback-seeking behaviors: A regulatory mode perspective. *Language Awareness*, 1–20. https://doi.org/10.1080/09658416.2026.2702383. . (Accessed 20 July 2026)
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Leow, R. P. (2015). *Explicit learning in the L2 classroom*. Routledge.
- Leow, R. P., & Zamora, C. C. (2017). Intentional and incidental L2 learning. In *The routledge handbook of instructed second language acquisition* (pp. 33–49). Routledge.
- Li, S., Hiver, P., & Papi, M. (2022). Individual differences in second language acquisition. *The Routledge Handbook of Second Language Acquisition and Individual Differences* (pp. 3–34). New York: Routledge, ISBN 9781003270546. https://doi.org/10.4324/9781003270546-2. https://www.taylorfrancis.com/books/9781003270546/chapters/10.4324/9781003270546-2
- Mahbodi, M., Papi, M., & Wolff, D. (2025). How learners' trust in teachers shapes their feedback-seeking behaviors: A mixed-methods study. *Language Teaching Research*, 0(0). https://doi.org/10.1177/13621688251352268
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417–528. https://doi.org/10.1111/0023-8333.00136
- O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge University Press.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Newbury House.
- Oxford, R. L. (2017). *Teaching and researching language learning strategies: Self-regulation in context* (2nd ed.). Routledge.
- Papi, M., Abdi Tabari, M., & Sato, M. (2024). The importance of seeking feedback for benefiting from feedback: A case of second language writing. *The Modern Language Journal*, 108(2), 489–512. https://doi.org/10.1111/modl.12923
- Papi, M., Bondarenko, A. V., Wawire, B., Jiang, C., & Zhou, S. (2020). Feedback-seeking behavior in second language writing: Motivational mechanisms. *Reading and Writing*, 33, 485–505. https://doi.org/10.1007/s11145-019-09971-6
- Papi, M., Eom, M., Zhang, Y., Zhou, Y., & Whiteside, Z. (2023). Motivational dispositions predict qualitative differences in oral task performance. *Studies in Second Language Acquisition*, 45(5), 1261–1286. https://doi.org/10.1017/S0272263123000220
- Papi, M., & Hiver, P. (2025). Proactive language learning theory. *Language Learning*, 75(1), 295–329. https://doi.org/10.1111/lang.12644
- Papi, M., Rios, A., Pelt, H., & Ozdemir, E. (2019). Feedback-seeking behavior in language learning: Basic components and motivational antecedents. *The Modern Language Journal*, 103(1), 205–226. https://doi.org/10.1111/modl.12538
- Pawlak, M. (2025). Explicit and implicit (automatized) knowledge of second language pronunciation: Implications for theory, research, and classroom practice. *Foreign Language Annals*, 58(2), 346–366. https://doi.org/10.1111/flan.12803
- Qian, Y., Kadivar, N., & Papi, M. (2026). Exploring information-seeking behavior in second language learning: A preliminary study. *Innovation in Language Learning and Teaching*, 1–22. https://doi.org/10.1080/17501229.2026.2681105
- Qian, Y., & Papi, M. (2026). Interaction-seeking behavior in second language learning: The role of regulatory modes. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.70362. . (Accessed 3 September 2026)
- Rebuschat, P. (2015). *Implicit and explicit learning of languages*. John Benjamins.
- Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning*, 69(3), 652–708. https://doi.org/10.1111/lang.12345
- Sato, M. (2023). Skill learning theories and language teaching: Different strokes for different folks. In Y. Suzuki (Ed.), *Practice and automatization in second language research: Perspectives from skill acquisition theory and cognitive psychology* (pp. 63–86). Routledge.
- Sato, M., & McDonough, K. (2019). Practice is important but how about its quality? Contextualized practice in the classroom. *Studies in Second Language Acquisition*, 41, 999–1026. https://doi.org/10.1017/S0272263119000159
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158. https://doi.org/10.1093/applin/11.2.129
- Shrestha, N. (2021). Factor analysis as a tool for survey analysis. *American journal of Applied Mathematics and statistics*, 9(1), 4–11.
- Skehan, P., & Foster, P. (2013). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Investigating complexity, accuracy and fluency in SLA*. John Benjamins.

- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60(2), 263–308. <https://doi.org/10.1111/j.1467-9922.2010.00562.x>
- Suzuki, Y. (Ed.). (2023). *Practice and automatization in second language research: Perspectives from skill acquisition theory and cognitive psychology*. Routledge.
- Suzuki, Y., & DeKeyser, R. M. (2026). Explicit knowledge and skill acquisition in second language learning. In C. A. Chapelle, & Danijela Trenkic (Eds.), *In The Encyclopedia of Applied Linguistics (Second Language Acquisition and the Multilingual Mind)* (2th ed.). Oxford, UK: Wiley.
- Swain, M. (2006). Languaging, agency and collaboration in advanced second language proficiency. In H. Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95–108). London: Continuum.
- Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language teaching*. John Benjamins.
- Teimouri, Y., Papi, M., & Tahmouresi, S. (2022). Individual differences in how language learners pursue goals: Regulatory mode perspective. *Studies in Second Language Acquisition*, 44(3), 633–658. <https://doi.org/10.1017/S0272263121000413>
- Tigchelaar, M., Bowles, R., Winke, P., & Gass, S. (2017). Assessing the validity of ACTFL Can-Do statements for spoken proficiency: A Rasch analysis. *Foreign Language Annals*, 50(3), 584–600. <https://doi.org/10.1111/flan.12286>
- Toth, P. D. (2011). Teacher-and learner-led discourse in task-based grammar instruction: Providing procedural assistance for morphosyntactic development. *Language Learning*, 61, 141–188. <https://doi.org/10.1111/j.1467-9922.2008.00441.x>
- Ullman, M. T. (2025). The declarative/procedural model. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition: An introduction* (4th ed., pp. 121–156). Routledge.
- Webb, S., Yanagisawa, A., & Uchiyama, T. (2020). How effective are intentional vocabulary-learning activities? A meta-analysis. *The Modern Language Journal*, 104, 715–738. <https://doi.org/10.1111/modl.12671>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2