

A revised model of feedback-seeking behavior in SLA: toward improved conceptualization and measurement

Mostafa Papi ^{1*}, Hyejin An¹, Kevin Hirschi ²

¹School of Teacher Education, Florida State University, 1114 W. Call St., Stone Building, Tallahassee, FL 32306, United States

²College of Education and Human Development, The University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, United States

*Corresponding author. School of Teacher Education, Florida State University, 1114 W. Call St, G129 Stone Building, Tallahassee, FL 32306, USA. E-mail: mpapi@fsu.edu

.....

This study proposed and tested a revised model of feedback-seeking behavior in second language acquisition. The revised model included feedback elicitation, monitoring, and use. In two studies, data were collected from 207 learners enrolled in Spanish classes at a large public university in Florida and 294 learners of English at a multi-branch private language institute in Iran. Exploratory factor analyses generally combined CF Monitoring and CF Use, whereas confirmatory factor analyses favored the hypothesized three-factor model, with stronger support in Study 2. The scales demonstrated acceptable convergent validity and internal consistency, although the distinction between Monitoring and Use was limited in Study 1 and improved in Study 2. Hierarchical multiple regression analyses further showed that the overall FSB index significantly predicted self-assessed L2 proficiency in both contexts, although the unique contributions of the individual components varied across samples. Overall, the findings supported the tripartite model of FSB, defined as learners' proactive efforts to elicit, monitor, and use CF to enhance their L2 production. Theoretical and practical implications are discussed, and future research directions are outlined.

.....

Keywords: proactive language learning, feedback-seeking behavior, oral corrective feedback, second language acquisition

Introduction

Feedback-seeking behavior (FSB) in second language (L2) learning refers to learners' proactive efforts in seeking and using corrective feedback (CF) on their L2 performance (Papi et al. 2019, 2020; Carless 2026). To highlight the primacy of the learner in the feedback process, Papi et al. (2019) introduced this notion to the field of second language acquisition (SLA), which has been dominated by a cognitive and psycholinguistic approach that views CF as a teaching technique and focuses its empirical lens on the techniques, timing and scope of CF (e.g. Long 1996; Gass and Mackey 2006).

FSB was introduced to emphasize learners' proactive role in the CF process and to reframe CF as a resource learners actively seek, process, and utilize to advance their L2 learning goals.

FSB has been the subject of growing empirical research over the last seven years, since the topic was introduced to the field (Carless 2026). Researchers have taken an interest in exploring the individual and contextual factors that affect and shape FSB, as well as how this behavior influences learning and performance outcomes (Wang and Zheng 2026). Studies in different contexts show that learners' FSB was predicted by learner characteristics, such as their self-esteem, mindsets and achievement goals (Yao and Zhu 2024; Guo and Xu 2026; Zhang 2023), future L2 selves (e.g. Zhan et al. 2023; Li and Zhang 2026), self-efficacy beliefs (Xu and Wang 2024), grit (Zhang and Jiang 2025), perceived costs and benefits of FSB (Man et al. 2026), and use of cognitive, metacognitive, social, and motivational self-regulation strategies (Xu 2021). Contextual variables such as teacher academic support (Zhang 2025), students' trust in their teachers (Mahbodi et al. 2025; Papi et al. 2026), and the availability of low-cost, flexible feedback opportunities have also been found to contribute to student FSB in both in-person and online environments (Kessler 2023). Finally, learners' FSB has been shown to contribute to their feedback literacy and ecology (Li and Han 2024), as well as to their L2 performance (e.g. Zhan et al. 2023; Papi et al. 2024).

The findings of the FSB studies conducted in the field, however, have not been symmetrical for both components of the FSB model. That is, whereas feedback monitoring, an FSB strategy which involves paying attention and trying to learn from the CF, has been associated with adaptive individual or contextual variables, feedback inquiry, which involves directly asking for CF, has been either weakly associated or not associated with such variables. For instance, in the context of L2 writing, Papi et al. (2024) found that only feedback monitoring predicted improvement in revision quality after students received written CF from their teachers. In contrast, feedback inquiry was unrelated to any of the L2 written performance measures. In the context of oral CF, studies have found feedback monitoring to be associated with adaptive motivational variables such as a growth mindset, development goals, and the ideal L2 self; by contrast, feedback inquiry has not shown parallel connections with those variables (e.g. Bondarenko 2020; Papi et al. 2019, 2020). In one study, feedback inquiry was in fact associated with a desire for the most implicit or no oral CF at all (Papi et al. *in press*).

Researchers have argued that these results could be attributed to the limited corrective function of feedback inquiry, which can instead serve as an impression-management or social-bonding strategy when feedback is sought from teachers or peers (e.g. Mahbodi et al. 2025; Papi et al. 2019, 2020). For instance, some questionnaire items, such as 'I ask my teachers to tell me how I can improve my (L2) speaking', which are often used for measuring feedback inquiry, do not necessarily tap into students' desire for correcting their errors and may represent a request for general L2 learning strategies. In other words, rather than tapping into a strategy exclusively for seeking correction, feedback inquiry may have served as a means of establishing or enhancing socio-emotional relationships with feedback sources, such as the teacher.

CF, however, is commonly seen in the field of SLA to have a corrective function. 'It is generally considered to be a form of negative evidence, or information that a particular utterance is deviant vis-à-vis target language norms' (Gass and Mackey 2006: 7). As a dimension of proactive language learning (Papi and Hiver 2025), FSB components are also expected to align with that definition and contribute to L2 knowledge through their corrective function. Given the empirical evidence and theoretical arguments, the construct of feedback inquiry does not appear to qualify as a meaningful strategy for seeking CF; thus, it needs to be redefined and re-operationalized accordingly. When viewed as a request for corrective rather than directive feedback, feedback inquiry has been associated with more intrinsic forms of motivation, complementing the monitoring strategy (Zhou et al. 2025). The present study aims to address this issue and reconceptualize feedback inquiry as an FSB construct that taps into learners' direct requests for corrective rather than directive feedback on their problematic L2 output. In addition, given the emerging interest in exploiting proactive behaviors through digital technology (Liu and Zhao 2026) and learners' reports of seeking feedback from artificial intelligence tools such as ChatGPT (Yan et al. 2026), including such technologies as potential sources of feedback appears both necessary and aligned with current developments in SLA research.

asking for CF on one's L2 performance, and feedback monitoring, which involves paying attention to and learning from CF. From the earliest studies, however, results have shown that CF inquiry and monitoring are associated with different types of individual, contextual, and performance variables. For instance, in their original study of 287 university foreign language learners in the United States, Papi et al. (2019) found that students with a growth mindset and a learning-approach goal engaged in more feedback monitoring than inquiry. In addition, only students with performance goals engaged in feedback inquiry, suggesting that learners who engage in feedback inquiry might have secondary goals, such as making a positive impression on their teachers (see also Yao and Zhu 2024). Drawing on the expectancy-value perspective, Wei (2026) found that a growth mindset did not directly predict FSB but did so indirectly through motivational factors such as self-efficacy, utility value, and interest, depending on participants' proficiency level. In a translation context, Man et al. (2026) reported that the perceived utility value of feedback positively predicted both feedback inquiry and monitoring. In contrast, a fixed mindset was a positive predictor of the self-presentation cost of FSB, which negatively predicted feedback monitoring but not inquiry. Zhang (2025) found that both ideal and ought-to L2 writing self were associated with feedback monitoring but not inquiry. The impression management function of feedback inquiry in the context of L2 speaking was also supported by the findings of Mahbodi et al.'s (2025) study, which found that L2-Spanish students' trust in their teachers' competence positively predicted their feedback monitoring but negatively predicted their feedback inquiry. Papi et al. (2026) found feedback monitoring to be associated with learner preference for the most explicit types of oral CF, whereas feedback inquiry was only associated with the most implicit types or a preference to receive no CF on their speaking errors at all.

Two studies have shown that feedback monitoring is more consistently associated with L2 learning and performance than feedback inquiry is. In China, Zhan et al. (2023) found that for mid-achieving students, feedback monitoring positively and feedback inquiry negatively mediated the relationship between high school students' ideal English writing self and their writing performance. For high-achieving students, the relationship was inverse: monitoring negatively and feedback inquiry positively mediated the relationships. In a study of ESL writing students in the United States, Papi et al. (2024) found that after controlling for the number of teacher comments on ESL learners' essays, feedback monitoring positively predicted their revision quality, whereas feedback inquiry did not.

While feedback inquiry and monitoring, as two subcomponents of an FSB model, should consistently be associated with adaptive individual, contextual, and performance variables, the studies show that the inquiry component has not met these expectations. As discussed earlier, this asymmetry can be attributed to the lack of a corrective focus in the operationalization of this construct. Papi et al. (2019) drew on Ashford and associates' work in organizational psychology (e.g. Ashford 1986), which does not seem to directly extend to the context of SLA where L2 learning is commonly viewed as acquiring a set of skills that need to be gradually acquired through correction of erroneous utterances, among other things (Long 1996; Lyster and Ranta 1997). To address this shortcoming, the present study will explore how creating a corrective conceptual focus for this construct (CF elicitation) can lead to differences in learners' self-assessed oral L2 proficiency. Oral L2 proficiency was selected as our focal outcome measure because the present study explores FSB in oral L2 interaction, where learners have frequent and immediate opportunities to elicit, monitor, and utilize CF during real-time language use. In other words, oral FSB is fundamentally interaction-driven and theoretically expected to influence oral interaction skills, which justifies using oral L2 proficiency as the outcome variable. Moreover, in L2 learning contexts, beyond eliciting and monitoring corrective feedback, consistent and intentional use of the CF is essential for the acquisition of linguistic features. Given the central role of CF uptake and use in L2 learning, the present study incorporates a third component to the FSB model representing the extent to which learners actively use the CF they receive. Reconceptualizing feedback inquiry to capture the corrective function of CF and the inclusion of CF use in the revised FSB model are grounded in major theoretical perspectives on CF processes in SLA, to which we turn now.

Theoretical background

Despite the persistence of negative views of CF in the field (e.g. [Truscott 2023](#)), there is a growing consensus on its utility in L2 development. From a theoretical perspective, CF is a central focus of inquiry in SLA ([Lyster and Ranta 1997](#)) and is considered essential for driving language development ([Long 1996](#)). The utility of CF in L2 learning can be attributed to its function in providing negative evidence that can be used to raise a learner's awareness of their erroneous L2 utterances and lead them to modify their following L2 utterances. [Long's \(1980\)](#) interaction hypothesis posits that CF serves as a fundamental mechanism that helps learners notice the gap between their problematic production and what is correct, and modify their output as they negotiate meaning. Acquisition is posited to be more likely with increased awareness of both target and incorrect forms ([Schmidt 1990](#)). From a skill-acquisition perspective, [Dekeyser \(2020\)](#) posits that CF provides declarative knowledge that can help learners reformulate their L2 utterances, which can be proceduralized through further practice. Usage-based theories (e.g. [Ellis 2010](#)) also view CF as beneficial, as it can provide marked, target-like input for further pattern building. Such a transformation from corrective information to L2 modification, however, does not occur automatically or consistently for all learners, and it involves how learners seek, monitor, and make use of CF ([Papi et al. 2019](#)).

According to [Long \(1996\)](#), the key to successful feedback incorporation is how the learner navigates noticing, processing, and adjusting their output after receiving CF. However, not all learners engage equally with CF, and not all types of CF are noticed by learners or result in modified output ([Lyster and Ranta 1997](#)). The relationship between CF and learner beliefs ([Han and Hyland 2015](#); [Han 2017](#)), preferences ([Amrhein and Nassaji 2010](#); [Akiyama 2017](#)), motivation ([Waller and Papi 2017](#); [Papi et al. 2021](#)), and a variety of background characteristics ([Han and Hyland 2015](#)) also suggests that learner experience and context shape their response to CF. For instance, [Han \(2017\)](#) found that learner beliefs about CF were associated with learner engagement with CF. [Akiyama \(2017\)](#) found that CF uptake was maximized when the CF type matched the learner's preference. Collectively, these variables represent important precursors to learner engagement with CF and align with the notion of feedback monitoring and use, which will be examined in this study.

Furthermore, SLA research and theory have centralized students' reactions to CF as their focus of analysis and thinking. However, a close reading of several CF studies in the literature has long suggested that learners elicit CF. For instance, in [Lyster and Ranta's \(1997\)](#) cornerstone taxonomy of CF, two interaction excerpts commence with a learner's turn and a noun phrase followed by a question mark ('L'eau érablé?' on p. 47 and 'Le... le girafe?' on p. 48), suggesting the learners' indirect requests prompted the teacher to provide CF. In recent literature, [Zhang and Zhang's \(2023\)](#) analysis of feedback episodes in L2 writing found that some learners took the initiative in resolving issues in CF. From an interactionist perspective, CF elicitation involves learners initiating CF sequences that are believed to drive development in communicative episodes ([Long 1996](#); [Gass and Mackey 2006](#)). By explicitly requesting clarification or correction ([Lyster and Ranta 1997](#)), learners can maximize feedback ([Zhang and Zhang 2023](#)), thereby increasing opportunities to modify their output.

When learners habitually and strategically prompt interlocutors for correction, they increase the probability, salience, and immediacy of feedback episodes, all of which have been shown to predict uptake and subsequent accuracy gains across a wide variety of contexts ([Li 2010](#)). The notion of CF elicitation aligns with the interactionist claim that development can be accelerated when learners create, rather than merely participate in, feedback-rich exchanges that direct attention to language form ([Mackey et al. 2000](#)). Further opportunities for learners to seek feedback are presented by technologies such as ChatGPT ([Yan et al. 2026](#)). Conceptualizing FSB as learners' effort in eliciting, monitoring, and using CF, therefore, aligns the model with core interactionist mechanisms—request, feedback, modification—that are believed to contribute to L2 development, creating alignments between the revised FSB model and SLA theory.

Research questions

Based on the literature reviewed above and the need for a reconceptualization of inquiry as a strategy with a corrective function and the addition of a component representing students' use of CF, the present study aims to test a revised FSB model that includes three hypothetical components: CF Elicitation, CF Monitoring, and CF Use. We will examine the construct validity, convergent validity, and discriminant validity of the new scales using factor-analytic methods. Additionally, the predictive validity of the FSB constructs will be assessed in relation to students' self-assessed interactive oral L2 proficiency. Based on these objectives, the following research questions will be addressed:

- RQ1** To what extent does the revised conceptual model of FSB demonstrate structural validity through its hypothesized factor structure?
- RQ2** How does the revised model of FSB demonstrate convergent and discriminant validity?
- RQ3** How do the constituent components of the FSB model predict students' self-assessed L2 proficiency?

Methods

Study 1. Spanish Learners in the US

The study was conducted at a university in Florida, a US state with a significant Spanish-speaking population. Spanish is the most common home language among English language learners in the state, comprising about 75 per cent of this population. Given the popularity of the language, universities in Florida offer a range of courses to help students develop Spanish proficiency. This setting provides an ideal environment to explore feedback-seeking behavior among Spanish L2 learners.

Participants

Two hundred seven students (77 males, 126 females, and 4 others) enrolled in L2-Spanish courses at a university in Florida participated in the study. Participants' ages ranged from 17 to 58 years ($M = 21.62$, $SD = 3.31$), and the vast majority were undergraduate students ($n = 203$). Nearly all participants were native English speakers ($n = 206$), with one participant reporting Bengali as their first language. The students represented a range of academic majors, including Psychology ($n = 32$), Biology ($n = 24$), and Computer Science ($n = 11$). Approximately 70 per cent ($n = 145$) were enrolled in 1,000- or 2,000-level courses (beginner level), while the remaining participants were enrolled in 3,000-level or higher courses (intermediate to advanced levels). The mean score for L2 proficiency was 2.18 ($SD = 0.82$), indicating that participants were, on average, at an intermediate level. Finally, their mean Length of L2 Study was 3.49 years ($SD = 3.01$).

Procedures

Following Institutional Review Board (IRB) approval, instructors of Spanish classes at the same university were contacted to facilitate data collection. A researcher visited the classes of the instructors who agreed to participate in the study. The students were informed about their voluntary participation, the anonymity of their data, and the confidentiality of their data. The data were collected in person using paper questionnaires.

Instruments

A questionnaire (Appendix A) was used to collect data. It consisted of three sections: Demographic Information (e.g. age, first language, and major), a short version of ACTFL Spanish Proficiency Self-Assessment (Tigchelaar et al. 2017), and 15 items measuring CF elicitation (5 items), monitoring (5 items), and use (5 items). Participants responded to the items using a 6-point Likert scale ranging from 1 (*Never/Rarely*) to 6 (*Constantly*). The instrument also included items for measuring other types of proactive behaviors, which are not reported in this study. The Spanish Proficiency

Data analysis

To further evaluate the convergent and discriminant validity of the FSB model, confirmatory factor analysis (CFA) with maximum likelihood estimation was performed using AMOS 29. The CFA tested the fit of alternative models and provided evidence for the distinctiveness and coherence of the identified factors. Finally, hierarchical multiple regression analyses were conducted using SPSS 29 to examine the predictive validity of the FSB components.

Exploratory factor analysis. The initial EFA yielded a three-factor solution, as determined by eigenvalues greater than 1 and scree plots. The initial analysis yielded three factors, with pragmatics-related items creating their own factor. Because the three pragmatic items did not represent a theoretically meaningful construct and instead reflected a mixture of elicitation, monitoring, and use, they were removed prior to a second EFA. The follow-up analysis identified a two-factor solution, with Factor 1 comprising eight items that captured both CF Use and Monitoring, accounting for 62.61 per cent of the total variance, and Factor 2 included four items that measured CF Elicitation, accounting for an additional 5.89 per cent of the variance, not aligning with the hypothesized three-factor model.

Although the chi square tests were significant for all models, this index is known to be highly sensitive to large sample sizes (e.g. $N > 200$; Kline 2016), limiting its usefulness as a sole indicator of model fit. Therefore, additional goodness-of-fit indices were considered, including the Comparative Fit Index (CFI; acceptable $>.90$), the Tucker–Lewis Index (TLI; acceptable $>.90$), and the Root Mean Square Error of Approximation (RMSEA; acceptable $<.08$). Alternative models were further compared using chi-square difference tests, the Akaike Information Criterion (AIC), and the Browne–Cudeck Criterion (BCC). The results for the three models are presented in Table 1, showing that the three-factor model and the two-factor model demonstrated significantly better fit than the one-factor alternative across the fit indices. In addition, the three-factor model showed a marginally better fit than the two-factor model, with the chi-square difference test being significant.

Convergent validity was supported for the three-factor model, as the average variance extracted (AVE) (Table 3) exceeded the .50 criterion (Fornell and Larcker 1981) for each construct (CF Elicitation = .56; CF Use = .78; CF Monitoring = .75). However, squared inter-construct correlations (SICs) exceeded the corresponding AVEs between CF Monitoring and CF Use, indicating insufficient

	χ^2/df ($N = 207$)	CFI	TLI	RMSEA	$\Delta\chi^2$	AIC	BCC
3-Factor model	154.32/51***	.954	.93	.099	—	232.32	237.57
2-Factor model	169.04/53***	.948	.92	.103	14.72***	243.04	248.02
1-Factor model	271.034/54***	.903	.86	.140	116.72***	343.03	347.88

^{***} $P < .001$.

	CF elicitation	CF use
CF use	.65 ^{***}	—
CF monitoring	.71 ^{***}	.91 ^{***}

*** $P < .001$.

	<i>k</i>	<i>M</i>	<i>SD</i>	α	ω	<i>CR</i>	<i>AVE</i>
Elicitation	4	3.94	1.19	.81	.82	.82	.56
Monitoring	4	4.42	1.15	.92	.92	.92	.75
Use	4	4.69	1.13	.93	.93	.93	.78

Coefficient omega (ω) tests the reliability of measurement scales that do not satisfy assumptions of tau-equivalence. CR, composite reliability; AVE, average variance extracted.

discriminant validity and suggesting that the conceptual distinctiveness between the two constructs may not be adequately supported due to the high correlation between Monitoring and Use (Table 2).

Finally, as presented in [Table 3](#), internal consistency for the three constructs was high, with McDonald's ω , Cronbach's alpha, and composite reliability values exceeding the acceptable threshold of .70. Descriptive statistics ([Table 3](#)) showed that participants reported the highest engagement in CF Use ($M = 4.69$, $SD = 1.13$), followed by CF Monitoring ($M = 4.42$, $SD = 1.15$), and CF Elicitation ($M = 3.94$, $SD = 1.19$).

Although the three-factor model yielded the best overall fit, the modest improvement over the two-factor model, along with evidence of high correlations, suggests potential overlap among the factors, especially CF Monitoring and Use. Taken together, these results suggest that FSB is not adequately represented as a single construct and is better conceptualized as comprising at least two, and potentially three, dimensions. As presented in [Table 4](#), all items loaded significantly on their intended latent constructs, with standardized loadings ranging from .47 to .92 ($P < .001$). Overall, the results support the adequacy of the three-factor measurement model ([Table 4](#)).

Multiple regression

Hierarchical multiple regressions were conducted to examine the extent to which each FSB construct (CF Elicitation, Monitoring, and Use) explained variance in the L2 proficiency in an incremental way. In addition, Length of Study was included as a covariate, given its expected influence on L2 proficiency (Hakuta et al. 2000; Zhang et al. 2020). The analysis, however, resulted in collinearity due to the high correlation between Monitoring and Use. Therefore, the two variables were combined to create a composite score, referred to as Monitoring/Use, and entered into

Table 4 Standardized and unstandardized CFA factor loadings for the three-factor model (US data).

Latent constructs	β	<i>B</i>
CF elicitation		
Q1	.76	—
Q2	.47	0.66
Q3	.88	1.05
Q4	.82	1.00
CF monitoring		
Q5	.84	—
Q6	.86	1.05
Q7	.92	1.18
Q8	.84	1.05
CF use		
Q9	.89	—
Q10	.88	0.99
Q11	.88	1.12
Q12	.88	1.21

N = 207. All coefficients significant at the *P* < .001 level.

Table 5 Regression results with the proficiency self-rating scores as the outcome variable (US data).

		Beta	<i>t</i>	<i>P</i>	Tolerance	VIF
Model 1	Length of L2 study	.47	7.36***	<.001	1.00	1.00
	Elicitation	.23	3.67***	<.001	1.00	1.00
Model 2	Length of L2 study	.46	7.16***	<.001	0.97	1.03
	Elicitation	.20	2.21*	.03	0.52	1.94
	Monitoring/use	.05	0.61	.55	0.51	1.96

P* < .05, **P* < .001.

another regression analysis. The results showed that, in the base model with Elicitation and Length of Study as predictors, both variables significantly predicted L2 proficiency. When Monitoring/Use was added to the second model, Elicitation remained a significant predictor, whereas Monitoring/Use did not emerge as a significant predictor.

To further explore the relationships, we conducted another hierarchical regression analysis in which Monitoring/Use was entered before Elicitation (Table 6). The results in the base model showed that Monitoring/Use significantly predicted L2 proficiency along with Length of Study when Elicitation was not added to the analysis, suggesting that Elicitation accounts for the variance explained by Monitoring/Use in this data. Finally, to assess the predictive power of the FSB construct as a whole, an FSB Index was created by combining Elicitation and Monitoring/Use, which was entered as a predictor alongside Length of L2 Study. As shown in Table 6, the two variables significantly predicted 27 per cent of variance in L2 proficiency.

Study 2: English learners in Iran

Participants

A total of 294 students (82 males, 212 females) enrolled in an English language institute in Iran participated in Study 2. Most participants were high school students (*n* = 234), with smaller groups of undergraduate (*n* = 45) and graduate students (*n* = 15). Their mean age was 17.94 years (*SD* = 5.06). Nearly all participants were native speakers of Farsi (*n* = 283), and

Table 7 Comparison of tested models (Iran data).

	χ^2/df ($N = 294$)	CFI	TLI	RMSEA	SRMR	$\Delta\chi^2$	AIC	BCC
3-Factor model	101.74/51***	.98	.98	.06	.02	—	179.74	183.36
2-Factor model	174.06/53***	.96	.95	.09	.03	72.32***	248.06	251.49
1-Factor model	367.07/54***	.90	.87	.14	.06	265.34***	439.07	442.42

*** $P < .001$.

Table 8 Descriptives and reliabilities of final scales in the three-factor model (Iran data).

	k	M	SD	α	ω	CR	AVE
Elicitation	4	4.81	1.06	.90	.90	.90	.69
Monitoring	4	4.77	1.04	.92	.92	.91	.72
Use	4	4.86	1.00	.90	.90	.90	.70

Coefficient omega (ω) tests the reliability of measurement scales that do not satisfy assumptions of tau-equivalence. CR, composite reliability; AVE, average variance extracted.

Table 9 Pearson correlations (Iran data).

	CF elicitation	CF use
CF use	.71***	—
CF monitoring	.76***	.83***

*** $P < .001$.

Results

Exploratory factor analysis. In the initial EFA, items tapping CF on pragmatics again emerged as a source of structural distortion and were thus removed. The final EFA yielded a stable two-factor solution, supported by eigenvalues greater than 1 and the scree plot. Similar to Study 1, Factor 1 consisted of eight CF Use and Monitoring items and accounted for 63.86 per cent of the total variance. Factor 2 captured four CF Elicitation items and explained 5.39 per cent of the variance.

Confirmatory factor analysis. As in Study 1, CFAs were conducted to test alternative models, as well as the convergent and discriminant validity of the scales (see the section ‘Study 1’ for details). The evaluation of the CFI, TLI, RMSEA, and SRMR indices (Table 7) indicated that the three-factor model provided a significantly better fit than the two-factor and one-factor alternatives. Convergent validity was also supported, with the AVEs exceeding the recommended .50 thresholds for all constructs (Table 8). In addition, AVEs exceeded the corresponding SICs, indicating reasonable discriminant validity. Global model fit indices and model comparison results provided strong support for the three-factor solution. In addition, the Pearson correlation of .83 between Monitoring and Use (Table 9) indicates improved discriminant validity of the two variables compared with the .91 correlation in Study 1 (Table 2), which is below the common threshold of .85 that raises concerns about discriminant validity (Brown 2015). Finally, internal consistency reliability for the three constructs was high, as evidenced by McDonald’s ω , Cronbach’s alpha, and composite reliability values (Table 8). Finally, all items loaded significantly on their intended latent

Table 10 Standardized and unstandardized CFA factor loadings for the three-factor model (Iran data).

Latent constructs	β	<i>B</i>
CF elicitation		
Q1	.78	—
Q2	.85	1.01
Q3	.86	1.16
Q4	.86	1.14
CF monitoring		
Q5	.79	—
Q6	.89	1.16
Q7	.86	1.18
Q8	.90	1.24
CF Use		
Q9	.78	—
Q10	.87	1.06
Q11	.83	1.07
Q12	.87	1.08

N = 294. All coefficients significant at the *P* < .001 level.

Table 11 Hierarchical regression with ACTFL long scale as outcome variable (Iran data).

		β	<i>t</i>	<i>P</i>	Tolerance	VIF
Model 1	Length of L2 study	.12	2.17*	.03	1.00	1.00
<i>R</i> ² = .09	Elicitation	.27	4.73***	<.001	1.00	1.00
	Length of L2 study	.13	2.37*	.02	1.00	1.00
Model 2	Elicitation	-.05	-0.65	.51	0.42	2.38
<i>R</i> ² = .16	Monitoring	.42	5.06***	<.001	0.42	2.38
	Length of L2 study	.12	2.27*	.02	0.99	1.01
Model 3	Elicitation	-.09	-1.09	.28	0.40	2.51
<i>R</i> ² = .18	Monitoring	.28	2.57*	.01	0.24	4.10
	Use	.20	2.01*	.04	0.28	3.57
The index model	L2 study length	.12	2.17*	.03	1.00	1.00
<i>R</i> ² = .15	FSB index	.37	6.76***	<.001	1.00	1.00

P* < .05, **P* < .001.

constructs, with standardized loadings ranging from .78 to .90 further supporting the adequacy of the three-factor measurement model (Table 10).

Multiple regression analyses

A hierarchical regression model was run, where Length of Study (covariate), Elicitation, Monitoring, and Use were entered in order, with L2 proficiency (Long ACTFL self-assessment scores) as the outcome variable. The results of the analyses (Table 11) indicate that both Length of Study and CF Elicitation were significant predictors in the base model, accounting for 9 per cent of the variance. When CF Monitoring was added in Model 2, it emerged as a significant predictor, whereas CF Elicitation was no longer significant. The model explained 16 per cent of the variance. In the final model, which included all three FSB constructs, CF Monitoring and CF Use emerged as significant predictors of L2 proficiency, while CF Elicitation remained non-significant. In addition, no serious multicollinearity issues were observed in this dataset, with VIF values below 5 and tolerance values above .20.

Finally, an additional multiple regression analysis was conducted using the FSB Index, which was formed by summing the mean scores for all three FSB variables. Consistent with Study 1,

the FSB Index emerged as a significant predictor, and along with Length of L2 Study, accounted for 15 per cent of the variance in the outcome variable (see [Table 11](#)), underscoring the overall importance of FSB in explaining L2 proficiency.

Discussion

The purpose of this study was to expand the theoretical scope and enhance the psychometric characteristics of the FSB constructs, originally developed by [Papi et al. \(2019\)](#). Motivated by the results of previous studies regarding feedback inquiry, the present study recalibrated the construct of inquiry, focusing on the corrective nature of students' CF elicitation while engaged in oral interaction. Therefore, the construct was named CF Elicitation to reflect these features. In addition, building on theoretical and empirical grounds in CF research, a scale was developed to measure the use or uptake aspect of FSB. The original questionnaire ([Papi et al. 2019](#)) was revised and used to collect data from learners of Spanish in the United States and learners of English in Iran. CFA results provided support for the emergence of three factors, particularly in Study 2, following the modification of items used in Study 1.

The first factor, CF Elicitation, represents learners' intentional attempts to elicit CF when they are not certain about the accuracy of their oral L2 utterances. That is, when they suspect they might have made grammatical, lexical, or pronunciation errors while speaking English, they ask other individuals (teachers or peers) or use digital resources (e.g. ChatGPT) for CF. CF Elicitation thus functions as a gateway strategy that helps students gain access to the feedback needed when it is not already available. Researchers have proposed that students taking such initiatives signal motivational readiness to process and use CF ([Mackey et al. 2000](#); [Gass and Mackey 2006](#)), which leads to the production and saliency of CF sequences ([Li 2010](#); [Lyster and Ranta 1997](#)) and maximizes the feedback they receive ([Zhang and Zhang 2023](#)).

The second factor included items that traditionally measured the monitoring component of FSB ([Papi et al. 2019](#)), representing learners' attention to and attempts to understand the CF provided on their various L2 errors (e.g. grammar, vocabulary, pronunciation). This FSB strategy has been linked to various adaptive motivational, affective, and contextual variables and has been shown to predict L2 performance (e.g. [Papi et al. 2019, 2020, 2024](#)). In addition, learners' attention to CF has been the focus of much SLA research and theorizing. For instance, [Schmidt \(1990\)](#) emphasized that learners can gain from CF when they notice, pay attention to, and become aware of the CF that fills a gap in their L2 knowledge. [Dekeyser \(2020\)](#) views CF as declarative information that learners can understand and use to improve through practice. The notion of feedback monitoring encompasses all the cognitive processes that learners engage in in relation to CF and equates the notion of cognitive engagement with CF. Such engagement concerns 'whether or not the information contained in the feedback is attended to, noticed, and processed in the right way' ([Pawlak 2013](#): 69).

The third factor, CF Use, which was added to the original FSB model in this study, included items measuring learners' attempts to use the CF provided on errors in their oral L2 production. Although CF Use and Monitoring did not demonstrate desirable discriminant validity in the first study, modifications in Study 2 led to improved distinctions. More specifically, CFA results suggested that CF Monitoring and CF Use represent closely related yet conceptually distinct FSB dimensions rather than a single undifferentiated construct. CF Use equates the behavioral dimension of CF engagement, where 'the focus is on whether and in what way learners uptake oral corrections' ([Ellis 2010](#): 342). The addition of the use dimension to FSB also aligns with the interaction perspective, which views CF as a prompt that contributes to acquisition not only by raising learners' awareness of their errors but also through leading them to adjust their erroneous L2 output ([Long 1996](#)). From [Dekeyser's \(2020\)](#) perspective, CF can lead to proceduralized L2 knowledge only when it leads to modified practice.

The results of hierarchical multiple regression analyses showed that the FSB predictors functioned differently across the two contexts. In the L2-Spanish context of the USA, CF Monitoring/Use was a significant predictor of L2 proficiency alongside Length of L2 Study. However, when

CF Elicitation was added to the model, it significantly predicted L2 proficiency, but CF Monitoring/Use was no longer significant, suggesting that CF Elicitation accounts for the variance explained by Monitoring/Use in L2 proficiency. In the context of Iran, Elicitation was also a significant predictor when it was entered into the base model. However, it lost significance when CF Monitoring and CF Use were added in the next two models. Contrary to the previous studies, which did not produce results favorable to feedback inquiry (e.g. Papi et al. 2024), CF Elicitation emerged as a predictor of L2 proficiency in both L2 Spanish and L2-English data in this study. These results confirm that this study has achieved its first objective of re-operationalizing feedback inquiry in a way that represents genuine requests for CF during oral interaction rather than for other purposes, such as impression management (e.g. Papi et al. 2020). Whereas the original scales developed for measuring Feedback Inquiry (Papi et al. 2019) included eight items asking the teacher or others 'to give me feedback', 'show me strategies/suggestions' on 'how I can improve', only three items specifically focused on the request 'to correct me when I make a mistake speaking.' By contrast, the new CF Elicitation scale developed in this study exclusively focuses on requests for correction of grammatical, lexical, and pronunciation errors among others, achieving the first objective of this study in capturing the corrective function of CF.

Relatedly, the inclusion of digital tools (e.g. ChatGPT) as potential sources of feedback in the CF Elicitation items reflects recent developments in SLA research. Learners are increasingly engaging in proactive learning behaviors such as feedback seeking (e.g. Liu and Zhao 2026) through AI-powered technologies and digital platforms, where immediate and individualized feedback opportunities are accessible (Kessler 2023; Liu and Hwang 2024; Yan et al. 2026). Although the present study did not specifically examine the distinct role of AI-mediated feedback sources, including them within the broader conceptualization of feedback sources, it reflects the growing role of digital technologies in proactive language learning and aligns the revised FSB framework with emerging trends in SLA research.

Differences in the L2-Spanish and L2-English data reflect differences in the sociocultural contexts of Florida and Iran. Students' use of Spanish in Florida is not limited to the classroom context. Proactive students can gain access to various contexts where they can use Spanish in real-life situations. Even though it is not commonly expected of native speakers to correct L2 speakers' erroneous utterances in social contexts, students' direct elicitation of CF can result in the reception of the desired CF and help learners develop their L2 proficiency. CF Elicitation can therefore represent learners' proactivity more effectively than CF Monitoring/Use, which only comes into play if CF is already provided. In the context of Iran, however, students' use of English is primarily limited to their classroom, where plenty of CF is provided on their faulty oral L2 production, obviating the need to ask for more CF. In the abundance of the CF provided, CF Monitoring and CF Use seem to play the primary role in developing Iranian learners' L2 proficiency.

The results of this study support the original rationale that a model of FSB in L2 learning with three components representing how learners elicit, reflect on, and use CF to improve their L2 knowledge and performance provides a more accurate and comprehensive understanding of learner involvement in the feedback process than the original FSB model (Papi et al. 2019). Proactive learners are expected to seek out CF on their erroneous L2 production, try to fully understand the CF on their errors, and utilize the CF to improve the accuracy of their L2 production. The inclusion of three components representing elicitation, monitoring, and use in an FSB model in SLA, thus, expands the previous view of FSB as inquiry and monitoring (Ashford 1986; Papi et al. 2019).

Conclusions

FSB was originally defined as 'learners' intentional, calculated, and strategic attempts to gather feedback information' on their L2 performance (Papi et al. 2020: 486). In light of the findings of the present study, FSB is defined as learners' self-directed efforts to elicit, monitor, and use corrective feedback to enhance their L2 production. The new instrument provides more comprehensive

Finally, the strong relationships between the FSB constructs, as well as the overall FSB index, and L2 proficiency underscore the importance of CF in learning outcomes in diverse learning contexts. Although this was not an experimental or longitudinal study, the association between different FSB and L2 proficiency further supports the value of feedback in L2 learning, which contradicts previous claims about the ineffectiveness of CF in L2 acquisition (e.g. [Truscott 2023](#)).

According to Brown (2015), ‘in applied research, a factor correlation that equals or exceeds .85 is often used as the cutoff criterion for problematic discriminant validity’ (p. 147). Although the correlation between Monitoring and Use did not reach this cutoff point after modifications to the items, the findings suggest that Monitoring and Use may represent closely related, and potentially sequential, phases of CF processing. In other words, learners’ attention to and processing of feedback (monitoring) may naturally co-occur with, or lead directly to, attempts to incorporate that CF into subsequent use. From this perspective, the two constructs may represent interconnected aspects of a broader behavioral pattern rather than entirely independent entities. At the same time, the enhanced distinction shown in Study 2 supports their conceptual distinction. Taken together, these findings highlight the need for further research to more clearly delineate the boundary between Monitoring and Use and determine if the two are best modeled as separate constructs or integrated within a unified measurement framework.

The FSB model can be instrumental for teachers in helping their students navigate feedback environments and benefit from the available feedback. The learner-centered perspective underlying

- Carless, D. (2026) 'Feedback Literacy', *International Journal of TESOL Studies*, 8: 1–16. <https://doi.org/10.58304/ijts.260101>
- Dekeyser, R. (2020) 'Skill Acquisition Theory', in B. Van Patten, G. D. Keating, and S. Wulff (eds.) *Theories in Second Language Acquisition*, pp. 83–104. Routledge.
- Ellis, R. (2010) 'EPILOGUE: A Framework for Investigating Oral and Written Corrective Feedback', *Studies in Second Language Acquisition*, 32: 335–49. <https://doi.org/10.1017/S0272263109990544>
- Fornell, C., and Larcker, D. F. (1981) 'Evaluating Structural Equation Models with Unobservable Variables and Measurement Error', *Journal of Marketing Research*, 18: 39–50. <https://doi.org/10.2307/3151312>
- Gass, S. M., and Mackey, A. (2006) 'Input, Interaction, and Output: An Overview', *AILA Review*, 19: 3–17. <https://doi.org/10.1075/aila.19.03gas>
- Guo, Q., and Xu, Y. (2026) 'Exploring Students' Feedback-Seeking Behavior: The Roles of Self-Esteem, Self-Doubt, and Achievement Goals from a variable-and Person-Centered Approach', *Studies in Higher Education*, 1–18. <https://doi.org/10.1080/03075079.2025.2612104>
- Hakuta, K., Butler, Y. G., and Witt, D. (2000) *How Long Does it Take English Learners to Attain Proficiency?* Santa Barbara, CA: University of California Linguistic Minority Research Institute. <https://eric.ed.gov/?id=ED443275>
- Han, Y. (2017) 'Mediating and Being Mediated: Learner Beliefs and Learner Engagement with Written Corrective Feedback', *System*, 69: 133–42. <https://doi.org/10.1016/j.system.2017.07.003>
- Han, Y., and Hyland, F. (2015) 'Exploring Learner Engagement with Written Corrective Feedback in a Chinese Tertiary EFL Classroom', *Journal of Second Language Writing*, 30: 31–44. <https://doi.org/10.1016/j.jslw.2015.08.002>
- Kessler, M. (2023) 'Written Corrective Feedback in an Online Community: A Typology of English Language Learners' Requests and Interlocutors' Responses', *Computers and Composition*, 67: 102752. <https://doi.org/10.1016/j.compcom.2023.102752>
- Kline, R. B. (2016) *Principles and Practice of Structural Equation Modeling* (4th ed.). New York: Guilford Press.
- Li, F., and Han, Y. (2024) 'The Influences of International Master's Students' Feedback-Seeking Behavior on Their Feedback Literacy and Feedback Contexts: An Ecological Perspective', *Journal of English for Academic Purposes*, 71: 101424. <https://doi.org/10.1016/j.jeap.2024.101424>
- Li, J., and Zhang, L. (2026) 'Unveiling the Antecedents of Feedback-Seeking Behavior in L2 Writing: The Impact of Future L2 Writing Selves and Emotions', *Assessing Writing*, 67: 101009. <https://doi.org/10.1016/j.asw.2025.101009>
- Li, S. (2010) 'The Effectiveness of Corrective Feedback in SLA: A Meta-Analysis', *Language learning*, 60: 309–65.
- Liu, G. L., and Zhao, X. (2026) 'A Scoping Review of AI-Mediated Informal Language Learning: Mapping Out the Terrain and Identifying Future Directions', *ReCALL*, 38: 111–30. <https://doi.org/10.1017/S0958344025100359>
- Liu, L. A., and Hwang, G. J. (2024) 'Effects of Metalinguistic Corrective Feedback on Novice EFL Students' Digital Game-Based Grammar Learning Performances, Perceptions, and Behavioral Patterns', *British Journal of Educational Technology: Journal of the Council for Educational Technology*, 55: 687–711. <https://doi.org/10.1111/Bjet.13400>
- Long, M. (1996). 'The Role of the Linguistic Environment in Second llanguage Acquisition', in W. Ritchie and T. Bhatia (eds.) *Handbook of research on Second Language Acquisition* (pp. 413–68). New York: Academic Press.
- Long, M. H. (1980) *Input, Interaction, and Second Language Acquisition*.
- Lyster, R., and Ranta, L. (1997) 'Corrective Feedback and Learner Uptake: Negotiation of Form in Communicative Classrooms', *Studies in Second Language Acquisition*, 19: 37–66.
- Mackey, A., Gass, S., and McDonough, K. (2000) 'How do Learners Perceive Interactional Feedback?', *Studies in Second Language Acquisition*, 22: 471–97. <https://doi.org/10.1017/S0272263100004010>
- Mahbodi, M., Papi, M., and Wolff, D. (2025) 'How Learners' Trust in Teachers Shapes Their Feedback-Seeking Behaviors: A Mixed-Methods Study', *Language Teaching Research*, 40: 1–26. <https://doi.org/10.1177/13621688251352268>

- Yao, Y., and Zhu, X. (2024) 'Modeling Chinese High School EFL Writers' Language Mindsets, Goal Orientations, and Feedback-Seeking Behavior: The Interaction Effect Between Language Mindsets and Writing Competence', *Reading and Writing*, 37: 539–61. <https://doi.org/10.1007/S11145-022-10339-6>
- Zhan, J., Yao, Y., and Zhu, X. (2023) 'The Mediating Effect of Feedback-Seeking Behavior on the Relationship Between Ideal L2 Writing Self and Story Continuation Writing Performance: A Multigroup Structural Equation Modeling Approach', *System*, 113: 102976. <https://doi.org/10.1016/j.system.2022.102976>
- Zhang, G. (2023) 'English Writing Mindsets and Feedback-Seeking Behaviors of Chinese College EFL Students', *International Journal of Education and Humanities*, 11: 475–9. <https://doi.org/10.54097/ijeh.v11i3.15152>
- Zhang, X., Winke, P., and Clark, S. (2020) 'Background Characteristics and Oral Proficiency Development Over Time in Lower-Division College Foreign Language Programs', *Language Learning*, 70: 807–47. <https://doi.org/10.1111/lang.12396>
- Zhang, X., and Zhang, R. (2023) 'Feedback, Response, and Learner Development: A Sociocultural Approach to Corrective Feedback in Second Language Writing', *Sage Open*, 13: 21582440231157680. <https://doi.org/10.1177/21582440231157680>
- Zhang, Y. (2025) 'The Impact of Teacher Academic Support and L2 Writing Self on Feedback-Seeking Behavior', *Reading and Writing*, 38: 1197–215. <https://doi.org/10.1007/S11145-024-10557-0>
- Zhang, Y., and Jiang, Z. (2025) 'Modeling the Interplay Between Teacher Support, Anxiety, and Grit in Predicting Feedback-Seeking Behavior in L2 Writing', *Assessing Writing*, 64: 100920. <https://doi.org/10.1016/j.asw.2025.100920>
- Zhou, H., Carless, D., and Nieminen, J. H. (2025) 'Students' Motivations for Feedback Seeking: The Value of Combined Monitoring and Inquiry', *Assessment and Evaluation in Higher Education*, 1–17. <https://doi.org/10.1080/02602938.2025.2596351>