Content list available at http://ijltr.urmia.ac.ir

Iranian Journal of Language Teaching Research

ORIGINAL ARTICLE

Urmia University

Development and Psychometrical Validation of a Computer-Assisted Language Learning Evaluation Scale: An Ecolinguistic Approach

Sara Sadat Hosseini [a], Hassan Soleimani [a, *], Fatemeh Hemmati [a], Jafar Afshinfar [a]

[a] *Payame Noor University, Iran*

ABSTRACT

From the complex dynamic systems theory perspective, evaluating CALL pedagogy involves miscellaneous variables. It seems the ecosystem of the CALL environment necessitates an array of extralinguistic (Ecolinguistically situated) issues to be considered in any appraisal of the CALL milieu. To this end, this paper aims to develop and validate a scale for CALL evaluation called the Ecolinguistics CALL Evaluation Scale (ECES). It enframes the complex processes of language use in CALL and technology mediation as facets of a complex adaptive system (CAS). The psychometric testing of a questionnaire was undertaken with 219 EFL academics to investigate the sociocultural, sociopolitical, and ideological values and norms within the CALL ecosystem, thereby representing the tripartite interaction of human beings, Ecolinguistics, and CALL. The model being probed included 14 components of ECES. Confirmatory Factor Analysis (CFA) was run using LISREL Software (Ver. 8.80) to examine the trait structure of the ECES questionnaire. Results yielded that all items had significant contributions to their constructs, and their respective t-values all indicated that the contribution of items to their constructs was also statistically significant. This scale might equip a more streamlined method that is readily adaptable to a variety of contexts, stakeholders, and criteria.

Keywords: big computer-assisted language learning; complex adaptive system; ecolinguistics; scale development and validation

ARTICLE HISTORY

Received: 7 Jan. 2024 **Revised version received:** 27 July 2024

Accepted: 10 Feb. 2025 **Available online:** 10 Mar. 2025

* Corresponding author: Department of Applied Linguistics and TEFL, Payame Noor University, Tehran, Iran

Email address: h_soleimanis@pnu.ac.ir

doi 10.30466/ijltr.2025.55680.2842

Introduction

Computer-assisted language Learning (CALL) evaluation is prominent since it is deemed essential to substantiate CALL's effectiveness (McMurry et al., 2016). CALL studies computer applications or technologies in second or foreign language teaching and learning (Chapelle, 2001; Levy & Stockwell, 2013; Warschauer, 2004). Since CALL applications are utilized extensively in today's learning contexts, the urgency of verifying and measuring the efficacy of CALL has risen due to this trend. That being the case, latent constructs as manifest measurement scales could be utilized to quantify intangible phenomena like attitudes, actions, and hypothetical situations that we assume to exist based on our theoretical knowledge of the world that cannot be directly assessed (DeVellis, 2017). Various scales have perspicuously established a more rigorous, methodical approach to evaluation that can aid stakeholders in mitigating and handling potential hurdles that may arise throughout the decision-making process (Clark & Watson, 1995; DeVellis, 2017; Nunnally, 1975). Prominent authors in CALL (e.g., Chappelle, 2010; Hubbard & Colpaert, 2019; Kessler, 2018; Levy & Stockwell, 2013; Warschauer, 2011) have proposed a host of scales for evaluating CALL. Multiple "state-of-the-art" CALL evaluation studies have been released within the last few years, showcasing critical survey research and practice in CALL. All else being equal, policies and practices at the national, institutional, and classroom levels have come to the forefront in many settings to address meso and macro issues, raising new questions that demand innovative solutions (Dafouz & Smit, 2020). To meet this need, it is imperative to conduct a comprehensive program evaluation (Liu et al., 2011). Nonetheless, the field is rife with conflicting perspectives and lacks sufficient research on developing an optimal framework in the complex and dynamic CALL contexts, where the domain is rapidly evolving.

As with the broader field of applied linguistics, "CALL can be located at the crossroads of several disciplines" (Levy, 1997, pp. 47-75). In as much as CALL, complexity theory, and applied linguistics as separate tripartite interwoven processes are pretty in their embryonic stage (Schulze, 2017; Soleimani, 2020), the establishment of a commensurate evaluative agenda of the complex adaptive system (CAS) in CALL and Ecolinguistics and its undercurrent dimensions seemed far from expectations. In the wake of the recent eco-paradigm transition in CALL (similar to a shift to CAS), proposed by Colpaert (2013), and the conceptualization of a CALL ecology model (Marek &Wu, 2014), "technology used for CALL is not an end in itself, but a means to an end that is based on fully understanding the educational ecology" (p. 571). Thus, much of the current debate among CALL researchers concerns the establishment of a coherent evaluative agenda in putting current CALL issues into perspective, which begins with an attempt to pinpoint the disciplinary influences on newly emerging areas of study. Therefore, "Ecolinguistics" as a subfield of linguistics, with its dedication to ecological and dialectical epistemologies, acts as one of the core features of the scale domain that still appears to be underappreciated and bears significant theoretical and practical implications for the model (Chen, 2016).

Accordingly, the technology-mediated facets of CAS are adopted as complex processes of language use in CALL and diverse dimensions of the Ecolinguistic approach (Steffensen & Fill, 2014). This paper intends to inquire into the sociocultural and ideological values, norms, and ecosystem of the society where CALL is used, thereby representing the tripartite interaction of human beings, Ecolinguistics and CALL. This paper explicitly aims to propose a reliable and purpose-driven CALL evaluation framework drawn from more systematic procedures and seasoned evaluators. The proposed framework differs from the formal evaluation tasks as it is less constrained and covers an extensive domain encompassing any CALL activity or material. By mapping the frameworks against ecolinguistic components, we aimed to incorporate dimensions compatible with the current CALL perspectives in the EFL context and develop a scale that underpins views about language ecology embracing the four theories of ecological linguistics: symbolic, natural, sociocultural, and cognitive as proposed by Steffensen and Fill (2014). We also intended to incorporate brand-new dimensions such as developing environmental knowledge and

an awareness of the pressing need for handling the issues of concern in a range of contexts, cognitive and affective abilities and dispositions, and behavioural strategies to apply and make constructive decisions in virtual learning contexts. As Warschauer (2002, 2004, 2011) investigates the application of computers beyond the classroom and unpack the "digital divide," CALL researchers felt the need to seek measures in domains of Ecolinguistics, cultural studies, cybercultures, and this, in turn, can strengthen CALL and move it further away from a tendency to paint a somewhat trouble-free and utopian vision of technology in education.

Literature Review

Ecolinguistics and CALL

Taking into account the concepts of the social or socio-cognitive perspective of constructivism (e.g., Bandura, 1986; Piaget, 1971; Vygotsky,1978), the field of Ecolinguistics examines language with its environment and complements the study of ecology and focuses on the interdependence of systems (Derni, 2008). As integrative CALL (Warschauer, 2004), like Ecolinguistics, prioritizes an authentic environment and a socio-cognitive perspective on language, it may be argued that the two fields are complementary. The 'Sole Agent' fallacy in CALL, posited by Bax (2003), challenged the conventional notion that technology is the mere factor in effective technology implementation. Ecolinguistics and the 'Sole Agent' fallacy in CALL focus on all elements and agents interacting in an environment. The study of language ecology has benefited from the advent of new computational measuring tools, which allow for the foresightful prediction of both the direction and pace of linguistic alterations. Measurement of interlingual attraction, construction of community language pressure profiles, and geocoding of language use patterns are quantitative approaches to studying language shifts in an ecological context that may be implemented using computers (Mackey, 2001). CALL also has applications in discourse analysis, critical discourse analysis, and corpus analysis to reveal language ecology.

CALL technology integrates all aspects of Ecolinguistics, including the capacity to construct and recruit non-symbolic structures and human technologies extending sensory, executive, and cognitive systems. In cognitive sciences, distributed cognition sees technology as systemic entities that transmit cognitive processes not just in a given circumstance but also "through time in such a manner that the outcomes of earlier events might alter the nature of subsequent occurrences" (Hollan et al., 2000, p.176).

Scale Development

In tandem with the prominent systematic evaluation frameworks, many measurement scales have also been devised upon which a compendium of interventions and attitudes in various applications have been assessed (Petticrew & Roberts, 2008). Operationally defined, scale development is the process of creating a set of items to measure a latent variable (Salkind, 2007). Measurement scales provide a macroscopic view of the numerical values of phenomena beyond direct measurement. Therefore, scale development as an intricate process necessitates a firm theoretical and methodological rigour that still needs further scrutiny (Clark &Watson, 1995; DeVellis, 2017; Nunnally, 1967; Vandewaetere & Desmet, 2009). Attitudes towards computer-assisted learning (A-CAL), attitudes towards foreign language learning (A-FLL), and, more precisely, attitudes towards computer-assisted language learning (A-CALL) are all illustrative instances of non-observable constructs that have been the crux of research efforts to develop and validate scale items as in the works of Vandewaetere and Desmet (2009). The same was true for the psychometrical validity of the questionnaires and the new scale of willingness to communicate

in a second language (L2 WTC) in digital and non-digital EFL scenarios, further investigated by Lee and Drajati (2020).

The technology-enhanced Language Learning (TELL) framework developed by Levy and Stockwell (2013) also reflected complexity theory by considering CALL effectiveness based on six dimensions: pedagogy, tasks, technology, learner, teacher, and context. This framework acknowledged CALL programs' complex and dynamic nature and the importance of considering multiple perspectives and dimensions. Along the same lines, in TeLL, an update and a principled framework for English for Academic Purposes (EAP) courses, Chau and Lee (2014) reviewed recent TeLL research and its implications for EAP pedagogy, curricula, assessment, and instruction. The study highlighted an increasing focus on vocabulary, grammar, and writing while noting less attention to speaking, listening, and reading. It also identified a shift from a tool-centric view to a technology-pedagogy-human alliance, with emerging trends in multi-purpose, multi-genre, and multi-role/skill designs. Despite these advancements, the authors underscored the need for a holistic framework tailored to EAP-specific TeLL.

Complementing this, in another qualitative research study of Intelligent Technology-Enhanced Language Learning (ITELL), Novawan et al. (2024) explored AI's role in higher education language teaching, highlighting positive impacts like enriched teacher perspectives and efficient assessments while noting challenges such as depersonalization and ethical considerations within which a hybrid model balancing AI-driven personalization with human interaction was recommended.

Alongside, the evaluative criteria for CALL software and materials, developed by Chapelle and Jamieson in 1987 and 1988, emphasized usability, interactivity, and feedback as key criteria for assessing the effectiveness of CALL. These criteria aligned with complexity theory, which suggested that complex systems, such as CALL programs, must be user-friendly and interactive and provide feedback to be effective. In contrast, McMurry's Evaluating framework (2016) as an Integrated Approach to Effectiveness Research took a more holistic approach to CALL evaluation, drawing on four key areas: pedagogy, technology, institution, and user perspectives. This approach reflected the complexity theory concept of emergent properties, which suggests that the effectiveness of CALL programs depends on the interplay of various factors, including the pedagogical approach, technology, the institutional context, and the users. At times, the extant scales in research are practical to use. However, sometimes, they are not contingent upon the criteria that we intend to measure, so the necessity of developing a multidimensional evaluation scale is felt as we try to do in-depth research on the issue as to what is available, how it has been defined and evolved, available definitions, the common themes, and dimensions based on the pre-available scales and questionnaires, the prominent figures, as well as the methodologies that have been employed for this particular research.

Given the sizable amount of literature on scale development in terms of theoretical and methodological aspects, some limitations persist. Therefore, scale validation, which ensures that a scale accurately measures its intended constructs, becomes crucial to guarantee its reliability and applicability. Some studies found themselves ill-equipped in terms of the suboptimal description of the construct domain, measurement model, underutilization of some techniques helping establish construct validity (MacKenzie et al., 2011), unfitting data reduction techniques, small sample sizes, disproportionate psychometric properties, application to only a single form of treatment or manual, the extensive time required to fill out the questionnaire (Hilsenroth et al. 2005), disequilibrium concerning the assessment of construct validity (Smith, 2005), and ultimately in terms of their narrow scope of the manifold interaction of specific variables that influence the evaluation priorities and objectives, e.g., differing stakeholder evaluation needs, organizational, political, sociopolitical, ideological, environmental or resourcing factors. Furthermore, using apposite evaluation techniques and tools, knowing what counts as evidence

and how it is applied, and the roles of practitioners, educators, and researchers in assessing and solving real-world programs are some practical challenges associated with performing an assessment. Aside from all the restrictions addressed above, there still exist some potential ecological variables that form the central construct of this current study but are still uncovered in the research and suggest future research which might yield new insights in hypothesizing and testing potential variables that could be accounted for during item generation of the scale development process.

Apropos to the stance elaborated in the literature, the researchers argue that the integration of exploratory falsificatory goals may attain a more problem-based, transdisciplinary approach, individual group analyses and qualitative and quantitative methods drawn from CDST as one of the main theoretical rationales in its conception and design of this study. To accomplish these objectives, the following research questions were formulated:

RQ1. What are the tentative components of an Ecolinguistic CALL evaluation scale (ECES)?

RQ2. What is/are the underlying factor structure/s of the Ecolinguistic CALL evaluation scale (ECES)?

Method

The newly-constructed ECES presented a validated evaluation tool to measure ideological, sociological, psychological, and biological dimensions with a few subcomponents defined by the researchers. It is a good fit for developing and implementing Ecolinguistics and CALL respective domains as it satisfied the researchers with establishing their criteria for achieving the intended results. In light of the systemic reality of ecology, ecological vision and its substrates were captured in the four parallel spheres of ecological linguistics proposed by Steffensen and Fill (2014). These notions and practices were picked because their significant conceptualization bore a convergent ecological orientation toward creating a complex theoretical underpinning in CALL.

Design of the Study

An exploratory sequential mixed methods design was employed to address the research questions of this study (Creswell & Plano Clark, 2018). Under this technique, we explored a phenomenon, identified themes, designed an instrument, and ultimately tested it. The rationale of the qualitative phase of this study was to determine variables and questions that inform, complement, and build on the quantitative data collection procedure for the subsequent phase. Alternatively, the qualitative data were gathered through the focus group method (FGM). Cognitive interviews were conducted to explore the respondents' perceptions, and a set of baseline tentative evaluation criteria was gleaned from pre-existing literature reviews, relevant book chapters, and journal articles, which were subsequently synthesized to capture the essence of ECES. Subsequent to collecting qualitative data, confirmatory factor analysis and quantitative data collection were conducted to explain the findings. This way, the researchers secured a triangulated data set, which might reliably lead to robust findings.

Contexts and Participants

The sample in this study comprised 219 male and female Iranian EFL learners and academics (M=27.03, SD=6.65) who participated of their own volition in late Fall 2022. Regarding the sampling techniques, scale development bears its specific justification pertaining to non-

probability and non-random sampling, as the maximum amount of variance is better guaranteed. Hence, participants with positive and negative sides of the measured variable were included in the sample. To illustrate, the right sampling method was employed for each stage of scale development. For instance, we drew on purposive sampling for initial item generation and refinement to ensure content validity, ensued by convenience sampling for pilot testing to identify any issues with the items and overall scale structure. Finally, stratified random sampling was employed for large-scale validation to capture variability across subgroups and to develop a well-rounded and generalizable scale.

It stands to be mentioned that 341 EFL Iranian learners and academics filled out and returned the questionnaire; however, 22 filled out and returned forms were discarded because they either opted for the same choice across all items or left several items blank. The sample size of 219 participants was determined through Power Analysis, a statistical technique to ensure the findings were genuine and significant (Larsen-Hall et al., 2016). To compute the optimal sample size for our study, we utilized the R-Package "SEM Power" (Moshagen, 2021), which allowed us to determine the necessary number of participants to achieve reliable results. Subsequently, all 219 respondents who filled out the questionnaire were eligible for inclusion. Since no exclusion criteria were established in advance, we ended up with a diverse pool of respondents. Descriptions of the data are provided in depth in Table 1.

Table 1
Frequencies (N), Percentages (%) of Participants' Characteristics

		Total (N= 219)	
		n	%
Age	15-20	29	13.2
	21-25	86	39.3
	26-35	69	31.5
	36 and above	35	16.0
Gender	Male	97	44.3
	Female	122	55.7
Degree	BA	33	15.1
	MA	168	76.7
	PhD	18	8.2
Place of teaching	Language Schools	180	82.2
	University	39	17.8
Teaching experience	3-6	48	21.9
	7-10	86	39.3
	11-15	61	27.9
	15 and above	24	11.0
Group association	Language teaching communities	181	82.6
	Civic activists	18	8.2
	Animal/environmental rights activists	6	2.7
	Political activists	9	4.2
	Socio-political working group	5	2.3

Component Model of the Ecolinguistic CALL Evaluation Scale (ECES)

Developing the 46-item ECES questionnaire, the researchers were able to measure the following 14 constructs: Time, Change, and Continuity (2 items); People, Places, and Environment (2 items); Ecolinguistic, Ecojustice, and Ecofeminism (5 items); Multimodal Interactive Learning (5 items); Ecological Discourse Analysis (3 items). Technology and Society (5 items), Global Connections (3 items), Individuals, Groups, and Institutions (3 items), Classroom Discipline in

CALL (3 items), Pedagogical Usability of Instructional EFL E-learning Materials (11 items), Digital Skills Categorization (8 items) and Digital Identity (6 items).

Data Collection Procedure and Analysis

Conducting Cognitive Interviews

To explore respondents' perceptions and to enhance the ECES, we conducted cognitive interviews as part of our stage one data collection. These interviews involved a diverse sample of participants, each lasting approximately 25 minutes. The semi-structured format included open-ended questions designed to gather in-depth feedback on key concepts in Ecolinguistics and ecological literacy and their integration into CALL (see the appendix for the list of sample questions).

Using Interview Data to Inform ECES Development and Revision

Insights from the cognitive interviews were crucial in refining the ECES. Therefore, participants' feedback led to specific revisions, ensuring the scale accurately captured the intended constructs. For instance, one of the original items, "Consider the role of digital technology in environmental education," was found to be too broad by participants. They suggested that the item mention specific digital tools to provide a clearer context. Consequently, the item was revised to "Assess the effectiveness of specific digital tools (e.g., virtual reality, interactive E-books, visual metaphors) in promoting environmental literacy)" in promoting environmental literacy, thereby addressing the feedback and enhancing the item's specificity and relevance. We followed the same procedure for the remaining items, ensuring each was revised based on participants' feedback to enhance clarity, relevance, and specificity.

Analyzing Cognitive Interview Data

To analyze the qualitative data on the cognitive interviews, the data were collected by conducting 15-25 face-to-face interviews among participants at two universities in Tehran to examine pre-testing methods. Experts' attitudes and perceptions were commonly sought through discussions to gain a deeper understanding of their perceptions, beliefs, opinions, and nuances to reach unanimity and avoid polarization. To this end, a set of a priori codes was assigned to data segments. Open thematic coding was employed as we collated excerpts to codify and identify the overarching themes (e.g., Technology in Language Learning, Environmental and Ecological Perspectives, Sociopolitical and Cultural Aspects, Pedagogical and Instructional Practices). We used various standards as guides while we thematized the data. Considering this theoretical stance, the data were coded in two distinct ways: Inductive, based on the emerging themes, and deductive, using the study's theoretical underpinnings. For example, when a code from the dataset was relevant to "Environmental and Ecological Perspectives," we mainly focused on the people, places, and environment. We discussed the nexus between language learning technology and environment, ecojustice, Ecolinguistics, ecofeminism, ecological discourse analysis, and the representation of nature in language learning materials. To illustrate the "Social and Cultural Aspects" category, we tried to explore the impact of technology on society and culture, including discussions on digital poverty, access to technology, global connections, cross-cultural flows, values, and the influence of social media platforms on the environmental aspects.

The dimensions and the item numbers for the scale are thoroughly explained (See the appendix for the complete list of items). This coding procedure was done to reach themes across the data.

After reviewing the dataset, the themes were aggregated, condensed, and applied to previous and subsequent datasets. Clusters of interrelated subthemes were identified through axial coding. The generated codes in each cluster were then used to label each cluster. Our coding approach allowed us to analyze and organize data related to our inquiry, with themes informed by the theoretical literature. One or two rounds seemed ideal until saturation or relatively few new insights were revealed (Beatty, 2007; Willis, 2007). On top of that, in-person interviews gave the researcher a better edge as they could observe the respondents' non-verbal communication and cues, thereby creating more encouragement and involvement on the respondents' part. Finally, a total of 62 items were proposed. The phase later continued with content validation and finalization of the evaluation model derived from the participants' and experts' opinions using the focus group discussions (FGDs). The ECES scale items were presented to academic specialists (i.e., Supervisory committee and connoisseurs in the field of environmental science to cross-check the dataset. The goals and objectives of the questionnaire were elucidated, considering their preferences and opinions about the initial set of criteria derived in the item generation step. Further, several consequential and influential field figures from top-tier CALL journals were identified and reached out via email to bestow their scholarly advice on the targeted yardstick dimensions to determine whether items appropriately described the property to be measured and added validity to the content.

Using Interview Data to Inform ECES Development and Revision

Insights from the cognitive interviews were crucial in refining the ECES. Therefore, participants' feedback led to specific revisions, ensuring the scale accurately captured the intended constructs. For instance, one of the original items, "Consider the role of digital technology in environmental education," was found to be too broad by participants. They suggested that the item mention specific digital tools to provide a clearer context. Consequently, the item was revised to "Assess the effectiveness of specific digital tools (e.g., virtual reality, interactive E-books, visual metaphors) in promoting environmental literacy," thereby addressing the feedback and enhancing the item's specificity and relevance. We followed the same procedure for the remaining items, ensuring each was revised based on participants' feedback to enhance clarity, relevance, and specificity.

Once the ECES questionnaire was developed, it was distributed to the masses in late Fall of 2022 for EFA and in late December 2022 for CFA. Utilizing their electronic devices (tablets, smartphones, and laptops), the respondents completed the questionnaire. Each participant was requested to rate their level of agreement with each statement. Responses were tracked on a five-point Likert scale from one (Strongly disagree) over (3) (Neutral) to 5 (Strongly agree). The items were formulated in line with Ecolinguistically relevant constructs at stake. Any set of drawn data was analyzed through a multitude of statistical techniques in four stages. First, the sample size was determined through Power Analysis. The desired power for statistical analysis should be .80 (Larsen-Hall et al. 2016). The R-Package "SEM Power" (Moshagen ,2021) was employed to compute the desired sample size for the present study. The computations were carried out under three conditions, i.e., the ideal power and RMSEA were set to be 0.80 and 0.05, respectively. The results suggested a sample size of 215 to obtain a power of 0.80.

Second, Cronbach's alpha reliability indices were computed for overall ECES and its components. Table 2 below illustrates Cronbach's alpha reliability indices for the ECES and its 14 components.

Table 2
Cronbach's Alpha Reliability Indices

	Cronbach's Alpha	N of Items
Time, Change, and Continuity	.715	2
People, Places, and Environment	.674	2
Ecolinguistic, Ecojustice and Ecofeminism	.825	4
Multimodal Interactive Learning	.794	4
Ecological Discourse Analysis	.784	3
Technology and Society	.841	4
Global Connections	.791	3
Individuals, Groups, and Institutions	.741	3
Culture	.767	3
Classroom Discipline in CALL	.772	3
Pedagogical Usability of Instructional EFL E-learning Materials	.830	4
Digital Skills Categorization	.831	4
Digital Identity	.799	4
Accessibility issues	.779	3
ECES	.862	46

The ECES questionnaire was tested empirically using thirteen separate Confirmatory Factor Analyses (CFA). The first two constructs, "Time, Change, and Continuity" and "People, Place, and Environment," could not be modelled by LISREL, so they were merged into a single construct. The results of the 14 CFA's led to the removal of 16 items out of 62 overall questionnaire components. A confirmatory Factor Analysis (CFA) was run to explore the underlying constructs of the ECES, and the omitted items were 9, 14, 22, 39, 40, 41, 42, 43, 44, 45, 50, 51, 52, 53, 58, and 59. Subsequently, the scale components were reduced from 62 to 46.

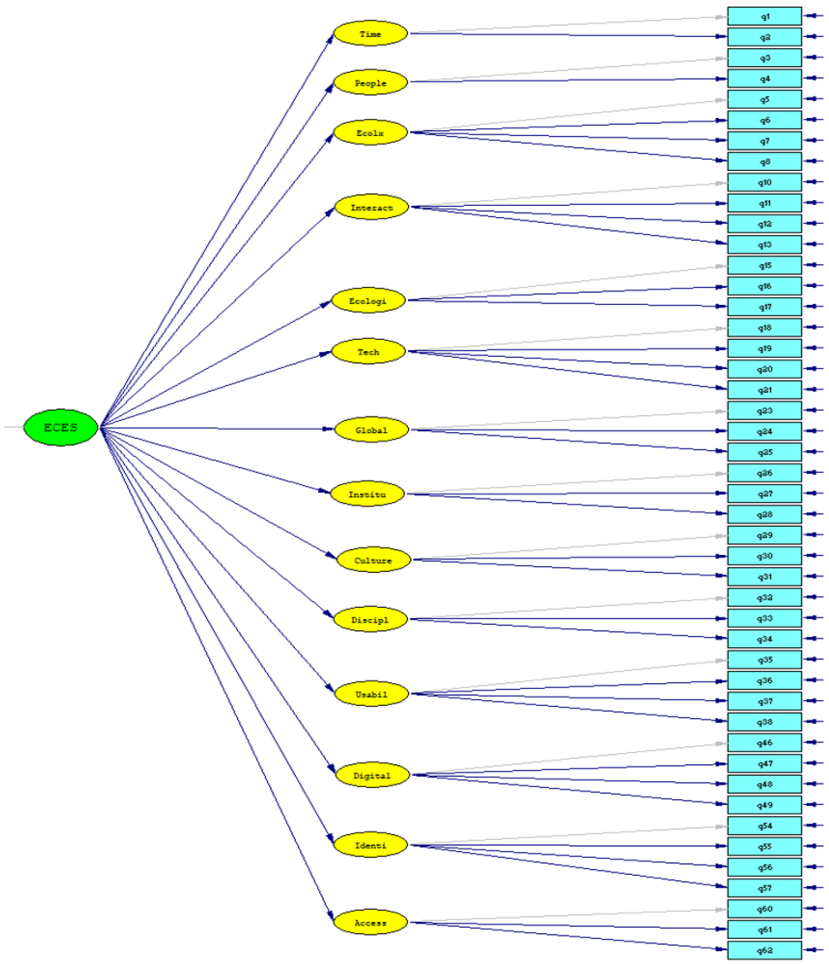

Figure 1. Underlying Factor Structures of ECES

Results

Testing Univariate and Multivariate Normality

Before discussing the results, it should be noted that the data collected in this study met the assumptions of univariate and multivariate normality. As presented in Table 3, skewness and kurtosis values were within the ranges of ±2 (Bachman et al. 2005). Thus, it was concluded that

the assumption of normality was retained. It should be noted that Zhu et al. (2019) suggested the criteria of ±3.

Table 3
Indices of Univariate and Multivariate Normality

Item	Skewness	Kurtosis	Item	Skewness	Kurtosis	Item	Skewness	Kurtosis
1	0.203	-0.856	22	-0.067	-1.251	43	0.040	-1.402
2	0.232	-0.665	23	0.046	-0.876	44	0.063	-1.320
3	0.300	-0.676	24	0.147	-0.846	45	0.051	-1.331
4	0.087	-0.924	25	0.319	-0.595	46	0.159	-0.797
5	0.178	-0.787	26	0.129	-0.830	47	0.198	-0.798
6	0.100	-0.860	27	0.078	-0.735	48	0.148	-0.691
7	0.175	-0.799	28	0.140	-0.898	49	-0.009	-1.229
8	0.261	-0.694	29	0.171	-0.835	50	0.057	-1.296
9	-0.048	-1.331	30	0.202	-0.844	51	-0.060	-1.308
10	0.199	-0.822	31	0.279	-0.582	52	0.179	-1.183
11	0.062	-0.815	32	0.045	-0.917	53	0.163	-0.741
12	0.188	-0.767	33	0.132	-0.792	54	0.073	-0.837
13	0.038	-0.899	34	0.147	-0.954	55	0.095	-0.695
14	-0.103	-1.288	35	0.257	-0.723	56	0.027	-0.890
15	0.094	-0.914	36	0.096	-0.942	57	0.002	-0.722
16	0.083	-0.943	37	0.062	-0.837	58	0.130	-1.211
17	0.151	-0.688	38	0.154	-0.875	59	0.052	-1.346
18	0.124	-0.855	39	0.073	-1.281	60	0.238	-0.854
19	0.045	-0.684	40	0.034	-1.273	61	0.134	-0.748
20	-0.027	-0.867	41	-0.085	-1.261	62	0.149	-0.811
21	-0.027	-0.757	42	-0.057	-1.329	Mardia		-1.376

The assumption of multivariate normality probed through Mardia's index was also retained. As illustrated in Table 3, Mardia's index of -1.37 was lower than ±3 (Bae & Bachman, 2010). Thus, it was concluded that the assumption of multivariate normality was also retained.

Confirmatory Factor Analysis of ECES

As discussed earlier, and through the process of testing components of ECES, the number of items was reduced from 62 to 46. The items that were dropped out due to their low contribution to their constructs were 9, 14, 22, 39 to 45, 46 to 48, 53, 58, and 59. Figure 12 delineates the final CFA model for ECES. All items had significant contributions, i.e.,> = .50, to their constructs.

Moreover, their respective t-values (Figure 13) all showed that the contribution of items to their constructs was also statistically significant, i.e., t-values >=1.96. It should be noted that LISREL assigns a value of one to the first item in each construct to enable the computation to be possible; however, these constant values do not affect the overall results of the models. The readers should refer to the standardized regression weights for the first items.

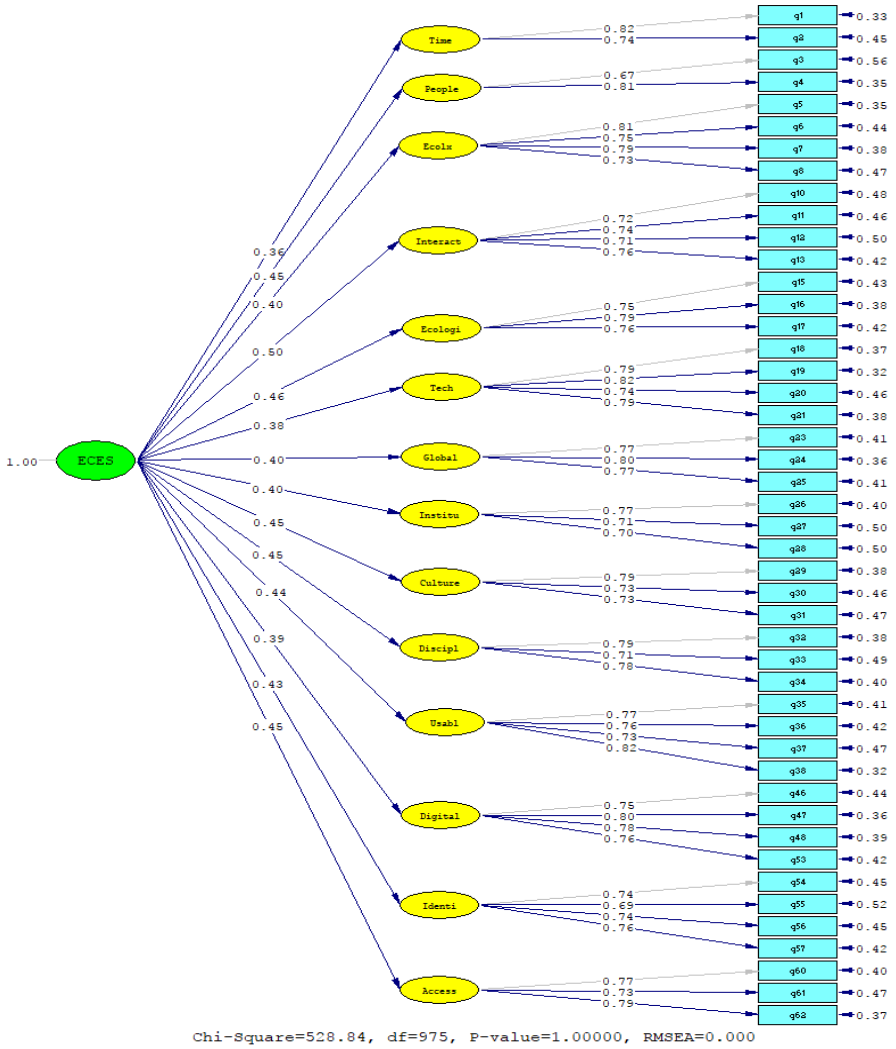

Figure 2. Final Model of Ecolinguistic CALL Evaluation Scale

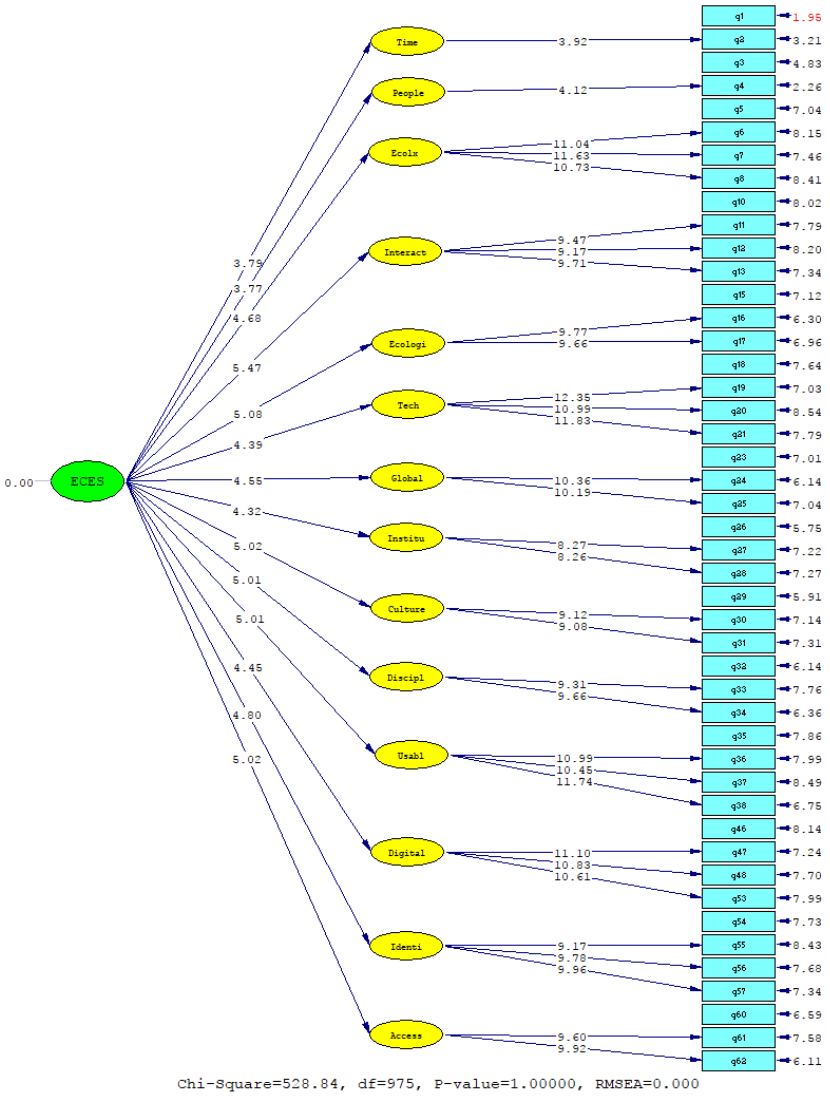

Figure 3. Final Model of Ecolinguistic CALL Evaluation Scale (t-values)

Figure 14 shows the structural models of ECES, standardized regression weights, and t-values, which show the contributions of constructs to ECES. As shown in Table 9, all constructs had moderate to significant contributions to ECES, i.e., >=.30, and significant (t-values > 1.96). These results supported the construct validity of the ECES questionnaire, hence the answer to the research question raised in this study.

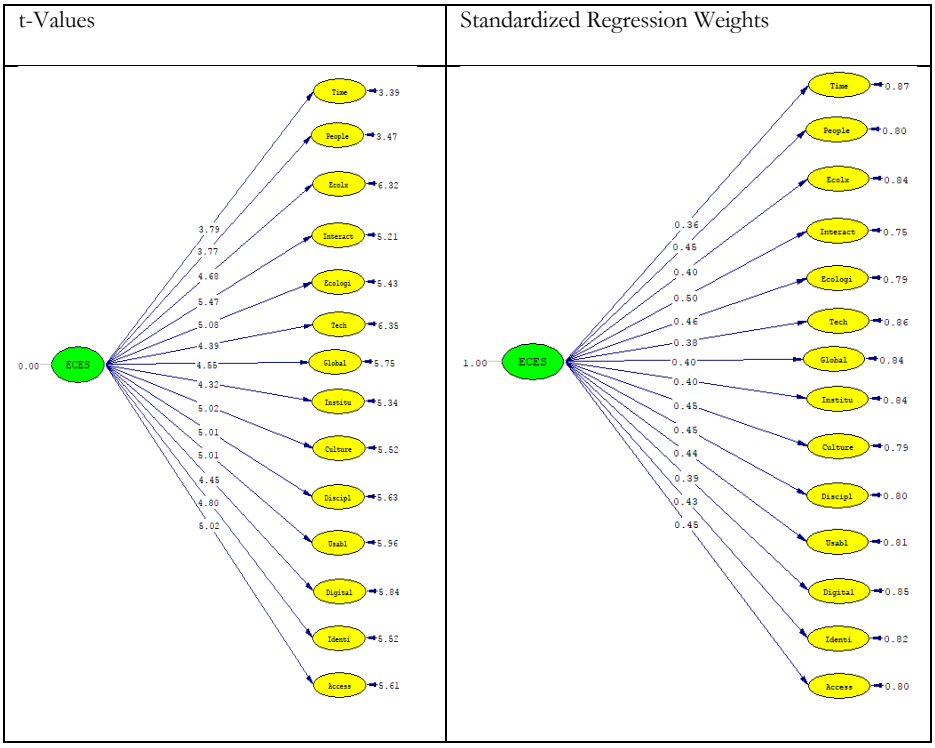

Figure 4. Structural Model of Ecolinguistic CALL Evaluation Scale

Finally, Table 4 shows the fit indices for the ECES model. All of the indices supported the fit of the model. Before discussing the results, it should be noted that model fit indices can be divided into absolute and incremental (relative and comparative) fit indices. As Khine (2013, p 17) stated, "Absolute fit indices measure how well the specified model reproduces the data. They assess how well a researcher's theory fits the sample data". On the other hand, "Incremental (relative, comparative) fit indexes measure the relative improvement in the fit of the researcher's model over that of a baseline model. The baseline model is usually the independence (null) model, which assumes covariances of zero between the endogenous variables" (Kline 2016, p. 266). The results are discussed below.

The chi-square badness of fit of 528.84 at 975 degrees of freedom was non-significant, i.e., p = 1.00. Its ratios over the degree of freedom, i.e., 528.84/975 = .542, were lower than 3. These results supported the fit of the model. The root mean square of error approximation (RMSEA) value of .000 and its 90 % CI [.000, .000] were lower than .05. These results further supported the model.

Table 4
Fit Indices of Final Model of ECES

Fit Indices	Labels	Statistic	D.F.	P-Value	Criterion	Conclusion
Absolute	X^2	528.84	975	1.000	>.05	Good Fit
	X^2 Ratio	.542	---	---	<=3	Good Fit
	SRMR	.032	---	---	<=.10	Good Fit
	RMSEA	.000	---	---	<=.05	Good Fit
	90 % CI for REMSEA	[.000, .000]	---	---	<=.05	Good Fit
	PCLOSE	1.00	---	---	=>.05	Good Fit
	GFI	.90	---	---	=>.90	Good Fit
Incremental	RFI	.91	---	---	=>.90	Good Fit
	NFI	.92	---	---	=>.90	Good Fit
	NNFI	1.00	---	---	=>.90	Good Fit
	IFI	1.00	---	---	=>.90	Good Fit
	RFI	.91	---	---	=>.90	Good Fit
Sampling Adequacy	Critical N	419.25	---	---	=>200	Adequate

The probability of close fit (PCLOSE) of one was higher than .05. The square root mean residual (SRMR) of .032 was lower than .05, and the goodness of fit index (GFI) value of .90 was higher than .90. These results all supported the fit of the model. All of the incremental fit indices were higher than the criterion of .90; i.e., relative fit index (RFI = .91), comparative fit index (CFI = 1.00), normed fit index (NFI = .92), non-normed fit index (NNFI = 1.00), and incremental fit index (IFI = 1.00) all supported the fit of the model; and finally, the critical N value of 419.25 was higher than 200; indicating that the present sample size was adequate for running confirmatory factor analysis.

Discussion

This study took a panorama toward generating a framework for the dynamic integration of a set of items and conducted psychometric testing of different dimensions of the ecolinguistic CALL evaluation scale (ECES). Based on this body of research, we generated a new scale drawn substantively from complex dynamic system theory (Larsen-Freeman,1997) as one of the main theoretical rationales of this study to develop a framework for the dynamic integration of a set of criteria. This supports the view posited by (Liu et al., 2011) that a proper and all-encompassing program assessment is required to help avoid rigid, paradigm-driven research.

As research question one was formulated to identify the tentative components of the ECES, we employed a qualitative data collection analysis involving focus group discussions and cognitive interviews, and a compilation of fundamental assessment criteria was derived from a comprehensive review of existing literature encompassing relevant book chapters and journal articles. Subsequently, a meticulous refinement process was undertaken to distill the quintessential elements of ECES, generating a list of items. The list was satisfactorily verified further ensued by its implementation. To our knowledge, this is an innovative, original scale encompassing 46 items and 14 components in each domain-specific subscale (i.e., Time, Change and Continuity; People, Places; Environment; Ecolinguistic, Ecojustice and Ecofeminism; Multimodal Interactive Learning; Ecological Discourse Analysis; Technology and Society; Global Connections; Individuals, Groups, and Institutions). As a result of the multiple-factor analyses, ECES was condensed into 14 factors and 46 items, and its reliability and validity were satisfied.

Additionally, the overall measurement results indicated that through the process of testing components of ECES, all items had significant contributions to their constructs, i.e.,> = .50, and proved to be a good fit with the data. Moreover, their respective t-values all indicated that the contribution of items to their constructs was also statistically significant, i.e., t-values >=1.96. Therefore, research question two, which aimed to determine the underlying factor structures of the ECES, was satisfactorily answered.

The first factor, "time, change and continuity," included CALL methodologies encompassing a multifaceted approach that could offer comprehensive language teaching and learning methodologies. The second item mentioned the circumvention of the pervasive issue of digital poverty within the nation. As these two items proved a good fit with the data, we posit that individuals who scored highly on factor 1 acknowledged incorporating a proactive approach to bridge the digital divide and providing essential digital infrastructure for all members of society. The second factor comprises the "People, Places, and Environment" and the potential impact of multimedia resources in challenging and mitigating the impact of anthropocentrism. Besides, the mastery of Techno-literacy and top-line digital growth was discussed. These two items proved a good fit with the data; therefore, we postulate that individuals who scored highly on factor 2 and its relevant items explored the potential impact of multimedia resources and digital literacy in addressing ecological challenges, presenting alternative perspectives, sustainable practices, and nurturing a sense of responsibility towards the environment.

The measurement model of "Ecological Discourse Analysis" and its subdimensions contained items about "Visualizing Nature through Technology-Mediated Language Learning," "Government Restrictions and Technological Barriers," and "Negative Impact of Top-Down Policies on the CALL Academic Community." We postulated that Visualizing Nature through Technology-Mediated Language Learning, such as digital story tools and animations, proved essential for cultivating a profound understanding of nature's harmonious facets. These resources seem to enable learners to immerse themselves in a virtual realm, facilitating the exploration of ecological concepts and nurturing a sense of interconnectedness. By seamlessly integrating these tools into language learning, learners might envision the splendour of nature, grasp its intricate dynamics, and develop a profound appreciation for the harmonious coexistence of the environment. The government Restrictions and Technological Barriers, on the other hand, imposed restrictions on internet access coupled with technological barriers, posed formidable challenges for users within the CALL community and technologically adept populations. These restrictions impeded the potential benefits of technology in language learning and hindered access to invaluable resources, collaborative platforms, and innovative learning opportunities. They believed that by curbing internet access and imposing technological barriers, governments inadvertently impede the growth and progress of the CALL community, hampering its ability to leverage technology for language learning purposes fully. The negative impact of Top-Down Policies on the CALL Academic Community detrimentally affected the identity, autonomy, self-efficacy, and agentive initiatives of the CALL academic community. When educational policies are enforced without due consideration for the needs and perspectives of the CALL community, they stifle creativity, curtail academic autonomy, and constrain the community's capacity to drive innovation in language learning. Therefore, the participants postulated that to cultivate a thriving CALL community, it is imperative to foster bottom-up approaches that empower educators and researchers to shape their practices, embrace their academic identities, and embark on proactive initiatives that contribute to advancing the field.

The current study enjoyed several merits compared to other similar evaluation studies. First, the scope of ECES practice was not constrained to the sole assessment criteria of CALL activities and materials, as was the common trend in previous studies (Beatty, 2010; Burston, 2003; Chapelle, 2010; Levy & Stockwell, 2013; Susser, 2001; Tomlinson, 2003; Villada, 2009). The researchers considered the needs of a diverse group of individuals and any relevant interested

parties (i.e., sociopolitical working groups, civic activists, animal/environmental rights activists, and social working groups) they encountered and reflected their merit of worth. Examining and devising items around simply one specific demographic (educators, learners, or programmers) may not address all of the issues inherent in the assessment. Consequently, we scrutinized various stakeholders' motivations for establishing CALL frameworks and their subsequent actions, limiting or assisting its growth. That being the case, we suggested that an enriched, dynamic paradigm be generated to address the diversity and complexity of requirements in a holistic, vital, and unifying manner. Therefore, a more accommodating and comprehensive-oriented framework was proposed to tease the methodological and evaluative yardsticks apart and identify the interactions between stakeholders, evaluands, and other criteria believed to be connected with CALL evaluation purposes.

Further, despite the recognition that Chappelle's Framework (2001) may function compatible in certain circumstances, especially when looking at the language issues involved in CALL, this may not be accommodating when addressing issues other than non-SLA concerns. Not even a modicum part of research existed to take account of the reciprocity of learners and environment to pin down the CALL and Ecolinguistic perspective. Among other things, non-SLA-related aspects, such as infrastructure and other administrative aspects, were entirely dealt with to benefit from an evaluation model that served well-inclusive of sociopolitical, ideological, and ecological considerations.

Along the same lines, the TeLL framework and ECES presented distinctive methodologies for assessing language learning technologies, each with unique advantages and limitations. The TeLL framework was leveraged to deliver tailored, adaptive learning experiences, enhancing engagement and outcomes through immediate feedback and data-driven insights. However, TeLL grappled with technological dependency, privacy concerns, and the substantial need for teacher training, which may inadvertently widen the digital divide. In the same vein, the ITELL study identified challenges such as increased student dependency on technology and potential demotivation due to over-reliance on AI tools.

Conversely, the ECES adopted a theoretically robust ecolinguistic perspective, evaluating CALL within its sociocultural and ecological contexts. This holistic approach offers a nuanced evaluation but is resource-intensive and complex, potentially limiting its applicability. Confirming and disconfirming various aspects, the TeLL framework aligns with ECES's recognition of the importance of dynamic, interactive learning environments, affirming TeLL's emphasis on real-time feedback and adaptability. However, ECES disconfirms TeLL's predominantly technological focus by advocating for a broader, holistic approach that considers ecological and sociocultural factors. While TeLL identifies emerging trends in multi-purpose, multi-genre, and multi-role/skill designs (Chau & Lee, 2014), ECES further emphasizes the necessity of incorporating ecological discourse analysis, ecojustice, and multimodal interactive learning.

Additionally, several noteworthy parallels and distinctions emerged when comparing the quantitative findings of ECES with WTC-R. For one, our study confirmed the assumptions of univariate and multivariate normality, as evidenced by skewness and kurtosis values within ±2 and a Mardia's index of -1.37, which are well within the acceptable range (Bachman et al., 2005; Bae & Bachman, 2010). Additionally, our CFA demonstrated robust construct validity for the ECES model, with all items showing significant contributions (t-values >= 1.96) and fit indices such as RMSEA = 0.000 and SRMR = 0.032, indicating a good model fit. The L2 WTC-R scale also showed a strong model fit, confirming its reliability and validity in measuring willingness to communicate in digital and non-digital contexts. These analyses underscore the complementary strengths of both scales, with the ECES offering a context-sensitive evaluation of CALL environments and the L2 WTC-R scale emphasizing technological integration and real-time feedback. These findings not only validate our study's contributions but also highlight the

importance of integrating sociocultural (Wang & Stockwell 2023), ecological, and technological considerations in CALL and TELL research.

These results underscored the reliability and validity of ECES in evaluating CALL environments. Conversely, the TeLL framework highlighted the effectiveness of integrating technology in EAP courses, supporting technology integration for enhanced language learning. Despite the different focal points of our study, which emphasize ecological linguistics within CALL versus the broader sociocultural integration of technology in EAP by TeLL, ITELL, and L2 WTC, all three showcase statistically significant improvements in language learning environments through the application of technology. This highlights the complementary strengths of our ECES model's precision in measurement and the L2-WTC, TeLL, and ITELL comprehensive approach to embedding technology in academic language learning contexts.

Furthermore, the user-friendliness of the ECES questionnaire could be enhanced by incorporating explicit interpretive directions. Furthermore, the process of interpreting the questionnaire entails carefully analyzing the various issues and recommendations presented in each section and then integrating them into the planning, execution, and assessment of CALL programs. This involves taking into account factors such as the cultural context in which the programs will be implemented, the specific needs of the students, the effective integration of technology, appropriate assessment methods, the professional development of teachers, and the ethical concerns that arise within CALL environments. Furthermore, we will articulate the components and processes that underpin a streamlined method, showcasing its adaptability and versatility. By incorporating these refinements, we can aim to present a more comprehensive and engaging framework that aligns with the needs and expectations of end users.

Overall, the current study results lead us to think that ECES can provide a departure from the prevalent practices in CALL assessment, as it represents a more systematic and methodical avenue for considering the social-ecological entanglement and facets of the scale that could otherwise be disregarded.

Conclusions

This study has given an account of the underlying factor structures of the Ecolinguistic CALL evaluation scale, which had not been exclusively scrutinized in the previous models of the CALL framework, thereby presenting a validated evaluation tool. We generated a framework for the dynamic integration of a 46-item scale with a 14-factor solution to explore research substantive contribution to the field of CALL. An examination of this study's quality revealed several potential areas of improvement. Additionally, we synthesized these insights in the central tenets of the Ecolinguistic CALL evaluation scale. These findings present significant implications for subsequent research. As already indicated, our proposed step-wise approaches and procedures in CALL evaluation could provide assessments that consider many gaps in the current state of CALL evaluation by including various clusters of participants. The successful implementation of the Ecolinguistic subscales and the Ecolinguistically-based task significantly bolstered learners' ecological perspectives while concurrently enhancing their comprehension of intricate ecological concepts. These outcomes substantiated the proposition that integrating Ecolinguistic dimensions into technology-mediated pedagogies holds promising potential for cultivating environmental literacy among English foreign language learners and teachers.

We reviewed conventional evaluation methods with CALL evaluation, reviewing and scrutinizing prevalent CALL frameworks and formal assessment assignments along the way. Our proposed use of standard evaluation procedures in CALL evaluation has the potential to provide

assessments that consider many gaps in the current state of CALL evaluation. The discipline of CALL might benefit from a better understanding of and use of conventional assessment procedures. Building more comprehensive integrated models aligns with Hubbard's (2008) recent call for more indigenous CALL theories. He found very few studies documenting the history of CALL theory and commented on those that did. Hubbard's call for a conceptual framework to encourage theory building in CALL research is something we wholeheartedly support. However, we would want to emphasize the need for a workable methodological framework as a secondary requirement for a theory's creation, verification, or refutation. CALL may glean many valuable insights from formal evaluations and publications. CALL evaluation may be learned from various theories, tenets, and methods. Publications on CALL assessment should be grounded on ideas comparable to those advocated by professional evaluators and represent experienced evaluators' knowledge.

Last but not least, a web of potential limitations must be considered. For one, the items constituting factors of "Time, Change and Continuity" and "People, Place, and Environment," each with two items that LISREL could not model. Hence, they were incorporated into a single construct based on the researcher's judgment. Nonetheless, follow-up studies should be supplemented to augment the explanatory power of each component and the overall explained variance using CFA that encompasses more multifaceted variables. The ECES may now need fixes to a number of previously unresolved issues. Second, a noteworthy constraint within the study was the exclusion of open-ended inquiries in the survey questionnaire, depriving respondents of the opportunity to articulate potential deficiencies or lacunae within the framework. This limitation restricts the qualitative exploration of participants' viewpoints and impedes the identification of areas necessitating refinement or deeper investigation within the said framework. Third, regional factors marginally impacted the findings since about half of the sample was concentrated in a few megacities. While this concentration in urban centers allowed for a convenient sampling approach, it raised questions about the extent to which the findings can be generalized to the broader population of language teachers in Iran, as the sample's geographical distribution may introduce some bias and limit the generalizability of the study's conclusions to other regions or rural settings.

Furthermore, the educational qualifications of the survey subjects merited careful consideration. The data revealed that approximately 85% of participants had master's or doctoral degrees. This observation raised a valid concern regarding the sample's representativeness concerning Iran's broader population of language teachers. Therefore, it is imperative to delve deeper into the educational profile of language teachers in the country to assess whether the high level of educational attainment within the sample aligns with the typical qualifications of the target population. The individuals may have also been exposed due to the predominantly obtained online sample. Further, while it was assumed that participants would respond truthfully and attentively to the questions in this study, there is a possibility that this assumption was impinged on by the participants' lackadaisical involvement since the researcher could not verify their veracity reliably. Other contributing factors may exist in addition to the abovementioned points; in particular, the participants had a huge hassle with connectivity due to the internet crackdown across the country and the political unrest, which may have affected the study's overall results.

Given the exploratory and multifaceted nature of this research, the authors anticipate that the list will be subject to inspection, revision, and refinement as groups debate it, engage with it, and define its value in education; evaluators attempted to be versatile and responsive since a single assessment method will not always yield the optimum results. Quality assessments consider stakeholders' values, aims, questions, and concerns, thereby giving informative data that may be used to make reasonable changes. Consistent with the overall finding of the majority of CALL research, the outcomes of this study indicate that more investigation is warranted (Felix, 2008).

Finally, it is essential to regularly update the items relevant to ECES to measure trending topics that we hope will enhance the substantive contribution of future CALL research.

References

Bachman, L. F. (2005). *Statistical analysis for language assessment*. (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511667350

Bandura, A., & National Inst of Mental Health. (1986). *Social foundations of thought and action*: *A social cognitive theory*. Prentice Hall. https://doi.org/10.4135/9781446221129.n6

Bax, S. (2003). CALL – Past, present, and future. *System, 31*, 13–28. https://doi.org/10.1016/S0346-251X(02)00071-4

Beatty, K. (2010). *Teaching and researching: Computer-Assisted Language Learning* (2nd ed.). Pearson Education Limited. https://doi.org/10.4324/9781315833774

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public opinion quarterly*, *71*(2), 287-311. https://doi.org/10.1093/poq/nfm006

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in Public Health*, *6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Burston, J. (2003). Software selection: A primer on sources and evaluation. *CALICO Journal*, *21*(1), 29-40. https://doi.org/10.1558/cj.v21i1.29-40

Chau, J., & Lee, A. (2014). Technology-enhanced language learning (TeLL): An update and a principled framework for English for Academic Purposes (EAP) courses.*Canadian Journal of Learning and Technology 40*(1). http://dx.doi.org/10.1558/cj.v21i1.29-40

Chapelle, C. (2001*). Computer applications in second language acquisition: Foundations for teaching, testing and research*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524681

Chapelle, C. (2007). Challenges in evaluation of innovation: Observations from technology research. *Innovation in Language Learning and Teaching*, *1*(1), 30-45. http://dx.doi.org/10.2167/illt041.0

Chapelle, C. (2010). The spread of computer-assisted language learning. *Language Teaching*, *43*(1), 66-74. https://doi.org/10.1017/S0261444809005850

Chen, S. (2016). *Language and ecology*: A content analysis of Ecolinguistics as an emerging research field. Ampersand, *3*(1), 108-116. https://doi.org/10.1016/j.amper.2016.06.002

Clark, L. A., & Watson, D. (1995). Constructing validity: basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. https://psycnet.apa.org/doi/10.1037/14805-012

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). SAGE. ISBN:9781483347011, 148334701X https://doi.org/10.1558/cj.v30i0.272-279

Colpaert, J. (2013). Peripatetic considerations on research challenges in CALL. *CALICO Journal*, *30*, 272-279.

Dafouz, E., & Smit, U. (2020). *ROAD-MAPPING English medium education in the internationalized university*. Palgrave Macmillan. http://dx.doi.org/10.14198/raei.2021.34.10

Derni, A. (2008). The ecolinguistic paradigm: An integrationist trend in language study. The *International Journal of Language Society and Culture, 24*, 21-30.

DeVellis, R. F. (2017). *Scale Development: Theory and Applications* (4th ed.). SAGE Publications, Inc. https://doi.org/10.1177/014662169101500413

Salkind, N. J. (2007). Questionnaires. In *Encyclopedia of measurement and statistics* (Vol. 0, pp. 809-812). Sage Publications, Inc., https://doi.org/10.4135/9781412952644

Felix, U. (2008). The unreasonable effectiveness of CALL: What have we learned in two decades of research? *ReCALL*, *20*(2), 141–161. https://doi.org/10.1017/S0958344008000323

Garrett, N. (2009). Computer-assisted language learning trends and issues revisited: Integrating innovation. *The Modern Language Journal*, *93* (focus issue), 719-740. https://doi.org/10.1111/j.1540-4781.2009.00969.x

Hartshorn, K. J., McMurry, B. L., West, R. E., Rich, P., Williams, D. D., & Anderson, (2016). An Evaluation Framework for CALL. Retrieved from https:// dx.doi.org/10.24093/awej/call10.8

Hilsenroth, M. J., Blagys, M. D., Ackerman, S. J., Bonge, D. R., & Blais, M.A(2005). Measuring psychodynamic-interpersonal and cognitive-behavioural techniques: development of the comparative psychotherapy process scale. *Psychotherapy: Theory, Research, Practice, Training*, *42*(3), 340–356. http://dx.doi.org/10.10p. 37/0033-3204.42.3.340

Hollan, J., Hutchins, E., Kirsh, D., 2000. Distributed cognition: toward a new foundation for human–computer interaction. *ACM Transactions on Human-Computer Interaction 7* (2), 174–196. https://doi.org/10.1145/353485.353487

Hubbard, P. (2008). Twenty-five years of theory in the CALICO Journal. *Calico Journal*, *25*(3), 387-399. https://doi.org/10.1558/cj.v25i3.387-399

Hubbard, P., & Colpaert, J. (2019). Toward transdisciplinarity in computer-assisted language learning. *calico journal*, *36*(2), 81-99. https://doi.org/10.1558/cj.37499

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field methods*, *18*(1), 3-20. https://doi.org/10.1177/1525822X05282260

Kessler, G. (2018). Technology and the future of language teaching. *Foreign language annals*, *51*(1), 205-218. https://doi: 10.1111/flan.12318

Kline, R. B. (2016). *Principles and practice of structural equation modelling*. 4th ed. Guilford Press. https://doi.org/10.4236/health.2013.511254

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, *18*(2), 141–165. https://doi: 10.1093/applin/18.2.141

Larsen-Hall, J. (2016). *A guide to doing statistics in second Language learning research Using SPSS*. (2nd ed.). Routledge. https://doi.org/10.4324/9781315775661

Levy, M. (1997). *Computer-assisted language learning*: *Context and conceptualization*. Oxford University Press.
https://doi.org/10.1017/S0958344000003803.ISBN:9780198236313,019823631X

Levy, M., & Stockwell, G. (2013). *CALL dimensions: Options and issues in computer-assisted language learning*. Routledge. https://doi.org/10.4324/9780203708200

Lee, J. S., & Drajati, N. A. (2020). Willingness to communicate in digital and non-digital EFL contexts: Scale development and psychometric testing. *Computer Assisted Language Learning, 33*(7), 688-707. https://doi.org/10.1080/09588221.2019.1588330

Liu, G. Z., Liu, Z. H., & Hwang, G. J. (2011). Developing multidimensional evaluation criteria for English learning websites with university students and professors. *Computers & Education, 56*(1), 65-79. https://doi.org/10.1016/j.compedu.2010.08.019

Lobo, R., Petrich, M., & Burns, S. K. (2014). Supporting health promotion practitioners to undertake evaluation for program development. *BMC Public Health*, *14*, 1-8. http://www.biomedcentral.com/1471-2458/15/1315

Marek, M.W., &Wu,W.-C. V. (2014). Environmental factors affecting computer-assisted language learning success: A complex dynamic systems conceptual model. *Computer Assisted Language Learning*, *27*(6), 1–19. http://dx.doi.org/10.1080/09588221.2013.776969

Mackey, W. F. (2001). The ecology of language shift. In A. Fill & P. Mühlhäusler (Eds.), *The ecolinguistics reader: Language, ecology, and environment* (pp. 67-74). Continuum. ISBN:9780826481733, 0826481736

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: integrating new and existing techniques. *MIS Quarterly, 35*(2), 293–334. https://doi.org/10.2307/23044045

McKeown-Ice, R., & Dendinger, R. (2000). Socio-political-cultural foundations of environmental education. *The Journal of Environmental Education, 31*(4), 37-45. https://doi.org/10.1080/00958960009598650

McMurry, B. L., Williams, D. D., Rich, P. J., & Hartshorn, K. J. (2016). An Evaluation Framework for CALL. *TESL-EJ, 20*(2), n2.

Moshagen M (2021). semPower: Power Analyses for SEM_. R package version 1.2.0, https://doi.org/10.1080/10705511.2014.950896

Novawan, A., Walker, S. A., & Ikeda, O. (2024). The new face of technology-enhanced language learning (TELL) with artificial intelligence (AI): Teacher perspectives, practices, and challenges. *Journal of English in Academic and Professional Communication*, *10*(1), 1-18. https://doi.org/10.25047/jeapco.v10i1.4565

Nunnally, J. C. (1975). Psychometric theory. 25 Years Ago, and Now. *Educational Researcher*, *4*(10), 7-21 https://doi.org/10.3102/0013189X004010007

Petticrew, M., & Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. John Wiley & Sons. https://doi:10.1002/9780470754887

Piaget, J. (1971). *The theory of stages in cognitive development*. McKay. ISBN: 9780679301639,0679301631

Smith, G. T. (2005). On construct validity: issues of method and measurement. *Psychological assessment*, *17*(4), 396. https://doi: 10.1037/1040-3590.17.4.396

Schulze, M. (2017). Complexity approaches to computer-assisted language learning. In S. L. Thorne & P. A. Schneider (Eds.), *Encyclopedia of language and education* (pp. 1-12). Springer.

Soleimani, H. (2020). *Computer-assisted language learning: Theory and practice*. Payame Noor University Press.

Steffensen, S. V., & Fill, A. (2014). Ecolinguistics: the state of the art and future horizons. *Language sciences*, *41*, 6-25. http://dx.doi.org/10.1016/j.langsci.2013.08.003

Susser, B. (2001). A defense of checklists for courseware evaluation. *ReCALL*, *13*(2), 261-276. https://doi.org/10.1017/S0958344001001021a

Tomlinson, B. (Ed.). (2003). *Developing materials for language teaching*. A&C Black. ISBN:9780826459176, 082645917X

Vandewaetere, M., & Desmet, P. (2009). Introducing psychometrical validation of questionnaires in CALL research: The case of measuring attitude towards CALL. *Computer Assisted Language Learning*, *22*(4), 349-380. https://doi.org/10.1080/09588220903186547

Villada, E. G. (2009). CALL evaluation for early foreign language learning: A review of the literature and a framework for evaluation. *CALICO Journal*, *26*(2), 363-389. https://doi:10.1558/cj.v26i2.363-389

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological Processes*. Harvard University Press. ISBN: 9780674576292

Wang, Y., & Stockwell, G. (2023). Social Justice and Technology in Second Language Education. *Iranian Journal of Language Teaching Research*, *11*(3 (Special Issue)), 1-18. doi: 10.30466/ijltr.2023.121403

Warschauer, M., & Kern, R. (2000). *Network-based language teaching concepts and practice*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524735

Warschauer, M. (2002). *Reconceptualizing the digital divide*. First Monday. https://doi.org/10.5210/fm.v7i7.967

Warschauer, M. (2004). *Technology and social inclusion: Rethinking the digital divide*. MIT Press. http://dx.doi.org/10.7551/mitpress/6699.001.0001

Warschauer, M. (2004). Technological change and the future of CALL. In S. Fotos & C. Brown (Eds.), *New perspectives on CALL for second and foreign language classrooms* (pp. 15–25). Lawrence Erlbaum Associates. https://doi.org/10.4324/9781410610775

Warschauer, M. (2011). *Learning in the cloud*. Teachers College Press. ISBN: 9780807752494

Zhu, X., Raquel, M., & Aryadoust, V. (2019). Structural equation modeling to predict performance in English proficiency tests. In *Quantitative Data Analysis for Language Assessment Volume II* (pp. 101-126). Routledge.ISBN:9781315187808

Hosseini, S.S., is a Ph.D. holder in TEFL from Payame Noor University, Tehran, Iran. She is a university instructor and has taught EFL in reputable language institutes. She completed her M.A. in TEFL at Kharazmi University. Intrigued by the nexus of technology, politics, sociology, and anthropology, she explores their interplay with Ecolinguistics and pedagogy in the CALL milieu. She has reviewed articles addressing Ecolinguistic and CALL issues for the Journal of *Computer-Assisted Learning*. Driven by a strong commitment to environmental and animal rights, she is dedicated to integrating Ecolinguistics values into her research in the field of education, particularly within TEFL, to promote sustainability.

Soleimani, H., is the Associate Professor of Applied Linguistics at Payame Noor University, Tehran, Iran. He has authored several books, including *Computer Assisted Language Learning: Theory and Practice* (PNU University Press) in 2021, *Non-parametric Statistics for Applied Linguistics Research* in 2010, and a chapter in *Issues in Materials Development* by Sense Publishers, the Netherlands, in 2016. He is the author of several articles. His special interests are SLA issues in general and computer-assisted language learning in particular. He teaches CALL, Materials Development, and First Language Acquisition at the PNU Ph.D. post-graduate school and Tehran University.

Hemmati, F., is the Associate professor of TEFL at Payame Noor University, Tehran, Iran. She has been teaching TEFL courses for more than 30 years and has authored two books focusing on EFL Writing. She has published more than 60 articles, and her main interests are classroom discourse analysis, writing skills, and teacher education.

Afshinfar, J., is an Assistant professor of TEFL at the Department of English, Faculty of Literature and Foreign Languages, Payame Noor University, Tehran, Iran. He has been lecturing at undergraduate and graduate levels and researching for more than 20 years. His research interests are primarily in the areas of first and second language acquisition, teaching methodology, materials development, and teacher education.