International Journal of TESOL Studies (2026)
260701, 1-20 https://doi.org/10.58304/ijts.260701

Trends in Research Methods

The Ergodicity Issue in Studies in Second Language Learning and Teaching: The Need for Intensive Longitudinal Data Collection and Analysis

Majid Elahi Shirvan
University of Bojnord, Bojnord, Iran

Received: 1 March 2026 / Received in revised form: 22 May 2026 / Accepted: 24 June 2026 / Available online: 10 July 2026

Abstract
This article addresses the ergodicity problem in Second Language Learning and Teaching (SLLT), emphasizing the need for methodological advancements that account for the dynamic, non-linear nature of language development. Traditional research methods in SLLT, often based on aggregated group data, assume ergodicity—a concept that equates group-level averages with individual-level processes, a premise that fails to capture the complexity of individual learners' trajectories. The article advocates for a shift toward intensive longitudinal data (ILD) collection and Dynamic Structural Equation Modeling (DSEM) to model the inherent non-ergodic nature of language acquisition. These methods facilitate a bottom-up generalization approach, where individual trajectories are examined for recurring dynamic patterns before testing these patterns across multiple cases to uncover shared mechanisms. By focusing on intra-individual dynamics, ILD and DSEM provide a nuanced understanding of language learning processes, revealing how emotional and cognitive factors fluctuate within individuals over time. This bottom-up approach contrasts with traditional top-down methodologies and offers a more accurate and personalized representation of language learning, enabling both the identification of common trends across individuals and the retention of individual variability. Despite challenges like sample size requirements and data complexity, the integration of ILD and DSEM is positioned as a transformative methodology for SLLT research, offering practical insights for future research.

Keywords
Ergodicity, second language learning and teaching, intensive longitudinal data, Dynamic structural equation modeling, bottom-up generalization

1 Introduction

Second language learning and teaching (SLLT) has long been conceptualized as a multifaceted and dynamic process, underpinned by cognitive, emotional, social, and contextual factors (Liao et al., 2025; Teng & Yip, 2025; Wen & Fang, 2026). However, traditional research methodologies in SLLT have

Email: elahishmajid@gmail.com; m.elahi@ub.ac.ir

primarily relied on cross-sectional or even longitudinal designs that aggregate data across groups to produce generalizable findings (Larsen-Freeman & Cameron, 2008). These conventional approaches assume ergodicity, a statistical principle which postulates that ensemble averages (group-level statistics) can be equated with time averages (individual-level processes), contingent on the conditions of homogeneity and stationarity (Gates et al., 2023; Molenaar, 2004). To illustrate, consider a study that measures the average foreign language anxiety (FLA) of 100 learners at a single point in time and concludes that learners generally experience moderate anxiety. If ergodicity holds, this group average should accurately represent how any given individual's anxiety fluctuates over time. In reality, however, one learner may experience consistently low anxiety that spikes sharply during oral tasks, another may show a steady decline in anxiety across a semester, and a third may oscillate unpredictably depending on classroom conditions. None of these individual trajectories is adequately captured by the group mean, illustrating why the ergodic assumption is problematic in SLLT contexts.

While this assumption serves as a foundation for many studies, it fails to account for the inherent non-linearity and intra-individual variability present in the language acquisition process. In SLLT, where development is shaped by idiosyncratic learning trajectories and the dynamic interactions between internal and external factors, ergodicity proves to be a limiting assumption. When ergodicity is assumed without verification, group-level findings may not accurately reflect individual trajectories, leading to what Molenaar (2004) terms the "ergodic fallacy." This misapplication can result in misguided interpretations and a failure to address intra-individual variability of learners and educators.

The ergodicity issue has gained significant attention in SLLT, particularly through the lens of Complex Dynamic Systems Theory (CDST), a metatheory that reconceptualizes language development as an emergent, person-specific and self-organizing process arising from the interactions of interconnected subsystems (Larsen-Freeman, 1997; de Bot et al., 2007). CDST has been instrumental in highlighting the complexity of second language development, emphasizing the need for repeated measurements and the modeling of intra-individual variability over time (van Dijk et al., 2025; Lowie et al., 2017). This approach allows for a more accurate depiction of the individual learning process, revealing how small changes can lead to significant outcomes and how language learning trajectories vary widely across individuals (Elahi Shirvan et al., 2020; MacIntyre & McGillivray, 2023; Verspoor & de Bot, 2021). Central to CDST is the idea that variability is not an anomaly but a driver of change, with fluctuations in individual learning trajectories representing the system's inherent potential for adaptation and evolution (Lowie & Verspoor, 2019). By shifting away from the assumption of ergodicity, researchers can better understand the dynamic processes of language acquisition and teaching, which are inherently non-linear and influenced by individual learner trajectories.

However, critics, such as Pallotti (2022), have argued that CDST's focus on descriptive case studies limits its ability to produce generalizable knowledge, a cornerstone of scientific inquiry. Nevertheless, proponents counter that CDST redefines generalization as a bottom-up process, starting with intensive examinations of individual trajectories to identify recurring dynamic patterns, which are then tested across multiple cases to uncover shared mechanisms (van Dijk et al., 2025). This approach contrasts sharply with traditional top-down methods, which derive hypotheses from general theories and assume uniform application across populations (Larsen-Freeman & Cameron, 2008).

This article critically addresses the ergodicity issue in SLLT and advocates for a shift toward the use of intensive longitudinal data (ILD) collection and analysis to overcome the limitations of traditional research approaches. ILD involves collecting frequent, time-dense measurements from individual participants over extended periods, which allows for modeling intra-individual variability and temporal dependencies in real-time (Hamaker et al., 2023). Data collection methods such as the Experience Sampling Method (ESM) and idiographic approaches, coupled with Dynamic Structural Equation Modeling (DSEM) as a data analysis technique, are highlighted as effective tools for capturing the non-ergodic nature of language learning processes while enabling bottom-up generalization (MacIntyre &

Legatto, 2011; Hamaker et al., 2018). These methodologies provide insights into individualized patterns of language development, emotional states (e.g., anxiety, enjoyment), and teaching efficacy, offering both ecologically valid and practically applicable findings (Elahi Shirvan & Taherian, 2025).

The article is structured to provide a comprehensive exploration of the ergodicity issue and its resolution through CDST-aligned methodologies. Following this introduction, the literature review synthesizes the foundations of CDST and the ergodicity problem in SLLT, incorporating recent empirical work. Subsequent sections detail the rationale and features of ILD, data collection methods (ESM and idiodynamic approaches), and the application of DSEM for analysis. Empirical examples illustrate the practical implementation of these methods, followed by discussions of affordances, challenges, implications, and future directions for SLLT research. The article concludes with a call for embracing non-ergodic perspectives to advance evidence-based, equitable practices in language education.

2 Literature Review

2.1 Complex Dynamic Systems Theory in SLLT

CDST emerged in the field of SLLT during the 1990s, marking a paradigm shift from traditional linear and stage-based models of language development (Larsen-Freeman, 1997). Building on earlier applications in developmental psychology, motor coordination, and first language acquisition, CDST was introduced to SLA to address the limitations of static, outcome-focused approaches (Thelen & Smith, 1994; Larsen-Freeman, 1997). Pioneering works by Larsen-Freeman (1997), Herdina and Jessner (2002), and de Bot et al. (2007) established CDST as a metatheory that conceptualizes language acquisition as a dynamic, non-linear process emerging from the interactions of interconnected subsystems, such as cognition, emotions, motivation, and environmental factors.

The introduction of CDST represented a significant departure from traditional models, which often relied on cross-sectional data and assumed linear progress. By the early 2000s, CDST had gained traction, influencing a wide range of SLA topics, including vocabulary acquisition, grammar development, and sociolinguistic competence (Verspoor et al., 2011). Recent reviews, such as those by Hiver and Al-Hoorie (2020) and van Dijk et al. (2025), underscore CDST's growing impact, with over 25 years of research demonstrating its applicability to diverse domains, including multilingualism, interactional discourse, and emotional dynamics in language learning.

A key principle of CDST is the self-organizing nature of language learning, which challenges the static, stage-based models that dominated SLLT research before. Language development is viewed as dynamic, constantly evolving through interactions with the learner's environment. Attractor states represent stable developmental points, but these states are fluid and can shift due to changing internal and external factors (de Bot et al., 2007). This principle underscores the importance of variability—both inter-individual (across learners) and intra-individual (within a learner over time)—which drives development rather than being dismissed as random noise (Lowie & Verspoor, 2019).

The application of CDST has had a profound impact on various areas of SLLT, including vocabulary acquisition, multilingualism, and the role of emotions and motivation in language development. Herdina and Jessner (2002) applied CDST in multilingual contexts, highlighting the interdependencies between languages and illustrating how the development of one language can influence others. More recently, Kruk et al. (2023) examined how positive and negative emotions interact in a dynamic model of language development, demonstrating the temporal fluctuations in learners' emotional states and their impact on language acquisition. In addition, MacIntyre and McGillivray (2023) delved into the complex role of anxiety in second language learning, illustrating how anxiety fluctuates in response to various factors and how these emotional dynamics shape learning outcomes. These studies collectively underscore the dynamic and non-linear nature of language learning, emphasizing the importance of considering both internal factors, such as emotions and motivation, and external factors in understanding the processes of second language acquisition. Despite its contributions, CDST research in SLLT is not without ongoing debates and unresolved issues. A recurring criticism concerns the lack of methodological standardization: while CDST calls for dynamic, idiographic approaches, there is no consensus on which methods best operationalize its core constructs (Hiver & Al-Hoorie, 2020). Critics such as Pallotti (2022) have also questioned whether CDST qualifies as a scientific theory in the formal sense, arguing that its explanatory power is largely post-hoc and its predictions are difficult to falsify. Furthermore, the field has yet to establish clear criteria for what counts as evidence of a complex dynamic system in language data, leading to inconsistent interpretations across studies. These unresolved tensions suggest that while CDST has productively reoriented the field toward process and variability, it requires more rigorous methodological grounding to fulfill its theoretical promise.

2.2 The ergodicity problem in CDST research in SLLT

The concept of ergodicity, borrowed from statistical mechanics, assumes that group-level statistics (ensemble averages) are equivalent to individual-level processes (time averages) when a system meets two strict conditions: *homogeneity*, where the same statistical model applies to all individuals, and *stationarity*, where statistical properties (e.g., means, variances) remain constant over time (Gates et al., 2023; Molenaar, 2004). In SLLT, these conditions are rarely met due to the dynamic, non-linear, and context-dependent nature of language learning and teaching (Lowie & Verspoor, 2019; Taherian et al., 2023). As a result, SLLT processes are non-ergodic, meaning that group averages cannot be reliably applied to individual learners or teachers.

Interestingly, the historical roots of this shift can be traced to Classical Test Theory (CTT), where the concept of a "true score" was originally defined as a distribution function derived from repeated measurements of a single individual over time (Lord & Novick, 1968). Practical constraints, such as the difficulty of collecting dense data, led to a shift toward group-based measurements, often involving single or few assessments across many individuals (Molenaar, 2004). However, CDST revives this idiographic focus, aligning with calls for process-oriented research that prioritizes individual trajectories over aggregated outcomes (van Dijk et al., 2025).

Empirical studies in SLLT have shown that intra-individual variability is often the primary driver of language development, rather than group averages. For example, Lowie and Verspoor (2019) found that variability in learners' language measures, such as lexical diversity and syntactic complexity, better predicted language proficiency than group averages, revealing the limitations of the ergodicity assumption in explaining individual development. In line with this, Elahi Shirvan et al., (2020) investigated the dynamics of foreign language enjoyment (FLE) using Ecological Momentary Assessment (EMA), showing that FLE fluctuates dynamically within individual learners, influenced by both internal emotional states and external factors. MacIntyre and McGillivray (2023) offer further evidence of the non-ergodic nature of language learning through their exploration of foreign language anxiety (FLA). Their analysis highlights how anxiety varies significantly at the individual level, with group-level averages failing to capture the nuanced and dynamic fluctuations that occur within individual learners. Elahi Shirvan et al. (2025) also examined the dynamics of anxiety from a person-specific perspective, using dynamic P-technique factor analysis to reveal how anxiety patterns vary across individuals over time, further highlighting the limitations of aggregate data in capturing individual variability. It is important to note, however, that the critique of ergodicity does not imply that group-level research is inherently invalid. Group-based analyses remain appropriate and informative when the research question is explicitly focused on population-level trends, policy-relevant averages, or comparisons between well-defined groups with reasonably homogeneous characteristics (e.g., comparing proficiency outcomes across two instructional conditions). The concern arises when group findings are uncritically extrapolated to individual processes, or when the research question is fundamentally about intra-individual dynamics.

Researchers should therefore treat the ergodicity issue as a call for epistemological clarity — aligning their method with their research question — rather than as a wholesale rejection of group designs.

However, critics of CDST, such as Pallotti (2022), argue that its focus on descriptive case studies limits its ability to produce generalizable knowledge, a cornerstone of scientific inquiry. Pallotti contends that CDST's emphasis on individual variability may hinder the development of broadly applicable theories, as case studies often lack the statistical rigor required for hypothesis testing. This critique highlights a tension between idiographic and nomothetic approaches in SLLT, with traditional research favoring the latter to establish universal laws. However, addressing the non-ergodicity issue does not necessarily limit generalizability; rather, it redefines how generalization should be approached in SLLT research (van Dijk et al., 2025). Through the bottom-up process of generalization, CDST emphasizes starting from the individual's trajectory and then expanding to broader patterns that emerge across cases, allowing for a more nuanced and flexible understanding of language learning processes. The next section will explore how generalizability is reconceptualized within CDST, highlighting how individual trajectories, when analyzed intensively and across multiple cases, can lead to broader insights without disregarding individual variability.

2.3 Generalizability in CDST research

CDST proponents argue that generalization is not abandoned but redefined as a bottom-up process (van Dijk et al., 2025; Elahi Shirvan & Taherian, 2025). Rather than imposing general theories on aggregated data, CDST starts with examinations of individual trajectories, identifying dynamic patterns (e.g., variability, attractor states) that form the basis for hypotheses. These hypotheses are then tested across multiple cases to uncover shared mechanisms, enabling generalization without assuming ergodicity (Hiver & Al-Hoorie, 2020). This bottom-up approach contrasts with traditional top-down methods, which derive hypotheses from general theories and test them on group data, often ignoring individual differences (Larsen-Freeman & Cameron, 2008). The CDST approach acknowledges the non-ergodic nature of SLLT, prioritizing intra-individual variability while seeking generalizable knowledge through iterative case comparisons (van Dijk et al., 2025). It is important to clarify, however, what constitutes sufficient evidence for bottom-up generalization in practice. There is no universally agreed threshold, but scholars such as Hiver and Al-Hoorie (2020) suggest that recurring dynamic patterns observed consistently across multiple intensively studied individual cases may provide an initial basis for proposing a shared mechanism. The criteria for comparing patterns across individuals should include the consistency of the pattern's direction (e.g., whether anxiety systematically precedes a drop in performance), the magnitude of temporal effects (e.g., autoregressive and cross-lagged coefficients of similar size), and convergence across methodologically diverse sources of evidence. As more cases are added and patterns are replicated, the confidence in broader claims increases, while individual variability is preserved as analytically meaningful rather than treated as noise.

Given the challenges of capturing individual variability and temporal dynamics in traditional research designs and the ergodicity and generalizability issues, how can we move beyond the limitations of aggregated data to better understand the complexities of SLLT? To address the limitations, scholars have increasingly turned to ILD methods. These methods involve frequent, densely spaced measurements over time, providing the granularity necessary to model intra-individual variability and temporal dependencies. The next section explores the rationale for adopting ILD approaches to address these challenges.

2.4 Intensive longitudinal data (ILD)

The rationale for utilizing ILD in SLLT stems from its ability to capture non-linear dynamics and temporal dependencies inherent in the language learning process. Traditional cross-sectional designs and sparse longitudinal studies often fail to account for the dynamic, non-linear changes that occur over time, limiting their ability to reflect the complexity of individual trajectories. By focusing on intra-individual variability, ILD enables researchers to track fluctuations in language performance, emotional states, and motivation in real time, providing a more accurate representation of the non-ergodic nature of SLLT. Additionally, ILD enhances ecological validity by collecting data in naturalistic settings, thus allowing researchers to examine language development as it occurs in real-world contexts (Gates et al., 2023).

ILD methods support advanced modeling techniques such as DSEM and ESM, which allow for the modeling of both temporal dependencies and individual patterns. These methods provide powerful tools for analyzing dynamic processes in SLLT (Hamaker et al., 2018). With frequent, densely spaced measurements collected over extended periods, ILD offers real-time data collection through modern technologies such as smartphones and wearable devices. This facilitates the capture of moment-to-moment fluctuations in affective states, such as willingness to communicate (WTC), FLA, and FLE, providing insights that traditional static data cannot offer (MacIntyre & Gregersen, 2022).

By prioritizing intra-individual variability, ILD and advanced modeling techniques provide a nuanced understanding of language development. This approach challenges traditional assumptions and facilitates bottom-up generalization, where patterns in individual trajectories are examined across cases to identify common mechanisms (van Dijk et al., 2025; Elahi Shirvan & Taherian, 2025). Furthermore, ILD methods bridge the gap between research and practice, offering actionable insights that inform pedagogical strategies, teacher training, and curriculum design—especially in diverse educational contexts where learners' experiences are highly individualized.

Given the potential of ILD to capture the complexity of individual learning experiences, how can these methods be integrated into SLLT research to enhance both theoretical understanding and practical applications in real-world educational settings? To integrate ILD effectively into SLLT research, it is essential to adopt data collection and data analysis methodologies that prioritize the capture of individual trajectories and temporal dynamics. The next two sections explore theses methodologies, discussing how these approaches contribute to a more nuanced and comprehensive understanding of language acquisition and teaching.

3 Intensive Longitudinal Data Collection Approaches

To address these limitations and the ergodicity issue, ILD collection methods, such as the ESM, and idiodynamic approaches, and integrated methodologies, have emerged as critical tools in SLLT research. These methods allow for the frequent, real-time collection of data in naturalistic settings, aligning with CDST's emphasis on non-linearity, variability, and ecological validity (Hamaker et al., 2023; MacIntyre & Gregersen, 2022). This section explores these data collection approaches, detailing their theoretical foundations, applications, advantages, and challenges.

3.1 Experience sampling method (ESM)

ESM is a technique designed to capture real-time experiences in naturalistic settings, minimizing recall bias and enhancing ecological validity (Gabriel et al., 2018). ESM involves prompting participants to report their thoughts, feelings, or behaviors at specific intervals or in response to specific events, typically via mobile devices or digital platforms (Gabriel et al., 2018). This method aligns with CDST's principles by capturing moment-to-moment fluctuations in variables like emotion, motivation, and engagement, which are critical to understanding the dynamic processes in SLLT (Elahi Shirvan & Taherian, 2025).

ESM has been widely applied in SLLT to investigate dynamic constructs, particularly affective states and their interactions with linguistic performance. Khajavy et al. (2021) employed ESM to examine the dynamic relationships between WTC, anxiety, and enjoyment among 38 Iranian university students majoring in Teaching English as a Foreign Language. Their findings revealed significant variability in all three variables over time, both within weekly sessions and from one week to another. This study underscores the utility of ESM in capturing the temporal dynamics of L2 motivation and affect, providing insights into how these factors fluctuate and interact in real-time.

Arndt et al. (2021) reviewed the potential of ESM for capturing second language exposure and use, emphasizing its advantages over traditional retrospective methods. They highlighted how ESM provides real-time data on language exposure that captures individual variation across multiple contexts, offering greater ecological validity and precision in second language acquisition research. These studies demonstrate the effectiveness of ESM in capturing the dynamic nature of second language acquisition, revealing how contextual and emotional factors fluctuate and interact with learners' language use and performance. By offering real-time, longitudinal data, ESM provides a more nuanced understanding of individual trajectories in L2 learning, reinforcing the importance of idiographic approaches in SLLT research. However, ESM is not without limitations that researchers must carefully weigh. Compliance is a persistent concern: participants may skip prompts, respond carelessly under fatigue, or alter their natural behavior in response to repeated signaling — a phenomenon known as prompt reactivity (Gabriel et al., 2018). Missing data is likewise common in ESM studies, especially over extended data collection periods, and can introduce systematic bias if non-response is related to the construct being measured (e.g., participants skipping prompts precisely when they are most anxious). Construct instability presents a further challenge: some psychological constructs may not be meaningfully measurable at the momentary level, particularly those that require reflection or contextual integration to be rated accurately. It is also important to distinguish between ESM as a general data collection strategy and specific implementation choices — such as signal-contingent versus event-contingent designs, fixed versus random intervals, and ecological momentary assessment versus daily diary formats — each of which carries different implications for data quality and analytical requirements (Shiffman et al., 2008).

3.2 Idiodynamic method

The idiodynamic method is another specialized approach for capturing moment-to-moment fluctuations in psychological states during specific tasks, making it particularly suited for studying dynamic processes in SLLT (MacIntyre & Legatto, 2011). Unlike ESM, which collects data at intervals throughout the day, the idiodynamic method focuses on real-time, task-specific experiences, often using video-stimulated recall or continuous self-ratings to track variables like WTC, FLA, or motivation (MacIntyre & Gregersen, 2022). Participants engage in a language task, such as a speaking activity, and immediately afterward rate their experiences on a second-by-second or minute-by-minute basis, often while reviewing a video of their performance.

The idiodynamic method is particularly valuable for studying the micro-level dynamics of SLLT, where variables like anxiety or engagement fluctuate rapidly during tasks (MacIntyre & Gregersen, 2022). By focusing on individual experiences, it addresses the non-ergodic nature of language processes, avoiding the pitfalls of group-based analyses (Lowie & Verspoor, 2019). The method's emphasis on real-time data collection also enhances ecological validity, as it captures experiences in the context of actual language use (Elahi Shirvan et al., 2020).

The idiodynamic method has been widely applied to study affective and cognitive dynamics in SLLT, particularly in the context of speaking tasks and classroom interactions. MacIntyre and Legatto (2011) introduced the method to examine WTC during oral communication tasks, asking participants to rate their willingness to speak on a continuous scale while reviewing video recordings of their performance. Their findings revealed rapid fluctuations in WTC influenced by task-specific factors, such as topic familiarity or interlocutor behavior, demonstrating the method's ability to capture micro-level dynamics.

Elahi Shirvan & Talebzadeh (2017) explored the fluctuations of foreign language enjoyment (FLE) in conversation using an idiodynamic perspective through ESM. Their research demonstrated how FLE fluctuated in real-time interactions, providing a more comprehensive dynamic understanding of how affective states impact language use and communication. The study highlighted the need for idiographic methods to capture these real-time, context-specific emotional dynamics, which cannot be fully understood through aggregate measures alone. Kruk (2021) extended the idiodynamic method to study FLA in virtual language learning environments, finding that anxiety levels varied significantly within a single task, driven by factors like technical difficulties or peer feedback. This study highlighted the method's sensitivity to context-specific fluctuations, aligning with CDST's interconnectedness principle (van Dijk et al., 2025).

In teacher-focused research, the idiodynamic method has been used to explore pedagogical dynamics. For example, Gregersen et al. (2014) applied the method to study teacher anxiety during classroom instruction, finding that moment-to-moment fluctuations were influenced by student engagement and lesson planning. These applications demonstrate the idiodynamic method's versatility in capturing both learner and teacher dynamics, providing insights into the non-linear processes of SLLT (Larsen-Freeman & Cameron, 2008). While the idiodynamic method and ESM share a commitment to capturing intra-individual variability in real time, they differ meaningfully in scope, procedure, and practical demands, and these differences should be kept conceptually distinct. ESM samples experiences across naturally occurring daily contexts at regular or random intervals, whereas the idiodynamic method is task-bound — it captures fluctuations during a delimited activity such as a speaking task, using video-stimulated recall to elicit second-by-second or minute-by-minute self-ratings. This makes the idiodynamic method particularly sensitive to micro-level task dynamics but less suited to capturing variability across contexts or over extended time periods. The idiodynamic method also carries specific practical burdens that ESM does not: it requires video recording of participants, which raises privacy considerations; it depends on participants' willingness and ability to engage in retrospective rating while watching themselves, which can introduce reflection or social desirability biases; and the resulting continuous-rating data require specialized coding and segmentation procedures before analysis. Researchers choosing between the two methods should base their decision on the temporal scale and contextual scope of the research question rather than defaulting to one method as universally superior.

Despite their advantages, these approaches present challenges, including participant burden, data complexity, and contextual variability. Frequent prompts can lead to participant fatigue, especially in long-term studies, requiring careful study design to balance frequency and engagement (Gabriel et al., 2018). Additionally, the large volume of data collected necessitates advanced analytical techniques like DSEM (see nest section), which can handle temporal dependencies and manage missing data (Hamaker et al., 2018).

3.3 Dynamic structural equation modeling as a data analysis method

DSEM represents a hybrid analytical approach that combines elements of time-series, multilevel modeling (MLM), and structural equation modeling (SEM) into one comprehensive framework (Elahi Shirvan & Taherian, 2025). In non-technical terms, DSEM can be thought of as a tool that tracks how a person's experiences (such as their anxiety or motivation) change from one moment to the next, examines whether those changes predict future experiences, and then compares these individual patterns across many people to identify what is broadly shared. Unlike traditional statistical models that produce a single average for a group, DSEM generates a model for each individual and then organizes these individual models into a coherent picture of group-level trends. This combination allows researchers to study both the temporal dynamics within individual learners and the variability across individuals, which is crucial for understanding complex and dynamic systems like SSLT (Hamaker et al., 2023). Table 1 indicates a detailed breakdown of how each of these models contributes to addressing the ergodicity issue and enhancing generalizability.

Table 1

Hybrid Analytical Approach of DSEM

Approach	Focus	Key Function	Description
Time-series Modeling	Analyzing data from a single individual over multiple time points to capture non-linear changes in language development.	Models how previous states influence future states, useful for studying fluctuations, transitions, and momentum in learning.	Helps address the ergodicity issue by modeling individual learner trajectories over time, capturing non-linear changes and fluctuations.
Multilevel Modeling (MLM)	Capturing variation in dynamics across individuals while considering the nested structure of the data.	Accounts for individual differences in developmental trajectories and varying impacts of predictors.	Enhances generalizability by accounting for individual differences in developmental patterns, allowing for more accurate comparisons.
Structural Equation Modeling (SEM)	Modeling both within-person processes and between-person differences simultaneously to identify causal relationships.	Incorporates path or factor analysis to identify directional dynamic associations between variables such as language proficiency and emotional states; note that DSEM supports lagged predictive inference rather than strict causal claims.	Contributes to generalizability by integrating individual and group dynamics, providing a comprehensive model for broader insights while maintaining individual variability.

3.3.1 DSEM decomposition: Person means and temporal deviations

A fundamental aspect of DSEM is the decomposition of individual trajectories into two components: person means (also known as attractors) and temporal deviations (also known as states) [See figure 1, Panel A]. This decomposition is critical for distinguishing between stable, trait-like characteristics and dynamic, state-like fluctuations in the data (Hamaker et al., 2018).

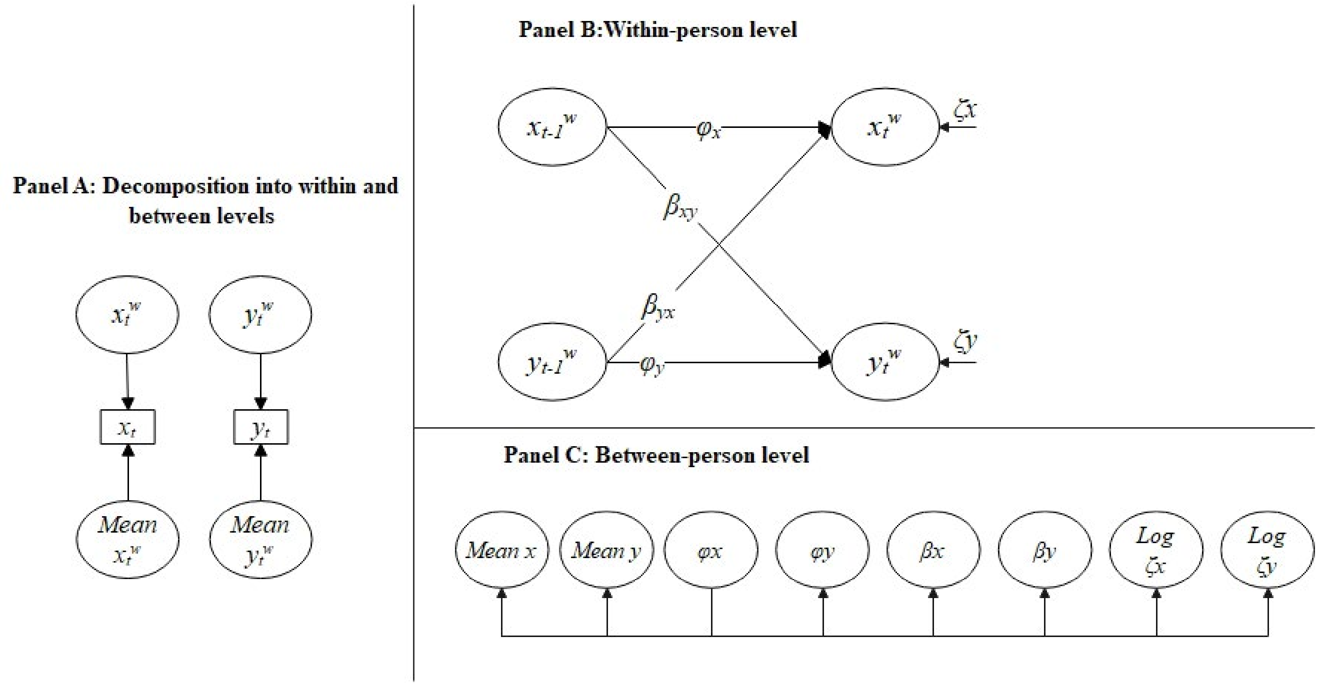
Person means refer to an individual's long-term equilibrium or average level with respect to a given variable (Elahi Shirvan et al., 2026). These are considered attractor states, which represent the stable baseline characteristics of an individual. These values are expected to remain relatively stable over time unless influenced by external factors or significant learning events.

Temporal deviations represent the short-term fluctuations or dynamic changes from the individual's baseline or attractor state (Solhi et al., 2026). These deviations capture the transient, state-like aspects of an individual's learning process. By analyzing these deviations, researchers can examine how learners react to specific learning experiences, teaching interventions, or contextual factors (e.g., peer interactions, feedback, or task difficulty) (Elahi Shirvan & Taherian 2025).

The decomposition into attractors and states is particularly valuable because it allows researchers to differentiate between stable individual traits and the dynamic, time-varying processes that influence SLA. This distinction is key in understanding how individual learners respond to various learning conditions over time, as well as how their internal states may influence their language learning trajectories.

Figure 1

DSEM



3.3.2 Levels and Parameters in DSEM

DSEM operates on two levels of analysis: the intra-individual level (within-person) and the inter-individual level (between-person). These two levels of analysis provide a comprehensive understanding of how dynamic processes unfold both within an individual and across individuals in a sample. This section explores the key parameters of DSEM at two levels:

At the within-person level, DSEM starts with individual-level (N = 1) time series analyses (see Figure 1, Panel B). Key parameters include autoregressive effects (e.g., φ), which measure the stability of a variable (e.g., negative affect) from one time point to the next, reflecting inertia or carryover. Cross-lagged effects (e.g., β) assess how changes in one variable predict subsequent changes in another, capturing lagged relationships. These parameters are estimated using time series models, allowing for the analysis of temporal dependencies within each individual. Residual variances at this level (e.g., ζ) represent unexplained variability after accounting for lagged effects, indicating random fluctuations or measurement error specific to each time point.

At the between-person level, the MLM component then extends this analysis to multiple cases simultaneously (see Figure 1, Panel C). This level estimates mean parameters (e.g., mean autoregressive and cross-lagged effects) and their variability across individuals. For instance, the average autoregressive effect (φ) and its standard deviation across persons quantify how consistently individuals exhibit stability, while the average cross-lagged effect (β) and its variability reflect differential responsiveness to events. Residual variances at this level (e.g., random effects on φ and β) capture between-person heterogeneity in these dynamic processes, highlighting how traits or contextual factors may influence parameter estimates. The SEM component further enhances this by modeling these individual differences using path analysis or factor analysis, providing a structured way to generalize findings.

4 Empirical example: L2 teacher boredom and creativity

Elahi Shirvan et al. (2024) conducted a study that highlights the advantages of applying DSEM within CDST research. Their investigation focuses on the daily fluctuations of boredom and creativity among L2 teachers, exploring how these two constructs interact and how control-value beliefs influence these interactions. The concept of "control-value beliefs" comes from control-value theory, which posits that emotions like boredom are shaped by an individual's perceived control over their actions and the value they assign to the activity. Specifically, control beliefs relate to an individual's perceived ability to influence a situation, such as how much control teachers feel they have over classroom management or their professional responsibilities. Value beliefs, on the other hand, refer to the importance a teacher places on their teaching tasks. Teachers who value their work more are likely to find it more engaging, which can impact their emotional experiences, including boredom and creativity. The data for this study were collected using a 5-day ESM facilitated by the Paco mobile app, involving 98 L2 teachers and yielding 423 observations. Participants reported their daily levels of boredom and creativity, while their control-value beliefs were assessed as time-invariant variables at the beginning of the study. The study addressed three key research questions:

1. How quickly do L2 teachers restore balance after experiencing boredom and creativity over the course of a workweek?
2. What is the dynamic relationship between boredom and creativity for L2 teachers over the course of a workweek?
3. How does the relationship between boredom and creativity vary across L2 teachers, considering their control-value beliefs?

The study conducted at two intra-Individual and inter-Individual levels, using three models to address the research questions:

4.1 Intra-individual level: A time-series model

At the intra-individual level, DSEM was used to analyze the temporal dynamics of boredom and creativity for each teacher throughout the workweek. This model examined autoregressive effects (e.g., how a teacher's boredom or creativity at one time point influenced the same variable at the next time point [RQ1]) and cross-lagged effects (e.g., how boredom influenced creativity, and vice versa, across time points [RQ2]). The autoregressive coefficients indicated that boredom showed greater persistence than creativity, meaning that teachers needed more time to recover from boredom. In contrast, creativity was more transient, returning to baseline levels more quickly. These findings provided insights into the temporal dynamics of boredom and creativity at the individual level.

However, the cross-lagged regression results showed variability across teachers: For some teachers, there were significant positive cross-lagged effects from boredom to creativity; for others, significant negative cross-lagged effects were observed from boredom to creativity, and for a subset of teachers, cross-lagged effects were non-significant, implying no direct relationship between boredom and creativity over time.

To better understand these differences, the study examined the relationship between boredom and creativity in the context of teachers' Control-Value beliefs (See Table 2). Some teachers exhibited strong reciprocal relationships between boredom and creativity, while others showed weaker or unidirectional effects. Participants were grouped into four clusters based on the nature of the relationship between boredom and creativity (positive, negative, or nonsignificant), with credibility intervals for standardized coefficients containing zero.

Table 2
Different Patterns of the Dynamic Interplay between Teacher Creativity and Boredom

Cluster	Control		Intrinsic		Achievement		Features
	M	SD	M	SD	M	SD	
1. Innovative teachers N=14 %=14.58% β_B: Sig and positive. β_C: Sig and negative.	5.86	.17	6.84	.11	4.68	.14	Proactive and innovative teachers with a strong sense of control, finding intrinsic value in their work, and moderately believing in significant achievement outcomes.
2. Adaptable teachers N=21 %=21.87% β_B: Nonsig. β_C: Sig and negative.	5.11	.21	5.98	.08	4.95	.09	Flexible teachers who maintain moderate control, value intrinsic aspects, and believe in meaningful achievements. They are open to adapting their methods
3. Traditionalist teachers N=54 %=54.16% β_B: Sig and negative. β_C: Sig and positive.	4.47	.25	5.02	.21	5.28	.24	They prefer conventional teaching methods, with a moderate sense of control, intrinsic value belief and high sense of achievement value beliefs in achieving outcomes through traditional approaches.
4. Dissatisfied teachers N=7 %= 7.29% β_B<0, CIs for β_C contained zero. β_B: Sig and negative. β_C: Nonsig.	3.86	.28	4.27	.08	4.15	.05	Dissatisfied teachers experience dissatisfaction, with low perceived control, intrinsic value, and achievement value beliefs.

Note. Abbreviations: Control, teacher's perceived control; intrinsic, teacher's teaching intrinsic values; achievement, teacher's teaching achievement values. *M*, mean; *SD*, standard deviation; CIs, credibility intervals. β_B, cross-lagged effects from boredom to creativity. β_C, cross-lagged effects from creativity to boredom.

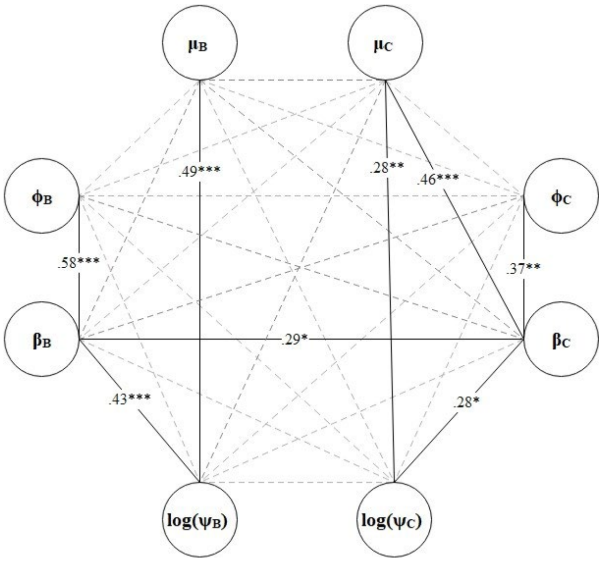
The analysis revealed that for different patterns of changes as clusters based on the nature of the relationship between boredom and creativity (positive, negative, or nonsignificant). Control-value beliefs, measured as time-invariant variable, explained these individual differences. A summary of these results is presented in Table 2.

4.2 Inter-Individual Level: MLM, and SEM

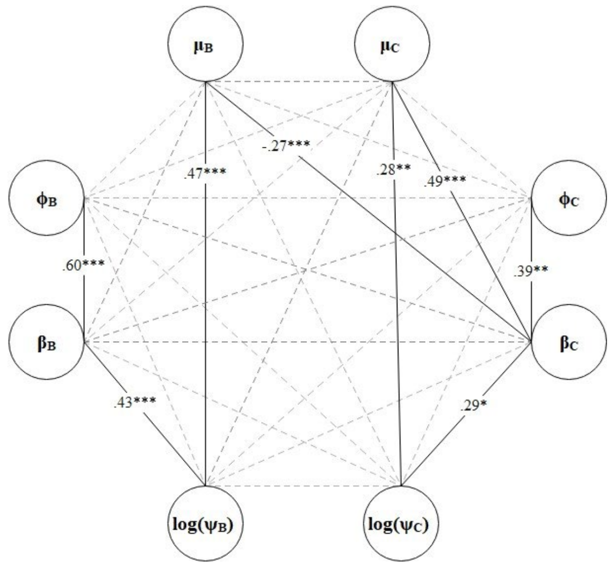
At the inter-individual level, DSEM incorporated MLM to examine variability in dynamics across teachers (see Figure 2). The model revealed significant between-person variation in how teachers experienced and responded to boredom and creativity in for clusters. For example, the correlations between the mean levels of boredom and creativity in Clusters 1 and 2 were non-significant, suggesting that for innovative and adaptable teachers, the trait level of boredom is not necessarily linked to the trait level of creativity. This finding implies that these teachers may be less dependent on external rewards, and their moderate to high levels of perceived control and intrinsic value beliefs may contribute to their resilience against the negative effects of boredom on creativity. In contrast, for teachers in Clusters 3 and 4, higher trait boredom levels were associated with lower trait creativity levels. These teachers, with lower perceived control, may feel restricted in their decision-making and ability to implement creative ideas in reshaping their environments.

Figure 2

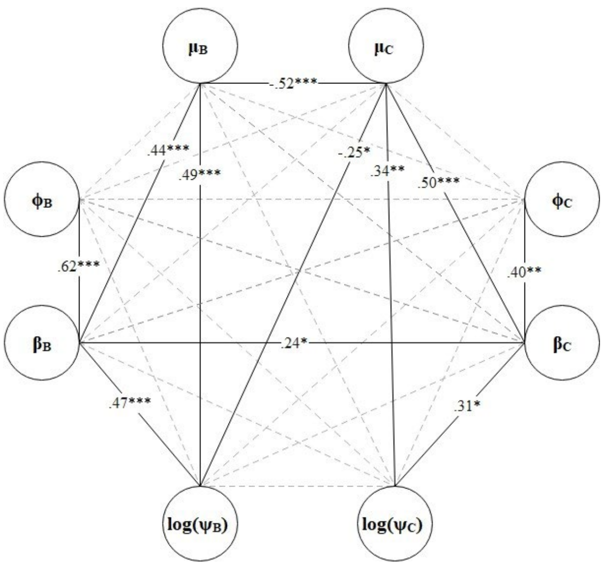
Between-Individual Variation in Four Clusters



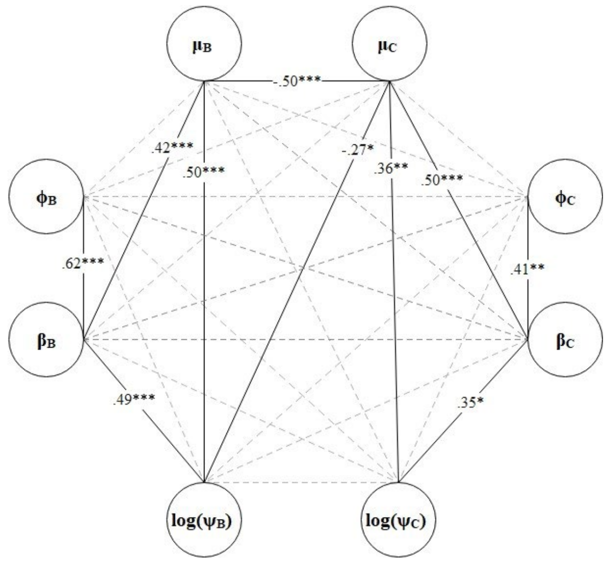
Cluster 1. Innovative teachers



Cluster 2. Adaptable teachers



Cluster 3. Traditionalist teachers



Cluster 4. Dissatisfied teachers

This research showcases how DSEM, along with additional clustering, can capture both within- and between-person dynamics of boredom and creativity, revealing unique temporal characteristics, inter-individual variability, and the role of control-value beliefs in shaping these emotional dynamics among L2 teachers.

Other two studies in SLLT have used DSEM to explore the dynamic relationships between various affective and cognitive variables over time, addressing the ergodicity issue. Pawlak et al. (2025) utilized DSEM to investigate the dynamic reciprocal associations between emotions during AI-assisted L2 writing tasks and data-driven learning. Their study revealed how emotions such as anxiety and enjoyment fluctuate throughout the writing process and interact with learners' engagement and performance. The findings underscore the importance of capturing these temporal dynamics to understand the complex, evolving nature of emotional states in second language learning. Similarly, Solhi, et al. (2024) applied DSEM to examine the impact of L2 teacher enjoyment on EFL learners' enjoyment and willingness to communicate. Their study demonstrated that positive emotional dynamics in the classroom are not only influenced by individual teacher emotions but also serve as a catalyst for enhancing learners' own emotional engagement and communication willingness. These studies illustrate how DSEM enables a deeper understanding of the temporal and reciprocal influences between emotional states, cognition, and behavior in language learning contexts.

5 Affordances and Contributions of ILD and DSEM in SLLT

In the context of SLLT, the integration of ILD and DSEM presents significant methodological affordances that address critical challenges in understanding the dynamic and non-ergodic nature of language learning. These methods contribute notably to the resolution of the ergodicity issue, a persistent concern in SLLT research where group-level statistics are often assumed to reflect individual processes. SLLT processes, however, are inherently non-ergodic, with individual learners and teachers following unique trajectories influenced by personal and contextual factors (Lowie & Verspoor, 2019). By focusing on intra-individual dynamics and modeling person-specific patterns, ILD and DSEM offer a robust framework to capture and analyze the variability and complexity of individual learning trajectories.

DSEM, in particular, facilitates a nuanced understanding of dynamic processes within language learning by allowing for the modeling of autoregressive and cross-lagged relationships over time. These relationships are essential for understanding how emotions, such as enjoyment and anxiety, influence one another and evolve in response to external and internal factors. By incorporating random effects and multilevel structures, DSEM addresses the critical need for a bottom-up generalizations within SLLT research. This dual perspective allows researchers to examine individual patterns while still identifying broader, shared trends across a population of learners (van Dijk et al., 2025). Through this approach, DSEM creates an analytical space where both general principles and individual variability are accounted for, enhancing the validity of findings across different contexts and learning environments.

A key methodological affordance of DSEM is its capacity to model the "attractor states" of learners, which represent the long-term equilibria of variables such as enjoyment, motivation, or anxiety. Attractor states serve as a dynamic baseline to which learners return in the absence of disruptions (Hiver et al., 2021). Temporal deviations from these states—manifested as fluctuations in emotional or motivational experiences—can be interpreted as affective states that influence a learner's ability to engage with language learning. This dual consideration of attractor states and affective states allows for a more comprehensive understanding of the fluctuating and dynamic interactions among various factors in the language learning process. For example, DSEM can estimate how long it takes for a learner to recover their baseline enjoyment after experiencing a drop due to a challenging task, or how individual differences in resilience affect recovery time.

Beyond autoregressive relationships, DSEM's capacity to model cross-lagged effects is particularly valuable in SLLT. Cross-lagged relationships examine the reciprocal influence between variables over time, such as how a learner's enjoyment might impact their anxiety levels, and vice versa. This methodological strength is especially relevant in the context of CDST, which emphasizes the co-evolution of multiple variables within an individual. By analyzing cross-lagged effects, DSEM provides critical insights into the dynamic interplay between affective and motivational factors, such as the role of anxiety in shaping enjoyment and the reverse influence of enjoyment on anxiety (van Dijk et al., 2025). This comprehensive approach enhances the capacity of researchers to uncover the temporal and predictive associations between emotional and cognitive processes in language learning.

Moreover, DSEM's ability to model system noise further strengthens its contribution to SLLT research. System noise refers to the unmeasured variables that can affect learners' emotional and cognitive experiences. By analyzing residual variances at both the intra-individual and inter-individual levels, DSEM allows researchers to capture nuanced differences in how learners are impacted by these unmeasured factors. This capacity to model system noise is vital for uncovering individual differences in how learners respond to external challenges, such as task difficulty or classroom dynamics, providing a richer and more detailed understanding of the variability in language acquisition (Hamaker et al., 2023).

6 Challenges of ILD and DSEM in SLLT

The integration of ILD) and DSEM into SLLT has significantly advanced research methodologies. However, these approaches present several challenges that must be addressed to optimize their utility in the field. This section discusses the methodological, analytical, and practical challenges of ILD and DSEM in SLLT.

6.1 Methodological challenges

6.1.1 Sample size requirements

A major challenge in using ILD and DSEM in SLLT research is the need for large sample sizes and multiple time points. DSEM generally requires a minimum of 50 participants and 25 time points per participant to ensure sufficient statistical power and model convergence (McNeish, 2019). In SLLT, where recruitment can be limited, particularly in specialized language programs or rural settings, this requirement can be difficult to meet. ILD, which often necessitates frequent measurements, exacerbates this issue. To overcome this, researchers can adopt small-sample DSEM techniques like Bayesian estimation, which has proven to enhance model stability with fewer participants (McNeish, 2019).

6.1.2 Data Collection Burden

The frequent data collection required by ILD methods, particularly through ESM or idiodynamic approaches, places a significant burden on participants, potentially leading to attrition and fatigue (Shiffman et al., 2008). In SLLT, where both teachers and learners have demanding schedules, frequent data collection can disrupt daily routines and reduce response rates. To mitigate this, researchers can optimize prompt timing, shorten surveys, and use adaptive methods that respond to specific triggers (e.g., lesson completion) (Shiffman et al., 2008).

6.1.3 Data Quality and Measurement Invariance

Maintaining data quality in ILD studies is challenging due to missing data, which is common in naturalistic settings like classrooms. Disruptions such as student absences can interrupt data collection

(Ortega & Iberri-Shea, 2005). Additionally, response bias and measurement invariance, particularly when participants self-report on constructs like anxiety or motivation, complicate the process. Ensuring consistency across time points and participants is essential, but resource-intensive, requiring validated scales and robust measurement models (McNeish & Hamaker, 2020).

6.2 Analytical challenges

6.2.1 Analytical complexity

DSEM is analytically complex, integrating multiple statistical techniques such as time-series analysis, MLM, and SEM. This complexity requires advanced statistical expertise, particularly in autoregressive and cross-lagged effects, random effects, and Bayesian estimation (Asparouhov et al., 2018). DSEM's steep learning curve can hinder its adoption, especially in SLLT, where many researchers lack this specialized knowledge. Solutions may include targeted training and the development of user-friendly software tools like Mplus.

6.2.2. Model convergence and overfitting

DSEM faces critical challenges regarding model convergence and overfitting. Convergence problems occur when the model fails to stabilize due to insufficient data or overly complex specifications (Hamaker et al., 2023). Overfitting occurs when a model captures random noise rather than meaningful patterns, thus undermining the generalizability of results (Hamaker et al., 2018).

6.2.3. Software accessibility

Access to software also hinders the widespread adoption of DSEM. MPlus offers an intuitive interface but requires a paid license, making it inaccessible for researchers in low-resource settings (Asparouhov et al., 2018). While open-source tools like R's dynr or Stan offer flexibility, they require significant programming expertise, which limits accessibility for researchers without coding skills (Elahi Shirvan & Taherian, 2025).

6.3. Practical challenges

6.3.1. Participant recruitment and retention

Recruiting and retaining participants for ILD studies is particularly challenging in SLLT due to the demanding schedules of learners and teachers. Frequent surveys can cause fatigue, leading to lower response rates and high dropout rates. Strategies like asynchronous ESM prompts or gamified interfaces could improve participant engagement, but they require thoughtful implementation to ensure data quality.

6.3.2. Ethical considerations

Ethical issues in ILD studies, especially those involving sensitive emotional constructs (e.g., teacher burnout or FLA), require careful attention to privacy, consent, and participant well-being (MacIntyre & Gregersen, 2022). Researchers must establish robust informed consent protocols and ensure that participants feel protected throughout the data collection process, particularly in environments with power dynamics, such as between teachers and students.

6.3.3 Resource constraints

The implementation of ILD and DSEM requires significant resources, including software, devices, and participant incentives. In under-resourced settings, limited access to these resources can impede research. Open-source tools and low-cost mobile apps may help alleviate these constraints, but their successful implementation requires infrastructure support and training. Taken together, these challenges are not equally pressing or equally tractable. The most serious and least easily resolved challenge is analytical complexity: the statistical expertise required to implement DSEM correctly is substantial, and errors in model specification — such as treating cross-lagged effects as causal or ignoring model convergence diagnostics — can produce misleading findings that are difficult for non-specialist readers to detect. This challenge is compounded by limited software accessibility and the scarcity of DSEM-trained methodologists in applied linguistics. The challenges of data collection burden and participant retention, while significant, are more manageable through thoughtful design choices such as adaptive prompting, participant incentives, and pre-registration of acceptable attrition thresholds. Ethical concerns and resource constraints are field-dependent: they are particularly acute in under-resourced or institutional research settings but may be less prohibitive in well-funded university contexts. Prioritizing these challenges rather than cataloguing them helps researchers make informed decisions about whether and how to adopt ILD and DSEM.

7 Conclusion

The integration of ILD and DSEM into SLLT has substantially advanced the way researchers capture the dynamic, non-linear, and individualized nature of language acquisition. These methods address the limitations of traditional approaches by focusing on intra-individual variability, temporal dynamics, and contextual interactions, aligning with CDST. DSEM's ability to model non-ergodicity, non-linear relationships, and identify subgroups of learners or teachers makes it an essential tool for personalized pedagogy and evidence-based practices in language education.

Despite challenges such as sample size requirements, data collection burden, and analytical complexity, these methods offer significant contributions to SLLT research. By providing granular insights into learners' and teachers' experiences, ILD and DSEM enable tailored interventions that promote engagement, reduce educational disparities, and improve learning outcomes. Furthermore, they bridge the gap between research and practice, ensuring that SLLT scholarship has direct real-world applications.

Looking forward, addressing the challenges of ILD and DSEM through innovations in data collection, analysis, and accessibility will enhance their impact. Ultimately, these methods will continue to shape the future of SLLT research, fostering more personalized, equitable, and effective language teaching and learning.

References

Arndt, H. L., Granfeldt, J., & Gullberg, M. (2021). Reviewing the potential of the Experience Sampling Method (ESM) for capturing second language exposure and use. *Second Language Research*, 026765832110200. https://doi.org/10.1177/02676583211020055

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388. https://doi.org/10.1080/10705511.2017.1406803

de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, *10*(1), 7–21. https://doi.org/10.1017/s1366728906002732

Elahi Shirvan, M., & Talebzadeh, N. (2017). Exploring the fluctuations of foreign language enjoyment in conversation: An idiodynamic perspective. *Journal of Intercultural Communication Research*, *47*(1), 21–37. https://doi.org/10.1080/17475759.2017.1400458

Elahi Shirvan, M., & Taherian, T. (2025). Dynamic structural equation modeling in response to the generalizability concern about CDST-inspired research. In W. Lowie, M. Rosmawati, & V. De Wilde (Eds.), *Research Methods in Complex Dynamic Systems Theory Approaches to Second Language Development* (pp. 35–56). John Benjamins.

Elahi Shirvan, M., Taherian, T., & Yazdanmehr, E. (2020). The dynamics of foreign language enjoyment: An ecological momentary assessment. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01391

Elahi Shirvan, M., Taherian, T., Pawlak, M., & Kruk, M. (2024). The dynamics of L2 teacher boredom and their link to creativity: A ban or boon for boredom. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.12595

Elahi Shirvan, M., Taherian, T., Kruk, M., & Pawlak, M. (2025). A person-specific perspective on the dynamics of anxiety in foreign language learning: A dynamic P-technique factor analysis. *Studies in Second Language Learning and Teaching*. https://doi.org/10.14746/ssllt.45784

Elahi Shirvan, M., Solhi, M., Taherian, T., & Thumvichit, A. (2026). Boredom fuels boredom: An experience sampling study on L2 learner and teacher boredom in reading classes. *Acta Psychologica*, 266, 106887. https://doi.org/10.1016/j.actpsy.2026.106887

Gabriel, A. S., Podsakoff, N. P., Beal, D. J., Scott, B. A., Sonnentag, S., Trougakos, J. P., & Butts, M. M. (2018). Experience sampling methods: A discussion of critical trends and considerations for scholarly advancement. *Organizational Research Methods*, *22*(4), 969–1006. https://doi.org/10.1177/1094428118802626

Gates, K. M., Chow, S.-M., & Molenaar, P. (2023). *Intensive longitudinal analysis of human processes*. Chapman & Hall.

Gregersen, T., MacIntyre, P. D., & Meza, M. D. (2014). The motion of emotion: Idiodynamic case studies of learners' foreign language anxiety. *The Modern Language Journal*, *98*(2), 574–588. https://doi.org/10.1111/modl.12084

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO Study. *Multivariate Behavioral Research*, *53*(6), 820–841. https://doi.org/10.1080/00273171.2018.1446819

Hamaker, E. L., Asparouhov, T., & Muthén, B. (2023). Dynamic structural equation modeling as a combination of time series modeling, multilevel modeling, and structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (2nd ed., pp. 576–596). Guilford Publications.

Herdina, P., & Jessner, U. (2002). *A dynamic model of multilingualism: perspectives of change in psycholinguistics*. Multilingual Matters.

Hiver, P., & Al-Hoorie, A. H. (2020). *Research methods for complexity theory in applied linguistics*. Multilingual Matters.

Hiver, P., Al-Hoorie, A. H., & Evans, R. (2021). Complex dynamic systems theory in language learning: A scoping review of 25 years of research. *Studies in Second Language Acquisition*, *44*(4), 1–29. https://doi.org/10.1017/s0272263121000553

Khajavy, G. H., MacIntyre, P. D., Taherian, T., & Ross, J. (2021). Examining the dynamic relationships between willingness to communicate, anxiety and enjoyment using the experience sampling method. *Second Language Learning and Teaching*, 169–197. https://doi.org/10.1007/978-3-030-67634-6_9

Kruk, M. (2021). *Investigating dynamic relationships among individual difference variables in learning English as a foreign language in a virtual world*. Springer.

Kruk, M., Pawlak, M., Taherian, T., Yüce, E., Elahi Shirvan, M., & Barabadi, E. (2023). When time matters: Mechanisms of change in a mediational model of foreign language playfulness and L2 learners' emotions using latent change score mediation model. *Studies in Second Language Learning and Teaching*, *13*(1), 39–69. https://doi.org/10.14746/ssllt.37174

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

Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.

Liao, S. J., Fang, F., Cao, X., & Wang, W. (2025). Almost like having my own teacher beside me": A learner's investment in Informal Digital Learning of English (IDLE) and English-medium Instruction (EMI) learning. *International Journal of TESOL Studies*, *250712*, 1-19. https://doi.org/10.58304/ijts.250712

Lord, F. M., and Novick, M. R. (1968). Statistical Theories of Mental Test Scores. Reading, MA: Addison-Wesley

Lowie, W. M., & Verspoor, M. H. (2019). Individual differences and the ergodicity problem. *Language Learning*, *69*, 184–206. https://doi.org/10.1111/lang.12324

Lowie, W., Van Dijk, M., Chan, H., & Verspoor, M. (2017). Finding the key to successful L2 learning in groups and individuals. *Studies in Second Language Learning and Teaching*, *7*(1), 127. https://doi.org/10.14746/ssllt.2017.7.1.7

MacIntyre, P. D., & Gregersen, T. (2022). The idiodynamic method: willingness to communicate and anxiety processes interacting in real time. *International Review of Applied Linguistics in Language Teaching*, *60*(1). 67-84. https://doi.org/10.1515/iral-2021-0024

MacIntyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied Linguistics*, *32*(2), 149–171. https://doi.org/10.1093/applin/amq037

MacIntyre, P. D., & McGillivray, M. F. (2023). The inner workings of anxiety in second language learning. *Annual Review of Applied Linguistics*, *43*, 88–104. https://doi.org/10.1017/s0267190523000065

McNeish, D. (2019). Two-Level dynamic structural equation models with small samples. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(6), 948-966. https://doi.org/10.1080/10705511.2019.1578657

McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods*, *25*(5), 610–635. https://doi.org/10.1037/met0000250

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research & Perspective*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Ortega, L., & Iberri-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics*, *25*, 26–45. https://doi.org/10.1017/s0267190505000024

Pallotti, G. (2022). Cratylus' silence: On the philosophy and methodology of Complex Dynamic Systems Theory in SLA. *Second Language Research*, *38*(3), 689-701. https://doi.org/10.1177/0267658321992451

Pawlak, M., Kruk, M., Elahi Shirvan, M., Taherian, T., & Karimpour, S. (2025). Dynamic reciprocal associations of AI-assisted L2 writing task emotions in data-driven learning: a dynamic structural equation modeling. *Computer Assisted Language Learning*, 1–30. https://doi.org/10.1080/09588221.2025.2498094

Shiffman S., Stone A.A., & Hufford M.R. (2008). Ecological momentary assessment. Annual Review of Clinical Psychology, *4*, 1–32. https://doi.org/10.1146/ annurev.clinpsy.3.022806.091415

Solhi, M., Elahi Shirvan, M., & Benlioğlu, B. (2024). Enjoyment begets enjoyment: An experience sampling study on the impact of L2 teacher enjoyment on EFL learners' learning enjoyment and willingness to communicate. *System*, *126*, 103493. https://doi.org/10.1016/j.system.2024.103493

Solhi, M., Elahi Shirvan, M., & Taherian, T. (2026). A dynamic bidirectional system of perceived teacher boredom, language learners' boredom, and L2 writing motivation: a feedback loop system. *Asian-Pacific Journal of Second and Foreign Language Education*, *11*(1), 1-24. https://doi.org/10.1186/s40862-026-00388-7

Taherian, T., Elahi Shirvan, M., Yazdanmehr, E., Kruk, M., & Pawlak, M. (2023). A longitudinal analysis of informal digital learning of english, willingness to communicate and foreign language boredom: A latent change score mediation model. *The Asia-Pacific Education Researcher*. https://doi.org/10.1007/s40299-023-00751-z

Teng, M. F., & Yip, J. C. (2025). A narrative inquiry into language teacher identity construction in the global GenAI context. *International Journal of TESOL Studies*, *250901*, 1-21. https://doi.org/10.58304/ijts.250901

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Mit Press.

van Dijk, M., Lowie, W., Smit, N., Verspoor, M., & van Geert, P. (2025). Complex dynamic systems theory as a foundation for process-oriented research on second language development. *Second Language Research*, *41*(4), 805-819. https://doi.org/10.1177/02676583241246739

Verspoor, M.H., de Bot, K., Lowie, W. (2011). *A dynamic approach to second language development: methods and techniques*. John Benjamins, Amsterdam

Verspoor, M., & de Bot, K. (2021). Measures of variability in transitional phases in second language development. *International Review of Applied Linguistics in Language Teaching*, *60*(1), 85–101. https://doi.org/10.1515/iral-2021-0026

Wen, Y., & Fang, F. (2026). From "Striving to Survive" to "Striving to Thrive": A Narrative Inquiry Into a Student's Motivational Development via Translanguaging in EAP and EMI Contexts. *International Journal of Applied Linguistics*. Advance online publication. https://doi.org/10.1111/ijal.70193

Majid Elahi Shirvan holds a PhD in Teaching English as a Foreign Language from Ferdowsi University of Mashhad, Iran. He is currently an Associate Professor at the University of Bojnord, Iran. His main research interest is the dynamics of psychological aspects of foreign language learning and teaching. He has published in different leading international journals such as Ecological Psychology, Foreign Language Annals, Studies in Second Language Learning and Teaching, Studies in Second Language Acquisition, and Language Teaching Research.

ORCID: https://orcid.org/0000-0003-3363-8273