Overview

From Simplicity to Complexity

As the contributions to this handbook show, the study of individual differences (IDs) in second language acquisition (SLA) is concerned primarily, though not exclusively, with why some learners are more successful than others in their rates and routes of development and their levels of ultimate attainment. Notably, it has now been over a decade since Dörnyei (2008, 2009) proposed to rethink ID variables in a situated, dynamic manner because ID variables are interconnected and interact with the environment while also developing and varying through time. The notion that complex dynamic systems theory (CDST) can offer new answers to existing questions surrounding language learners and their development has been taken up in both conceptual and empirical work (e.g., Gurzynski-Weiss, 2020; Sampson & Pinner, 2020). Since then, there has been a growing reorientation to the way in which scholars study IDs in SLA. This new approach to studying IDs has definitively put CDST on the research map. The increasing adoption of CDST methodological tools for researching why some language learners are more successful than others, and what makes them so, has resulted in a new research landscape in which CDST has firmly established its relevance and explanatory potential (Dörnyei, 2017).

All research methods for IDs have a number of built-in assumptions, some of which are unstated or implicit in the techniques of data elicitation and analysis. For more conventional, non-CDST research methods for IDs, a commonplace way of operating is to first isolate things from their environment, break things up into smaller parts, draw boundaries between things in order to study them more systematically, consider things as relatively stable and unchanging, and assume the underlying processes are relatively similar and predictable. These types of research, however, do not enable their users to investigate the combined interactions of multiple individual parts, or to study the effects of complex systems or situations where these outcomes are neither linear nor simple (Hiver & Al-Hoorie, 2016). Several methods in widespread use (e.g., cross-sectional correlational designs and pre-/post-experimental designs) operate from the assumption that relatively few elements isolated from their surroundings have a simple and predictable influence on learners and their second language development. Such methods do not readily lend themselves to studying IDs in ways that acknowledge their complex and dynamic realities and situate these phenomena firmly in context (MacIntyre et al., 2017). Instead, they are prone to oversimplifying predictive
analyses of outcomes, and often rely heavily on explanations that ignore temporal change, context-dependence, and complex causality.

In contrast, the CDST family of methodological tools takes a very different perspective as its point of departure—a systems view (see e.g., Larsen-Freeman, 2015). These methods posit that the reality of the human and social world is one in which, first, everything counts and everything is connected (i.e., the relational principle) and, second, everything changes (i.e., the adaptive principle) (Overton & Lerner, 2014). CDST research reconceptualizes the core of IDs in SLA, not exclusively as the conventional, modular independent variables they are often viewed and treated as, but as systems or systemic phenomena (e.g., with emergent outcomes and self-organized states) grounded in a context-dependent and dynamic view of development. Scholars championing this new way of thinking “to explain the dynamic development of real people in actual contexts” (Dörnyei, 2017, p. 87) admit that the reorientation which comes with using CDST challenges many of the field’s existing assumptions and suggests new approaches to inquiry (Hiver et al., 2021a).

Because the world is complex and dynamic, the unit(s) of analysis should reflect this and allow ID researchers to examine how parts of the whole relate to each other and change in context over time. The CDST family of research methods generally adopts complex systems as the unit of analysis. These complex systems (Hiver & Papi, 2019) consist of a number of elements or components situated in context. These components, at least one of which is an agent (i.e., the individual), interact with each other based on certain principles of interdependence. Over time, the components change as a result of their interactions with other components. From these interactions emerge system-wide outcomes and macro-level patterns of behavior.

Another consequential change of adopting the CDST family of methodological tools for IDs research is that time matters (Lemke, 2000) and that neither IDs nor SLA can be researched exclusively in static ways. As illustrated above, adaptive change is one pivotal characteristic of adopting the CDST perspective for research (Larsen-Freeman & Cameron, 2008). Initial conditions and histories have a critical role to play in understanding learners and their development (Verspoor, 2015). The added value of CDST research for IDs in SLA is that it allows researchers to refocus attention more explicitly on processes of change at various timescales and take a much more developmental perspective in their research.

**Technical Features**

In order to elaborate on how the CDST family of methodological tools operates, in this section I introduce a handful of quantitative and qualitative methods. For each, I review some technical details related to data elicitation and analyses, and discuss relevant strengths and challenges to applying these methods. Some, such as the idiodynamic method, have seen wider adoption than others. Generally speaking, however, CDST methods such as these remain less widely employed than other more conventional methods in IDs in SLA research.

**Idiodynamic Method**

Unlike many available methods that rely on averaging data from a number of units (e.g., individuals) and aim to represent the central tendency of the units and group-level variability, the idiodynamic method aims for a deeper, more individual-level understanding of the processes involved. From this perspective, individual idiosyncrasies are not noise to be eliminated in favor of a single, neat average (Van Geert & Steenbeek, 2014). When the units of a group are strongly similar to each other, then generalization from the group to the individual and vice versa is justifiable. However, this condition—known as ergodicity—is more often than not absent in IDs data as typically no one is average, and not all participants in a sample are equally influenced by an intervention (Lowie & Verspoor, 2019).
When the CDST researcher zooms in on one learner, for example, an important consideration is the timescale of the investigation. The timescale of a process can range from seconds to minutes to days to months to years. When it comes to investigating smaller timescales, researchers need a method to collect accurate, real-time data on the learner’s intra-individual variability during, say, a specific language task (see also Chinn & Sherin, 2014). The idiodynamic method was developed for this purpose (MacIntyre, 2012). The idiodynamic method aims to uncover fluctuations over time and the possible reasons underlying them, allowing the CDST researcher to use procedures that sharpen the focus on time-dependent variation within a single individual or unit. It also allows researchers to collect both online and offline measurements of learners’ variation in multiple IDs.

Specifically, the idiodynamic method involves video recording a learner’s speech, interactions, or task performance and then showing the participants (and others) the video to solicit stimulated recall ratings about their L2 speech, interaction, or task performance. The unique aspect of the idiodynamic method is that it allows the participant to provide ratings for every second of the task in a systematic manner. In phase 1 the participant engages in a task while being videotaped or audio recorded. Immediately after this, the audio or video data is loaded into the software (Anion Variable Tester is one free program). In phase 2, the participant is then asked to watch the video or listen to the recording and provide ratings on a specific ID variable (e.g., willingness to communicate) for every second of the task (quantitative data). As the hypothetical example in Figure 31.1 shows, these ratings can be done several times over using the same data for different variables and can even incorporate self-ratings, interlocutor ratings, or expert ratings. The software records and graphs reactions and provides a continuous quantitative measure that focuses on change in the variable(s) of interest. In phase 3, using stimulated recall, the participant is finally asked to provide explanations for their ratings (qualitative data). The software generates a graph and an Excel file of the ratings, while the qualitative data may be transcribed by the researcher for further analysis (e.g., L2 production measures such as complexity, accuracy, and fluency). Depending on the purpose of the investigation, the researcher might perform a horizontal analysis by comparing patterns across different participants (Figure 31.2). Alternatively, the researcher might perform a vertical analysis focusing on patterns within one individual.

Questions that the idiodynamic method enables a researcher to investigate may be observational or experimental. The idiodynamic method provides the researcher with interesting real-time, individual-level data related to processes of interest. In contrast to research questions based on observational or descriptive data, in the experimental case the researcher deliberately manipulates certain factors, such as the type of corrective feedback provided or the difficulty of the task, to investigate

L2 learner: “I’m really expecting to meet you tomorrow... sorry...um... I mean... I’m really looking forward to meeting you tomorrow...”

Figure 31.1 Idiodynamic Data for a Learner’s Speech Repair in an L2 Task.
their effect on different outcomes (see e.g., Figure 31.2). These outcomes and their associated processes can be psychological, such as engagement, anxiety, and willingness to communicate; they can also be linguistic, such as lexical choices, grammatical accuracy, the uptake of feedback, or fluency with language (e.g., Nagle et al., 2019). Example questions include (Hiver & Al-Hoorie, 2020):

- How do self-ratings (e.g., of motivation, anxiety, and competence) change over the course of a task?
- How do self-ratings relate to other-ratings (e.g., a researcher, a peer, or an expert observer)?
- To what does the learner attribute these fluctuations?
- What factors (e.g., proficiency level, interlocutor, task design, or instructional context) influence self-ratings and perceptions?
- Does the type of stimuli (e.g., audio only, video only, or audio + video) have an effect on the learner’s self-ratings or perceptions?
- What is the effect of training (e.g., pragmatic strategy instruction, corrective feedback, and task repetition) on self-ratings and perceptions?
- How do self-ratings and perceptions relate to other physiological measures (e.g., eye movements, heart rate, skin conductance, EEG, and fMRI)?

The idiodynamic method relies on procedures of data elicitation and analysis that focus on intra-individual variability (MacIntyre et al., 2020). These can address novel research questions with time-intensive sampling and data. This method can be applied before and after other treatments and interventions (Figure 31.2), especially in settings where more conventional between-subject research designs may be inappropriate or challenging to implement. The idiodynamic method also has important advantages for investigating patterns of change across time. For example, it can be used to determine how certain learning behavior of interest is generated and can address the relationship between variables over time. The presence of a video functions as a cue to the participant so that complete reliance on memory can be avoided. As the time span between events and their recall increases, memory becomes more prone to a variety of biases that reduce the potential to accurately capture that event. In this sense, then “a short video beats a long memory almost every time” (MacIntyre, 2012, p. 365), thereby reducing—though not eliminating entirely—the potential for bias in data elicitation. One drawback is that the amount of data generated can be enormous and messy, especially when a large sample of participants is used. Furthermore, these data may be
highly idiosyncratic and contextualized, and therefore the researcher needs to consider how to link the findings to theory and practice.

**Experience Sampling Method**

CDST research stresses the need to take two points into account in IDs research: first that individual development is complex and idiosyncratic, and second that it is structurally dynamic, undergoing constant change. The *experience sampling method* was designed to collect such in-situ and time-dense data, especially in intact settings (e.g., at work, in a classroom, within social institutions, or at home with one’s family) where there is some degree of familiarity with a well-defined complex system or with a developmental question of interest. Though it originated in the study of psychological phenomena, the experience sampling method (ESM)\(^2\), with its time-intensive data and analyses, has potential for any research field looking to explore the interaction between individuals, contexts, and processes such as learning, development, or human behavior more generally (Trull & Ebner-Priemer, 2014). This is because it can tap into complex macro behaviors, dynamic micro interactions within a system, and the emergence of new patterns of behavior.

The most characteristic feature of ESM is that it prompts individuals to respond to data elicitation stimuli at regular intervals (e.g., to indicate their affective state or level of alertness throughout the day each time they are signaled to do so). Such repeated measurements result in a type of data that allows researchers to model distinct processes of change for individuals—time is the primary axis that makes the data meaningful and coherent. ESM is, therefore, a means for collecting information about both the context and the content of human life. However, beyond simply cataloging which activities people engage in, where, and with whom, it is the individual differences (i.e., affect, cognitions, and motivations) accompanying such samples of experience that provide a window into the quality of that experience (Mehl & Conner, 2012).

ESM data is elicited by asking individuals to provide systematic self-reports either at regular intervals, such as every ten minutes (e.g., interval-contingent sampling); when they are signaled, such as when a buzzer sounds or participants are pinged (i.e., signal-contingent sampling); or following a particular event of interest, such as after every task ends (i.e., event-contingent sampling) (Shiffman et al., 2008). In the most commonly used technique, individuals provide a response to these stimuli whenever a signaling device prompts them to respond. Such signals can now be fully automated and sent to any one of the latest smart devices or wearables. The signal is a cue to complete the data elicitation measures at that precise moment, and intervals can be scheduled as regularly or infrequently as desired. Some variations include every five or ten minutes in an hour-long class, between every stand-alone activity in a group meeting or project, every two hours over a day-long period, or three times a day over a week. As a domain-specific illustration, students could be asked to respond at each interval to semantic differential scales about their emotional mood (e.g., alert–drowsy, relaxed–anxious, confused–clear, or involved–detached); Likert-style items indicating their opinion about how challenging, important, satisfying, or interesting they found the classroom task (e.g., strongly agree to strongly disagree); or free-response items indicating what they were doing and why, who they were using the L2 with, and what their L2 interaction was about as they were signaled.\(^3\)

One advantage of employing ESM in studies of IDs in SLA is that ESM can free the researcher from the need to be directly involved or physically present during any data elicitation. It also addresses a limitation of observational methods that rely on data from single individuals or single classrooms. ESM allows the tracking of many individuals’ IDs simultaneously, over time, and across situations. Compared to other self-report methods that are retrospective, ESM taps into the ways in which individuals experience, feel about, and report on those learning activities at that moment. For this reason, it adds contextual detail of respondents’ experience of language learning and language use in real time (Feldman Barrett & Barrett, 2001). And, unlike most traditional self-report
techniques, it does not rely on a single assessment moment but gathers repeated measurements across many occasions. By doing this, ESM combines the ecological validity of naturalistic behavioral observation with the non-intrusive nature of diaries and the rigor and precision of psychometric techniques (Bolger & Laurenceau, 2013).

Hektner et al. (2007) propose that there “are few important questions … that cannot benefit from the systematic sampling of experiences” (p. 12). On very large scales of time and context, one of the main foci of ESM relates to individual–environment interactions. At this level of analysis, questions might relate to person-to-person interactions and the role of context in interpersonal dynamics (Hiver & Al-Hoorie, 2020):

- How are particulars of human behavior, functioning, or performance (e.g., learning, communication, and language use) shaped by interactive environmental influences?
- What situational constraints and linkages are there for human behavior and for the social interaction it entails?

With an emphasis on individuals and how IDs contribute to their development, other more micro-level questions are more appropriate (Hiver & Al-Hoorie, 2020):

- How do individuals experience qualitatively distinct situations (e.g., being in an L2 classroom or a study-abroad context), activities (e.g., processing input, expending effort, allocating attention, and interacting communicatively), and outcomes (e.g., success, failure, excitement, interest, and comprehension)?
- In a given setting, what outcomes emerge that are more or less stable across individuals (e.g., learning and negotiation of meaning)?
- How do individuals differ from one another in their experience of similar situations, and how do such experiences and outcomes change through time?
- For particular target-related behavior (e.g., completing a task), what processes of change are part of these experiences?

Of course, ESM studies of IDs in SLA must also grapple with practical implementation issues. In practice, ESM includes dense data points collected longitudinally over a defined time span at the individual level. This quantitative data that ESM produces presents analytic challenges including dealing with missing data, creating strategies to summarize meaningful events within the data stream, and dealing with the effect of autocorrelation in a time series. Another drawback is the question of obtrusiveness—whether the repeated measurement procedure has an excessive influence on learners’ thoughts and behaviors in the moment (Shiffman et al., 2008). For example, stopping to respond may interfere with learners’ concentration or performance, or interfere with whatever the learners’ primary task is. Perhaps the biggest challenge is the heavy demand that these regular responses to the questions or items can impose on respondents. The nature of the data elicitation itself can be a factor that discourages potential participants, lowers overall completion rates, and causes data quality to deteriorate as time passes.

**Social Network Analysis**

A system, especially those in the study of IDs in SLA, is one of many nodes embedded within an interconnected web. This multinode web that includes interconnected systems, their relationships, and processes is called a network (Kadushin, 2012). Social network analysis methods draw heavily on the relational principle (i.e., that everything counts and that everything is connected), with practical consequences for sampling, developing measurements, and handling the resulting data in IDs research.
Social network analysis (SNA) operates on several assumptions (Knoke & Yang, 2020). First, individual systems and their behavior(s) are interdependent, not autonomous. Second, relational links between individual systems are interactive pathways through which systems maintain this interdependence. These stable, but not static, patterns of relational pathways make up the network structure. Third, the network’s structural environment both constrains and provides opportunities for the functioning of individual systems. And fourth, the functioning of individual systems feeds back into the network.

SNA methods make a distinction between relational data and attribute data. More conventional IDs in SLA research rely mainly on attribute data (Figure 31.3). However, SNA methods require relational data in order to express models in terms of relational concepts, information, properties, and processes (Yang et al., 2017). As an example of relational data, consider a network of multilingual friendships (Figure 31.4). The values shown in this dataset represent the number of friends that each individual has in common with one another—as might be shown on a social networking website. This data can be visualized in a sociogram (Figure 31.5) showing how each friend (i.e., system) in the network has ties to others on the whole. Among other things, these links might be language interactions, behavioral influence, group co-membership, material transactions, transfer of information, or personal evaluations of one another.

SNA has a dual focus on both relations and the structural network that results from “individual or group behavior and attitudes” (Carolan, 2014, p. 7), and multiple levels of analysis are possible. For example, a language learner might be part of a network that includes the community of language users, educational institutions and public policymakers, and the relations between all these. SNA methods can examine the behavior and functioning of the systems in the network as well as what this higher-order structure can reveal about the systems within it (Knoke & Yang, 2020). Both levels contain important features that can reveal important structures and processes.

In IDs research, there are several important implications of working with relational data. First, in SNA the unit of analysis is a unitary collective consisting of systems (e.g., individuals, communities, groups, organizations, and institutions) and the links between them (Borgatti et al., 2018). This is the network, and it should always be theoretically or empirically justified. SNA develops this information in a representation of points and paths. The points (aka nodes or actors) represent individuals or systems and their goals or actions. The paths (aka ties or edges) that run between points tie them together and represent the interactional and causal interdependencies they share. For example, Figure 31.6 shows a network with a higher density of cohesive relationships and distributed centrality (i.e., balanced interdependence).

<table>
<thead>
<tr>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>Cases</td>
</tr>
</tbody>
</table>

*Figure 31.3* A Data Matrix for Attribute Data.

<table>
<thead>
<tr>
<th>Cases</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>etc…</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>etc…</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>etc…</td>
</tr>
<tr>
<td>Cases</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td>etc…</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>etc…</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td>etc…</td>
</tr>
</tbody>
</table>

*Figure 31.4* A Relational Data Matrix Showing the Number of Friends Shared by Each Individual.
Next, setting network boundaries for sampling and for data elicitation is an important step with a network model. Boundary specification is more than simply identifying obvious boundaries of a situation and the systems in context. One common approach to boundary specification, the *positional* approach, constructs network nodes based on some a priori principle or criterion (e.g., all those with a specific quality or specific levels of a given ID variable). Others include the *event* approach, based on participation in a given activity (e.g., an experimental task or interaction with peers); the *relational* approach, which reflects social connectedness (e.g., students in a cohort or at a particular level of proficiency); and the *nominal* approach, which is based on the theoretical concerns and choices of the researcher (Kadushin, 2012).
Methods for Complexity Theory

SNA can be applied equally well to research questions at many different levels of analysis. At a descriptive level, working to develop and validate models of a network, we might investigate (Hiver & Al-Hoorie, 2020):

- Which other systems or actors (e.g., cognitive and non-cognitive factors, and L2 instructional conditions) is a particular system related to in a network?
- Which systems or actors are more influential and less influential within a network?
- What patterns or network features emerge from the relations between its systems or actors?
- How do these structures change or remain stable over time in a network?
- What is the nature of the relationships (i.e., unidirectional, symmetric, intermediary, or causal) between these systems in the network?

At a more inferential level that aims to interpret the utility and outcomes of network models for IDs in SLA research, questions include (Hiver & Al-Hoorie, 2020):

- How are networks generated, and how do they evolve and adapt?
- What purpose does, or should, a network serve?
- How can change be introduced in a system?
- What are the implications of two separate networks with similar functioning overall (e.g., similar processes and outcomes) but with key differences in network structure—and vice versa?
- How can temporal changes that impact the functioning and robustness of a network be tracked and forecasted?

SNA methods highlight the ways in which human and social activities (e.g., learning, creating, teaching, and communicating) bring systems into contact with each other. As this contact varies structurally in its value, frequency, duration, and direction, these issues are clearly relevant for IDs in SLA research. It is clear that many dimensions and forms of relational data can be studied using SNA. For instance, SNA has been applied to questions of language development, language processing, language maintenance, and language competence and use (see e.g., a special issue by de Bot & Stoessel, 2002). Its recent applications in the field relate to studies of learners’ interactional goals and emotions, and their willingness to communicate with peers across instructional settings (e.g., Butler & Liu, 2019; Gallagher, 2019; Gallagher & Robins, 2015).

In their study, Gallagher and Robins (2015) examined how interactional networks are constructed in the L2 classroom, and how relational ties affect learners’ willingness to communicate (WTC) in classroom settings with many formal and informal opportunities for L2 contact (e.g., through one-on-one and small group activities and classroom presentations). The researchers collected relational information to compare how networks self-organized between members of the same cultural-linguistic group and among learners with cross-cultural social ties. The researchers hypothesized that learners’ L2 WTC would be socially distributed according to relative positions within a social network, and would differ meaningfully between small-group interactions versus large-group interactions.

They sampled 75 English for Academic Purposes (EAP) students from a non-credit-bearing EAP university course—a relatively well-bounded system of member individuals who interact with each other toward a common goal. Students were asked to generate names of up to ten individuals with whom they had interacted in the last two weeks. To code these relational data, intracultural ties and cross-cultural ties were specified. The researchers also measured respondents’ L2 WTC by asking students to self-report the percentage of time they would choose to initiate L2 communication in various social situations.
The researchers predicted that:

- L2 WTC would be generally associated with greater network activity
- Large-group L2 WTC would be associated with greater status in the network
- Small-group L2 WTC would be similar across network ties

Their overall findings showed that group structure was organized predominately among intracultural ties, and cross-cultural ties did not appear to demonstrate equivalent self-organizing properties. Along with many detailed findings regarding within-network structure and relational ties, their findings illustrate that individuals similar in L2 WTC in one-on-one and small-group settings shared particular intracultural social pockets within the network. They also found that L2 WTC in large-group and presentation settings was distributed according to popularity/status, though this is limited to intracultural ties. Interestingly, they found no support for the notion that participants high in L2 WTC were more active among their network ties in any communication setting.

**Design-Based Research Methods**

*Design-based research* (DBR) refers to a group of methods designed to bridge research and practice in education (Barab & Squire, 2004). DBR used in the context of IDs in SLA research can specify a sequence of actions to deliberately achieve an intended language learning outcome. In this sense, DBR is a method that intentionally searches for evidence related to complex and dynamic questions such as what works for whom, when, in which settings, under what conditions, and why (Anderson & Shattuck, 2012). By employing mixed methods of data elicitation and analysis, DBR aims to better understand the messiness of what transpires in educational settings, all while treating context as a central part of the story.

There are many functions of DBR that make it suitable for IDs research. Many of the research questions that can be investigated through DBR are “big” questions (DBR Collective, 2003) in the sense that they deliberately explore for whom, how, when, and why something does or does not work in applied settings. Below, I offer concrete parallel examples of possible IDs in SLA questions (Hiver & Al-Hoorie, 2020):

- What are the important characteristics of a system and its functioning? (e.g., *What aspects of working memory capacity are most central to real-time L2 speech production?*)
- Why does system A function differently than system B in context Y? (e.g., *Why do learners with mastery-approach goal orientations exert greater effort in classroom settings than other learners?*)
- How can system-level problems be identified and targeted in research settings? (e.g., *How can institutions diagnose and address steadily declining student engagement in L2 classrooms?*)
- What happens to system-level fidelity when unanticipated processes and outcomes are encountered in time and context. (e.g., *How does a learner’s production of L2 phonological features augment or deteriorate over time as anxiety is reduced or increased?*)
- How do processes at different timescales improve our understanding of systems in context? (e.g., *What evidence can examining learners’ cognitive and emotional responses to corrective feedback in real time provide to help us understand the mechanisms underlying the effectiveness of corrective feedback?*)
- How can initial conditions feed into system change iteratively? (e.g., *How do initial self-competence beliefs affect learners’ sustained task performance, and how do they change as a result of ongoing performance?*)
- Under what conditions and settings can system change be enacted effectively? (e.g., *What combinations of IDs and contextual conditions are needed to spur learners’ development of explicit L2 knowledge?*)
DBR takes place in authentic (i.e., non-laboratory) environments and studies phenomena within the contexts they naturally occur (Tabak, 2004). A first step when embarking on a DBR study is to clarify the specific purpose of the study, the problem being addressed, or the focal question. This is when researchers should articulate any logical assumptions—similar to the Bayesian logic of “priors”—about the starting points, pathways of development, and prospective endpoints or goals of the system under investigation. The next step is to formulate a design that embodies these assumptions in practical settings. In other words, the researchers mark out a sequence of actions intended to deliberately achieve an intended outcome (Barab, 2014).

As yet, there are no existing applications of this method in IDs research in SLA. As a hypothetical illustration, the DBR objective might be an issue such as reversing the rapid decline in language enrollment and participation in language programs that affects some contexts. DBR research should be conceptualized as a process of generating desired change in a system, and for the present example this may mean finding ways to increase the perceived value, investment of effort, and support for student learning in this specific domain. With this purpose firmly in mind, the DBR team can begin to draw on actual language student capabilities, ID profiles, instructional histories, analysis of needs, current practices and policy initiatives, and other resources as the starting point they intend to build on (Collins et al., 2004).

Once this starting point is more or less clear, the DBR team must also achieve consensus about the targets or prospective endpoints they intend to reach as they generate desired change in the system (Penuel et al., 2011). What, in other words, constitutes success? This is where the DBR team should refer to previous IDs in SLA studies that provide evidence about the necessary conditions for optimal language development, while also benchmarking the level and type of change the system would be expected to undergo. The researchers would then use this opportunity to examine the conditions needed to support various students’ classroom learning (e.g., what combination of IDs seem to be predictive of sustained participation in language learning or, conversely, of disinvestment and non-participation?). They should also plan the processes they intend to test and revise iteratively to raise personal participation in language learning.

Once a study has been planned and structured, the research team generates multiple forms of data by integrating quantitative and qualitative data collection and analysis (Anderson & Shattuck, 2012). DBR expects researchers to generate an extensive and longitudinal dataset by collecting and curating a range of data sources. In practice, this might include integrating data sources such as samples of ongoing student work, student learning products, task performance measures, responses to interviews and elicited prompts, language production, notes from observation protocols, interactional patterns and classroom discourse, and ratings on other more quantitative measurements (e.g., surveys or assessments) (Barab, 2014). This is done to document system change and to support a systematic and complex analysis of the phenomenon under investigation.

The design of DBR studies is iterative in the sense that data feeds into subsequent cycles of design and implementation. However, the longitudinal dataset generated from a DBR study is usually analyzed retrospectively in order to provide a situated and dynamic account of the processes and outcomes under investigation and to ensure that resulting claims are valid and grounded in evidence (Tabak, 2004). DBR is built on a kind of applied pragmatism that matches the method(s) of data collection and analysis closely with the topic or question under investigation. By operating
from this problem–solution approach, DBR typically involves mixed methods using a variety of research tools and techniques (Penuel et al., 2011).

**Contributions to ID Research**

In IDs research, like many other domains of SLA, a researcher’s disciplinary orientation drives topic selection and question generation. The CDST family of methodological tools has provided the field with new insights regarding the focus of IDs in SLA research, the type of research questions appropriate for exploration, the types of evidence that can be collected, and the types of results and answers that are achieved. The questions in Table 31.1, for instance, show what this might look like from a conventional methodological paradigm and how this changes when employing CDST methods (see also MacIntyre et al., 2017). Much of the difference comes from the cross-sectional versus longitudinal differences inherent to these different paradigms. There is also the possibility in CDST to structure questions around the relational principle (e.g., how are learners’ task-specific beliefs and their emotional engagement related to their interaction behaviors in the L2 classroom?); around the adaptive principle (e.g., how does an engaged L2 learner’s task performance throughout a semester predict their learning gains?); or a combination of both (e.g., how does the interaction mode [face-to face vs. synchronous computer-mediated communication] and task type affect learners’ classroom engagement over the course of a semester?).

All theories, if they are to avoid becoming passing academic fads or bandwagons, must contribute something of substance that is new and worthwhile—something that pushes the field forward (Larsen-Freeman, 2017). CDST has indeed made such contributions in the study of IDs in SLA, pushing further and faster even than in related fields such as IDs in educational psychology. Respective chapters in this handbook show that CDST perspectives and approaches have permeated many areas of IDs research. Some of the most prominent strands of IDs in SLA in which CDST experienced rapid uptake include motivation (Dörnyei et al., 2015; Papi & Hiver, 2020), emotions (e.g., anxiety and enjoyment) (Gregersen et al., 2014), language learning strategies (Oxford, 2017), willingness to communicate (MacIntyre & Legatto, 2011), and the self (Mercer, 2011). In these and other areas a situated, complex, and dynamic way of seeing the object of study has become the state of the art—as has using complex and dynamic templates and designs to investigate those IDs. Ongoing work has expanded the reach of CDST by applying it to study newer IDs in SLA such as working memory (Jackson, 2020), group dynamics (Poupore, 2018), and

<table>
<thead>
<tr>
<th><strong>Table 31.1 Standard Research Questions Compared with CDST Research Questions</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard Research Questions</strong></td>
</tr>
<tr>
<td>Are anxiety and motivation associated with learners’ WTC?</td>
</tr>
<tr>
<td>What effect does TBL have on learners’ engagement in the L2 classroom?</td>
</tr>
<tr>
<td>Do students rely on working memory capacity more when responding to output-prompting CF or to input-providing CF?</td>
</tr>
<tr>
<td>Is higher meta-cognitive capacity predictive of syntactic and lexical complexity in learners’ L2 writing?</td>
</tr>
</tbody>
</table>

**Note.** CAF = complexity, accuracy, and fluency; CF = corrective feedback; TBL = task-based learning; WTC = willingness to communicate.
engagement (Hiver et al., 2021b). Current work has also widened in scope to explore the value of CDST for language teacher IDs.

Another contribution of CDST has been the introduction of new designs and innovating with existing methods (see examples in earlier sections). Conventionally, the debate around the merits of qualitative versus quantitative research is a dominant methodological area of contention. With CDST methods, however, this debate has been superseded by concern for the merits of individual-versus group-level (i.e., high or low $N$) designs and analyses, and the timescale or number of occasions (i.e., high or low $T$) appropriate for these designs. CDST research attempts to address many purposes, including both exploratory and falsificatory/confirmatory aims (Hiver & Al-Hoorie, 2020; Hiver et al., 2021a). Once the aim of the research is established, this can inform whether the unit of analysis should be the individual or the group (or both), what the timescale of investigation is, and whether the research should be designed around methods that are qualitative or quantitative (or both). More work remains to be done using these integrative designs, and this forward-facing agenda is what I now turn to in the final section of the chapter.

**Future Directions**

While some important advances in understanding have happened, these are early days for CDST research methods in IDs. CDST methods will continue to be developed and refined as the field finds new ways to employ them to study the complex interconnectedness, dynamism on multiple timescales, and nonlinear development of IDs. As Molenaar and Campbell (2009) propose, psychological processes are varied and complex, and this heterogeneity can go undetected if research designs are not sensitive to the complex and dynamic aspects of IDs.

What might the future of applying CDST methods hold? CDST expands the possibilities for future IDs in SLA research by encouraging design decisions at several distinct levels (Figure 31.7): aim, unit of analysis, and method. The arrows between these options indicate that CDST research can take both features into account at each level in order to make research maximally informative. Starting with the aim, the ID research might be exploratory or may attempt to test certain understandings or expectations (e.g., observationally or (quasi-)experimentally). Although the complex social world does not lend itself to universals that can be applied across all settings and populations, unlike what some may assume about CDST, it is still possible to form some generalizations and probabilistic predictions by comparison to other similar systems, under similar conditions and contexts, and with similar outcomes (Hiver & Al-Hoorie, 2020; Hiver et al., 2021a). Consequently, when using CDST research tools for IDs in SLA there is no reason to shy away from making predictions and then subjecting these predictions to empirical tests. Integrative CDST designs should be characterized by a flexible switch between these two aims, and adopting the exploratory–falsificatory approach proposed here can radically reorient researchers and their

![Figure 31.7 Integrative Framework for CDST Research Designs.](adapted from Hiver & Al-Hoorie, 2020)
aims, making them actively seek negative, disconfirming results rather than exclusively celebrating positive ones.

The next level, unit of analysis, has to do with whether the level of granularity in a CDST study is at the individual or the group level. Here, some have contrasted an idiographic, person-centered, individual-level approach with a nomothetic, variable-centered, or group-level approach (see also Larsen-Freeman & Cameron, 2008). Both individual- and group-based designs feature in CDST research. The former is focused on finding what is unique in each individual, while the latter looks for generalizations that apply across many individuals. This unit also applies to timescales and processes of change in which the nomothetic approach emphasizes inter-individual variability and the overall mean trajectory of all cases, whereas the idiographic approach emphasizes intra-individual variability and the unique developmental trajectories of each individual (Verspoor et al., 2011).

Group-based research tends to be the norm in IDs in SLA research, though individual-based research may be more compatible with the assumptions of CDST (Lowie, 2017). Individual-based research allows the researcher to hold a close lens to development and change without averaging away individual idiosyncrasies (Molenaar & Campbell, 2009). Of course, the notion that one level of data is superior to another may be a point of (mis)interpretation considering that a complex system is not restricted to representing an individual. If a group is the system chosen as the unit of analysis for IDs research, or if it is any higher-level system other than an individual, then it may be that group-level data are more relevant for that particular study. Ideally, an integrative design should attempt to draw from both of these approaches as the individual-level and group-level of analysis are complementary in CDST research methods. Individual-based research designs allow meticulous analyses of single cases while group-based results uncover broader tendencies and how these results vary in the population (Hiver et al., 2021a).

The final choice is the method. CDST methods deal primarily with longitudinal data if they operate using the adaptive principle, but may also apply to cross-sectional data if concerned with the relational principle. Longitudinal data and designs are usually more CDST compatible because these focus on the outcomes or patterns that are reached at different points in time as well as the mechanisms that explain how an outcome is reached. Additionally, it is nearly impossible to study change and development (the adaptive principle) without also accounting for context and interconnectedness (the relational principle). As the templates and methods outlined previously show, integrative CDST designs can draw from both qualitative and quantitative methods to advance knowledge in a particular area of IDs in SLA. Both methods should support IDs in SLA research in a unified manner.

Many specialists in IDs would acknowledge that the issues they are tackling are fundamentally complex, broad, and systemic. If they frame and address an issue or topic from a narrow viewpoint, they may focus only on a single aspect of a multifaceted problem (Larsen-Freeman, 2019). The current limitations of what the IDs field and its eminent research minds have been able to address may be the result of a research paradigm that isolates and segments IDs and thus fails to address the complexity of the field. CDST research acts as a strong antidote to this. CDST research encourages the intelligent use and integrating of quantitative and qualitative methods appropriate for the phenomenon under investigation, and the future will undoubtedly see increased methodological integration that is more pragmatic and less constrained by researcher ideology.

Notes

1 Note that experimental research is not limited to those that rely on two-wave designs, and the CDST research ethos does not rule out experimentation and falsification.

2 Another widely used term is “ecological momentary assessment” (EMA). EMA originated in the field of behavioral medicine and tends to include objective physiological measurements related to health (e.g., pulse and blood pressure). Because the experience sampling method (ESM) originated in psychology, it is the term I adopt in this chapter.
3 This is reminiscent of the “immediate recall” method in feedback research which aims to measure learners’ noticing of feedback in flight.

4 Autocorrelation is the correlation of a data point or observation in a time series with other delayed observations as a function of the time lag between them.

5 While DBR may strike some as similar to action research or practitioner research, it tends to be much larger in scope and involves mixed methods of data collection and analysis.

6 While many group-based designs are also cross-sectional, these two terms should not be conflated. Cross-sectional research designs examine a sample of individuals at a particular point in time, and their primary concern is not change or growth over time. Group-based designs need not be cross-sectional in nature.

References


