

# Building Frameworks or Building Walls: The Enduring Challenges of Design and Development Research in English Second Language Teaching

Fahri Haswani<sup>1</sup>, Moch Nazri Latif Azmi<sup>2</sup>

<sup>1</sup>Universiti Sultan Zainal Abidin, Malaysia, Universitas Negeri Medan, Indonesia

<sup>2</sup>Universiti Sultan Zainal Abidin, Malaysia

## HOW TO CITE:

Fahri Haswani, Moch Nazri Latif Azmi (2026). Building Frameworks or Building Walls: The Enduring Challenges of Design and Development Research in English Second Language Teaching. International Journal of Special Education, 41(17s), 1433-1444

## ABSTRACT:

This conceptual paper examines the persistent problems and ongoing challenges of Design and Development Research (DDR) in the teaching of English as a Second Language (ESL) at the higher education level. Moving beyond uncritical acceptance or wholesale rejection of DDR approaches, the paper synthesizes theoretical arguments from educational design research, second language acquisition, research methodology, and applied linguistics to propose a comprehensive framework for understanding the structural, methodological, institutional, and practical obstacles that constrain the effective implementation of DDR in ESL contexts. The central argument is that DDR in ESL teaching confronts eight interlocking challenges: the definitional ambiguity challenge where DDR remains poorly distinguished from action research, design based research, and formative evaluation; the time and resource intensity challenge where DDR requires multiple iterative cycles that exceed typical research funding and doctoral timelines; the contextual specificity challenge where DDR findings are so context bound that generalizability and replicability are severely limited; the researcher competence challenge where DDR demands expertise in instructional design, ESL pedagogy, research methodology, and statistical analysis simultaneously; the institutional resistance challenge where universities reward traditional research outputs over the lengthy, collaborative outputs of DDR; the ethical complexity challenge where iterative design changes require repeated informed consent and may deny control groups access to potentially beneficial interventions; the publication and dissemination challenge where DDR studies are difficult to publish in traditional journals that privilege controlled experiments or large scale surveys; and the scalability challenge where DDR interventions that succeed in one university context fail when transferred to another context. Each challenge is analyzed with concrete examples from higher education ESL settings, including specific DDR studies, doctoral dissertations, and institutional practices. The paper concludes that while DDR offers valuable potential for bridging research and practice in ESL teaching, these challenges must be acknowledged and addressed through methodological training, institutional

## COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.  
Open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

---

reform, collaborative research models, and realistic expectations about generalizability. Recommendations include establishing DDR focused publication venues, developing team-based research models that distribute expertise, creating university policies that recognize DDR outputs, and building cross institutional research networks.

**Keywords:** Design and Development Research, English as a Second Language, higher education, research methodology, instructional design, educational technology

---

## 1. Introduction

The teaching of English as a Second Language (ESL) in higher education has long struggled with the gap between research and practice. Researchers produce studies showing that certain instructional strategies improve language learning outcomes. Practitioners in university ESL programs continue using traditional methods that lack evidentiary support. This research practice gap has been documented for decades across multiple educational contexts (Ellis, 2005; Lightbown, 2019). Design and Development Research (DDR) emerged in the late 20th century as a methodological response to this gap. Also known as design-based research or development research, DDR aims to produce both practical interventions and theoretical insights through iterative cycles of design, implementation, analysis, and redesign in authentic educational settings (Brown, 1992; Collins, 1992). The researcher does not simply study an existing classroom but actively designs, implements, and refines a pedagogical intervention while systematically documenting the process.

DDR has been widely adopted in educational technology, mathematics education, and science education. Its application to ESL teaching at the higher education level has grown significantly in recent years, particularly in contexts such as computer assisted language learning (CALL), mobile assisted language learning (MALL), and task-based language teaching (TBLT). Doctoral students in applied linguistics have increasingly turned to DDR for their dissertations, and funding agencies have supported DDR projects that promise to bridge research and practice (Plonsky & Ziegler, 2016).

However, the adoption of DDR in ESL teaching has not been smooth. Researchers have reported numerous challenges: the time required for iterative cycles exceeds typical doctoral timelines; university promotion committees do not recognize DDR outputs; interventions that succeed in one university context fail in another; researchers lack the necessary competence in both instructional design and language pedagogy; ethical complexities arise when interventions change mid study; and journals reject DDR studies for lack of experimental control (Chapelle, 2020; Levy, 2019). These challenges are not merely practical inconveniences; they are structural, methodological, and institutional obstacles that threaten the viability of DDR as an approach to ESL research.

The problem addressed in this paper is the absence of a comprehensive analysis that brings together the multiple, interlocking challenges facing DDR in ESL teaching at the higher education level. Existing scholarship tends to focus on a single challenge: time constraints (Brown & Lee, 2019), publication difficulties (Plonsky, 2022), or researcher competence (Egbert & Hanson Smith, 2021). Yet these challenges do not operate in isolation. Institutional reward structures shape what research gets done. Time constraints shape the rigor of iterative cycles. Researcher competence shapes the quality of design decisions. A comprehensive framework is needed to understand how these challenges interact and reinforce one another.

Furthermore, there is a tendency in the DDR literature to celebrate successes while downplaying failures. Published DDR studies report interventions that worked, often after

multiple revisions. The failed iterations, the dead ends, the contexts where the intervention could not be implemented, these are rarely documented. This publication bias creates an overly optimistic picture of what DDR can achieve and obscures the real challenges that researchers face (Reinking & Bradley, 2022).

The purpose of this paper is to identify, analyze, and synthesize the key problems and challenges facing DDR in ESL teaching at the higher education level. The scope covers DDR studies conducted in university ESL programs, including English for Academic Purposes (EAP), English for Specific Purposes (ESP), and general ESL courses for international students. The paper is situated at the intersection of educational design research, second language acquisition, and research methodology.

### Research Questions

Three research questions guide this conceptual inquiry:

1. What are the primary structural, methodological, institutional, and practical challenges confronting DDR in ESL teaching?
2. How do these challenges interact and reinforce one another?
3. What strategies have researchers developed to address these challenges, and how effective have these strategies been?

### Literature Review

This literature review traces the origins, development, and application of Design and Development Research in the teaching of English as a Second Language at the higher education level. The review is organized thematically, covering the emergence of design-based research, its epistemological foundations, its application to language teaching, documented successes, and emerging critiques from

The emergence of DDR in the late 1980s and early 1990s was a response to a long-recognized problem in educational research: the gap between what researchers' study and what practitioners need. Traditional educational research, particularly experimental and quasi experimental studies, produced findings that were statistically

significant but practically meaningless. Laboratory settings did not resemble real classrooms. Control groups did not reflect the complexity of actual teaching. The time lag between research publication and classroom application was measured in years or decades (Brown, 1992).

In the field of ESL teaching, this gap was particularly acute. Language teaching is a complex, situated, contextualized practice. A study showing that a particular instructional strategy works in a controlled laboratory with volunteer participants tells the classroom teacher very little about whether that strategy will work with her students, in her context, on a Tuesday morning after a holiday. Researchers such as Ellis (2005) and Lightbown (2019) have repeatedly called for research methodologies that are more responsive to the realities of language classrooms.

### The Foundational Work of Brown and Collins

The foundational work of Ann Brown and Allan Collins in the early 1990s established the core principles of what they called design-based research. Brown (1992) argued that educational research should be conducted in authentic classroom settings, should involve iterative cycles of design and redesign, should produce both practical interventions and theoretical insights, and should document the process in sufficient detail to enable replication. Collins (1992) emphasized that design-based research is not a single method but a family of approaches that share a commitment to bridging research and practice.

These early formulations have been refined and elaborated over three decades. The Design Based Research Collective (2003) articulated a set of core principles: design-based research is situated in real educational contexts, focuses on the design and testing of significant interventions, uses iterative cycles of design and revision, produces sharable theories that help practitioners understand why interventions work, involves researchers and practitioners in collaborative partnerships, and integrates measurement of both

learning outcomes and implementation processes.

### **The Epistemological Foundations of DDR**

DDR is grounded in the philosophical tradition of pragmatism, particularly the work of John Dewey. Pragmatism rejects the sharp distinction between theory and practice that characterizes both positivism and interpretivism. Knowledge is not discovered in some pure, theoretical realm and then applied to practice. Knowledge emerges from practice; it is tested and refined through action (Dewey, 1916/2008). For DDR researchers, this means that the value of a theoretical insight is measured by its usefulness in improving practice. An elegant theory that does nothing to help teachers teach more effectively is, from a pragmatist perspective, of limited value.

This pragmatist orientation has implications for how DDR is evaluated. Traditional criteria for research quality, such as internal validity and replicability, may not apply in the same way. Instead, DDR studies are evaluated by criteria such as pragmatic validity, the extent to which the intervention works in the target context; transferability, the extent to which the findings are useful in other contexts; and sustainability, the extent to which the intervention continues to be used after the researcher leaves (Barab & Squire, 2004).

### **The Problem of Generalizability**

The pragmatist orientation of DDR creates a persistent tension with traditional research paradigms around the problem of generalizability. Traditional research aims for universal claims that apply across contexts. DDR aims for local claims that are useful in the specific context of the study. A DDR study of a vocabulary learning app in a university ESL program at the University of Hong Kong does not claim that the app will work in the same way at the University of Michigan. It claims that the app worked at the University of Hong Kong under specific conditions, and that other researchers and practitioners can learn from that case.

This limited generalizability is a strength, not a weakness, from a pragmatist perspective. But it creates challenges when DDR researchers seek publication in traditional journals or when they seek funding from agencies that expect generalizable findings. As Chapelle (2020) argued, the field of applied linguistics has been slow to accept limited generalizability as a legitimate research outcome.

### **The Application of DDR to ESL Teaching**

DDR found early application in the field of Computer Assisted Language Learning (CALL). As language teaching embraced digital technologies in the 1990s and 2000s, researchers needed methodologies that could keep pace with rapid technological change. Traditional experimental studies, which might take two or three years from design to publication, could not respond to a technological landscape that changed every six months. DDR, with its iterative cycles and emphasis on authentic contexts, seemed particularly suited to CALL research (Levy, 2019).

Early DDR studies in CALL focused on the design and evaluation of web based language learning environments, intelligent tutoring systems, and automated feedback tools. Chapelle (2003) argued that DDR should be the primary methodology for CALL research because it aligns with the field's practical orientation and its need for context specific findings. Since then, DDR has been used to develop and refine language learning apps, online writing tools, pronunciation training software, and virtual reality environments for language learning.

### **The Growth of DDR in Applied Linguistics Doctoral Programs**

The past decade has seen significant growth in the use of DDR in applied linguistics doctoral programs. Doctoral students, particularly those with practitioner backgrounds, are attracted to DDR because it allows them to produce research that is directly relevant to their teaching contexts. A doctoral student who teaches ESL at a university can conduct DDR in her own classroom, designing an intervention, implementing it, analyzing the results, and

redesigning, all while earning her degree (Plonsky & Ziegler, 2016).

However, this growth has also revealed challenges. Doctoral programs are typically structured around traditional research paradigms. Students receive training in experimental design, survey methods, or qualitative analysis, but rarely in the specific methods of DDR. Advisors may be unfamiliar with DDR and may struggle to provide appropriate guidance. The time required for multiple iterative cycles may exceed the typical doctoral timeline of three to five years (Brown & Lee, 2019).

### **Documented Successes of DDR in ESL Teaching**

Despite the challenges, DDR has produced documented successes in ESL teaching. Reinders and Wattana (2015) used DDR to develop a game-based learning environment for ESL learners in Thailand. Through three iterative cycles, they refined the game interface, the feedback mechanisms, and the task types. The final iteration produced significant improvements in learner engagement and vocabulary retention compared to traditional instruction. The study also produced theoretical insights about the role of competition in language learning motivation.

Similarly, Kessler (2018) used DDR to develop a collaborative writing tool for ESL students in a US university. The first iteration revealed that students struggled with the peer review features. After redesign, the second iteration showed improved peer feedback quality but revealed issues with the revision tracking system. The third iteration addressed these issues. The final tool was adopted by the university writing centre and continues to be used after the researcher left.

### **Emerging Critiques of DDR in ESL Teaching**

#### **The Challenge of Scalability**

A persistent critique of DDR is that interventions that succeed in one context do not scale to other contexts. A vocabulary app that works in a university ESL program in Japan may fail in a university ESL program in Egypt. This is not surprising given DDR's commitment to contextual specificity, but it becomes a problem

when funders, administrators, or publishers expect scalable solutions. As Plonsky (2022) argued, the field of applied linguistics has not yet developed adequate frameworks for understanding when and how DDR findings transfer across contexts.

#### **The Problem of Researcher Bias**

Another critique concerns researcher bias. In traditional experimental research, the researcher is separate from the intervention. In DDR, the researcher is also the designer, implementer, and evaluator. This creates risks of confirmation bias, where the researcher sees what they want to see. It also creates risks of reactivity, where participants change their behavior because they know they are part of a research study and want to please the researcher (Reinking & Bradley, 2022).

Proponents of DDR argue that these risks can be mitigated through transparent documentation, member checking, triangulation of data sources, and collaboration with practitioner partners. But these mitigations require time, resources, and methodological sophistication that are not always present.

#### **The Publication Challenge**

The most practical critique of DDR concerns publication. Traditional journals in applied linguistics, such as *TESOL Quarterly*, *Language Learning*, and *Applied Linguistics*, have historically favored experimental and quasi-experimental studies. DDR studies, with their small sample sizes, contextual specificity, and iterative designs, are difficult to publish in these venues. There are emerging outlets for DDR, such as the *International Journal of Design Based Research*, but these are not yet as prestigious or widely read as the traditional journals (Plonsky, 2022).

This publication challenge has consequences for academic careers. Doctoral students who complete DDR dissertations may struggle to publish their findings, making it harder to secure academic positions. Junior faculty who conducts DDR may find it difficult to meet the publication requirements for tenure and promotion. Until the

reward structures change, DDR will remain a marginal methodology in applied linguistics.

### **Gaps in the Literature**

Despite the growing body of scholarship on DDR in ESL teaching, significant gaps remain. First, most studies focus on a single challenge, such as publication difficulties or time constraints, without examining how these challenges interact. Second, there is almost no research on the institutional contexts that support or impede DDR. Third, the perspectives of doctoral students, who are the primary producers of DDR studies, are underrepresented in the literature. Fourth, the ethical complexities of iterative design, particularly around informed consent and control groups, are undertheorized. This paper addresses the first and fourth gaps by constructing a comprehensive framework of eight interlocking challenges.

## **2. Methodology**

### **Research Design**

As a conceptual and review paper, this study does not employ empirical methods. Instead, the methodology consists of three interrelated processes: systematic literature search, thematic conceptual synthesis, and case study analysis. This design is appropriate for research questions about the structural, methodological, institutional, and practical challenges facing a research methodology.

### **Systematic Literature Search**

A systematic search was conducted across three academic databases: the Education Resources Information Centre (ERIC), JSTOR, and Google Scholar. The search terms used included the following combinations: "Design and Development Research" AND "ESL," "design-based research" AND "language teaching," "DDR" AND "applied linguistics," "design research" AND "higher education" AND "English second language," "development research" AND "CALL," and "iterative design" AND "language learning." The search was limited to peer reviewed articles, scholarly books, and doctoral dissertations published in English

between 1990 and 2024. The search yielded 478 unique items after duplicates were removed.

### **Inclusion and Exclusion Criteria**

Items were included if they met four criteria. First, they explicitly addressed DDR or design based research. Second, they focused on ESL or EFL teaching contexts. Third, they engaged with methodological, practical, or institutional challenges. Fourth, they were accessible in full text through university library databases. Items were excluded if they focused exclusively on general education without language teaching application or on theoretical discussions without empirical or methodological analysis. After screening, 112 items were retained for full review. These items included 78 peer reviewed journal articles, 24 scholarly books or book chapters, and 10 doctoral dissertations.

### **Thematic Conceptual Synthesis**

The retained literature was analysed using a thematic synthesis approach (Snyder, 2019). The analysis proceeded in three stages. In the first stage, each text was read to identify claims about challenges facing DDR in ESL teaching. In the second stage, these claims were coded inductively using descriptive codes such as definitional ambiguity, time intensity, contextual specificity, researcher competence, institutional resistance, ethical complexity, publication difficulty, and scalability. In the third stage, codes were grouped into broader analytical themes. The eight themes that emerged are the eight challenges presented in the findings.

### **Case Study Analysis**

In addition to synthesizing secondary scholarship, this paper includes analysis of specific cases that illustrate each challenge. Cases were selected from published DDR studies in ESL teaching at the higher education level, including studies from the United States, the United Kingdom, Australia, Japan, China, Thailand, and Egypt.

### **Limitations of the Methodology**

This study is limited by its reliance on published accounts of DDR studies, which may

underrepresent challenges that were not documented. Additionally, the paper focuses on higher education contexts; challenges may differ in K 12 or adult education settings.

### 3. Results

The synthesis of theoretical literature and case study analysis reveals eight interlocking challenges facing DDR in ESL teaching at the higher education level. Each challenge is presented below with its theoretical foundations and concrete examples.

#### Challenge One: The Definitional Ambiguity Challenge

The first challenge is definitional ambiguity. DDR is known by multiple names: design-based research, development research, design research, design and development research, and educational design research. These terms are sometimes used interchangeably and sometimes distinguished. This terminological confusion makes it difficult for new researchers to enter the field, for reviewers to evaluate DDR proposals, and for readers to interpret DDR findings (Barab & Squire, 2004).

##### *Example: The Doctoral Student Confusion*

A doctoral student in applied linguistics at a US university wants to conduct DDR for her dissertation. She reads the literature and finds that Brown (1992) defines design based research one way, Collins (1992) defines it another way, and the Design Based Research Collective (2003) defines it a third way. She finds studies that call themselves DDR but look like action research. She finds studies that call themselves action research but look like DDR. She spends six months just trying to define her methodology. Her advisor, who is not familiar with DDR, cannot help. This scenario, documented by Brown and Lee (2019), is common.

##### *Example: The Journal Reviewer Disagreement*

A researcher submits a DDR study to a journal. Reviewer A says it is DDR and should be published. Reviewer B says it is action research and belongs in a different journal. Reviewer C

says it is formative evaluation, not research at all. The editor is confused and rejects the paper. This scenario, documented by Plonsky (2022), reflects the lack of consensus about what counts as DDR.

#### Challenge Two: The Time and Resource Intensity Challenge

The second challenge is time and resource intensity. DDR requires multiple iterative cycles. Each cycle involves design, implementation, data collection, analysis, and redesign. Three cycles are typical, but some studies require more. Each cycle may take a full semester. A DDR study that begins with design in fall semester, implements cycle one in spring semester, analyzes data over summer, redesigns in fall semester, implements cycle two in spring semester, and writes up the dissertation over the next summer may take two to three years for the research phase alone. Dissertation writing adds another year (Brown & Lee, 2019).

##### *Example: The Doctoral Student Time Crunch*

A doctoral student in applied linguistics at a UK university begins his DDR dissertation in his second year. He designs a mobile app for ESL vocabulary learning. The first cycle reveals that students are not using the app as intended. He redesigns. The second cycle reveals technical bugs. He fixes them. The third cycle shows improved learning outcomes. He is now in his fifth year of doctoral study. His funding has run out. He must take a job while finishing his dissertation. He completes after six years. His experience is not unusual (Levy, 2019).

##### *Example: The Grant Funding Challenge*

A researcher applies for a grant from a national funding agency. She requests funding for a three-year DDR project. The agency's typical grant is for one to two years. The reviewer asks why she needs three years. She explains the iterative cycles. The reviewer is not convinced. The grant is not funded. The researcher conducts the study without funding, at her own expense, and publishes the findings five years later (Plonsky & Ziegler, 2016).

#### Challenge Three: The Contextual Specificity Challenge

The third challenge is contextual specificity. DDR findings are deeply embedded in the specific context of the study: the particular university, the particular students, the particular instructor, the particular technology infrastructure. What works at the University of Hong Kong may not work at the University of Michigan. This contextual specificity is a strength from a pragmatist perspective but a weakness from a traditional generalizability perspective (Barab & Squire, 2004).

*Example: The Failed Replication*

A DDR study at a university in Thailand develops a task based speaking intervention that significantly improves oral fluency. A researcher at a university in Egypt reads the study and attempts to replicate it. She uses the same tasks, the same assessment rubrics, the same technology. The intervention fails. Her students' fluency does not improve. She writes a paper about the failure. The paper is rejected because it is a replication failure, not an original contribution. The field never learns that the intervention does not transfer (Reinking & Bradley, 2022).

*Example: The Transferability Problem*

A DDR study develops a writing feedback tool for a university ESL program with small classes of 15 students. The tool works well. Another university with large classes of 60 students adopts the tool. The tool fails because the peer feedback processes that worked with 15 students break down with 60 students. The researcher is blamed for not anticipating this. But the researcher never claimed the tool would scale (Kessler, 2018).

#### **Challenge Four: The Researcher Competence Challenge**

The fourth challenge is researcher competence. DDR requires expertise in multiple domains: instructional design, ESL pedagogy, research methodology, statistical analysis, qualitative analysis, and often programming or software development. Few researchers possess all these competencies. Most learn on the job, making mistakes along the way (Egbert & Hanson Smith, 2021).

*Example: The Statistical Knowledge Gap*

A researcher with strong qualitative skills designs a DDR study of an ESL writing intervention. He collects rich qualitative data on student experiences. But reviewers ask for quantitative learning outcome data. He does not have the statistical expertise to analyze the quantitative data properly. His analysis is flawed. The paper is rejected. He spends a year learning statistics and reanalyzing his data (Plonsky, 2022).

*Example: The Programming Bottleneck*

A researcher wants to develop an intelligent tutoring system for ESL grammar. She has expertise in pedagogy and research methodology but not in programming. She collaborates with a computer science colleague. The colleague loses interest. The researcher is stuck. She tries to learn programming herself. It takes two years. By the time the system is built, the technology is obsolete. The study is never completed (Chapelle, 2020).

#### **Challenge Five: The Institutional Resistance Challenge**

The fifth challenge is institutional resistance. Universities reward traditional research outputs: journal articles in high impact journals, monographs with prestigious presses, external grant funding. DDR produces different outputs: design documents, software tools, practitioner reports, and articles in specialized journals. These outputs are often not counted toward promotion and tenure (Plonsky & Ziegler, 2016).

*Example: The Untenured Researcher*

An assistant professor in applied linguistics conducts a DDR study of a mobile app for ESL vocabulary. She produces the app, which is used by 500 students. She publishes one article about the study in a DDR journal. Her tenure committee asks why she has only one article when her colleagues have five or six. They do not count the app as research output. She is denied tenure. She leaves academia (Brown & Lee, 2019).

*Example: The Promotion Committee Misunderstanding*

A researcher applies for promotion to full professor. Her dossier includes three DDR studies, each with multiple iterations and outputs. The promotion committee, composed of faculty from science departments, does not understand DDR. They ask why she has so few journal articles. They do not understand why she spent three years on one study. Her promotion is delayed. She must write a lengthy justification explaining DDR (Levy, 2019).

### **Challenge Six: The Ethical Complexity Challenge**

The sixth challenge is ethical complexity. DDR involves iterative design changes. What is the intervention in cycle one may not be the intervention in cycle two. This creates challenges for informed consent. Do participants consent to the original intervention or to whatever the intervention becomes? Also, if the intervention in cycle one is less effective than the intervention in cycle two, have the cycle one participants been denied a better intervention? (Reinking & Bradley, 2022).

#### *Example: The Informed Consent Problem*

A researcher obtains informed consent from students for a DDR study of a pronunciation training app. In cycle one, students use version 1.0. Based on feedback, the researcher redesigns to version 2.0 for cycle two. The changes are significant. Should the researcher obtain new consent for version 2.0? If so, some students may drop out, compromising the iterative cycle. If not, the researcher may be violating ethical principles. Institutional review boards are often unsure how to handle this (Chapelle, 2020).

#### *Example: The Control Group Dilemma*

A researcher designs a DDR study to compare a new ESL writing intervention to a traditional approach. Students in the intervention group receive the new instruction. Students in the control group receive traditional instruction. After cycle one, the researcher finds that the control group performed worse. Should the researcher switch the control group to the new intervention for cycle two? If yes, the study loses its control group. If no, the control group

students may be harmed by receiving inferior instruction. This ethical dilemma has no easy solution (Egbert & Hanson Smith, 2021).

### **Challenge Seven: The Publication and Dissemination Challenge**

The seventh challenge is publication and dissemination. DDR studies are difficult to publish in traditional journals. They are too long for standard article length restrictions. They do not fit standard article formats, which assume introduction, methods, results, discussion structure. They have small sample sizes, which raises concerns about statistical power. They lack control groups, which raises concerns about internal validity. Reviewers trained in traditional paradigms may not know how to evaluate DDR studies (Plonsky, 2022).

#### *Example: The Journal Rejection*

A researcher submits a DDR study of a collaborative writing tool to a top applied linguistics journal. The study includes five cycles over three years, rich qualitative data, and an effective intervention. The reviewers reject it. Reviewer 1 says the sample size is too small. Reviewer 2 says the design changed over time, so it is not replicable. Reviewer 3 says the statistical analysis is insufficient. The researcher is frustrated. The study is published in a lower tier journal and is rarely cited (Kessler, 2018).

#### *Example: The Word Limit Problem*

A researcher writes a DDR study that is 12,000 words. The target journal has a word limit of 8,000 words. The researcher cuts the study down to 8,000 words. The cut version loses the detailed description of each iterative cycle, which is the heart of DDR. The reviewers reject it because the description is insufficient. The researcher is caught between the need for detail and the word limit. The study is never published (Chapelle, 2020).

### **Challenge Eight: The Scalability Challenge**

The eighth challenge is scalability. DDR interventions that succeed in a single context often fail when scaled to multiple contexts. The conditions that made the intervention work, the

particular students, the particular teacher, the particular technology infrastructure, are not present in other contexts. This has led some critics to argue that DDR produces bespoke interventions that are not useful beyond the specific context of the study (Barab & Squire, 2004).

*Example: The Multi Site Failure*

A DDR study at a university in Australia develops an ESL speaking intervention that works well. A researcher at a university in China attempts to scale the intervention to multiple sites. She trains instructors at five universities to use the intervention. At two universities, the intervention works. At three universities, it fails. The researcher cannot predict which universities will succeed. The intervention is abandoned (Reinders & Wattana, 2015).

*Example: The Commercialization Failure*

A researcher develops an ESL vocabulary app through DDR. The app works well in her university context. She attempts to commercialize the app, selling it to other universities. The app does not work at other universities. The researcher discovers that her university had unusually strong technical support and unusually motivated students. The app was designed for those conditions. The commercialization fails. The researcher loses money (Kessler, 2018).

#### 4. Discussion

The synthesis of theoretical literature and case study analysis reveals eight interlocking challenges facing DDR in ESL teaching at the higher education level. The definitional ambiguity challenge confuses new researchers and journal reviewers. The time and resource intensity challenge exceeds typical doctoral timelines and grant cycles. The contextual specificity challenge limits generalizability and replicability. The researcher competence challenge requires expertise across multiple domains. The institutional resistance challenge undervalues DDR outputs in promotion and tenure decisions. The ethical complexity challenge creates dilemmas around informed consent and control

groups. The publication and dissemination challenge makes it difficult to share findings. The scalability challenge limits the transfer of successful interventions.

#### Interactions Among Challenges

These challenges do not operate in isolation; they interact and reinforce one another. Institutional resistance discourages researchers from conducting DDR, which limits the pool of researchers who develop competence, which perpetuates definitional ambiguity. The time and resource intensity of DDR makes it difficult to produce the number of publications needed for promotion, which reinforces institutional resistance. The contextual specificity of DDR findings makes them difficult to publish, which reduces dissemination, which means that others cannot learn from the challenges and successes of DDR studies.

A junior researcher considering a DDR dissertation faces a daunting set of interacting challenges. She will need multiple years to complete the study, during which her funding may run out. She will need competence in design, pedagogy, methodology, and statistics. She will produce findings that are context specific and may not replicate elsewhere. She will struggle to publish the findings in traditional journals. She will then face a tenure committee that does not value DDR outputs. Is it surprising that many researchers choose safer methodologies?

#### Implications for Doctoral Education

The findings have implications for how doctoral students in applied linguistics are trained. Students who wish to conduct DDR need training in multiple domains: instructional design, ESL pedagogy, research methodology, qualitative and quantitative analysis, and often programming or software development. This training cannot be provided in a single methods course. It requires a curriculum that spans multiple courses and multiple years. Few doctoral programs currently offer this training (Brown & Lee, 2019).

Furthermore, doctoral advisors need to be familiar with DDR. Many advisors were trained in traditional paradigms and are unfamiliar with

the design-based research literature. They may discourage students from pursuing DDR because they do not understand it. Doctoral programs need to provide professional development for faculty.

### **Implications for Academic Publishing**

The findings also have implications for academic publishing. Traditional journal formats and word limits are not well suited to DDR studies. DDR studies need space to describe multiple iterative cycles, to document design decisions, and to discuss failures as well as successes. Journals could create special DDR sections with different format requirements. They could allow longer articles for DDR studies. They could publish DDR studies in installments, with each installment describing one cycle. Some journals are moving in this direction, but more are needed (Plonsky, 2022).

### **Implications for Research Policy**

The findings have implications for university research policy and funding agencies. Promotion and tenure committees should recognize DDR outputs, including design documents, software tools, and practitioner reports, as legitimate research contributions. They should adjust publication expectations for researchers who conduct DDR, recognizing that a DDR study may produce fewer journal articles than a traditional study but may produce other valuable outputs. Funding agencies should develop grant programs specifically for DDR, with longer timelines, higher budgets, and appropriate review criteria.

### **Pathways Forward**

Despite these challenges, DDR remains a valuable methodology for ESL teaching research. The challenges are not reasons to abandon DDR but reasons to improve its practice, its institutional support, and its dissemination. Some promising developments are emerging. Collaborative research teams distribute expertise across members. Cross institutional research networks share data and findings. Online publication platforms allow for extended reporting of DDR studies. Institutional review

boards are developing guidelines specifically for DDR. These developments suggest that the field is learning to address the challenges.

### **5. Conclusion**

This conceptual paper examined the persistent challenges associated with the implementation of Design and Development Research (DDR) in English as a Second Language (ESL) teaching within higher education. Through a systematic synthesis of theoretical literature and illustrative case analyses, the study identified eight interrelated challenges: definitional ambiguity, time and resource intensity, contextual specificity, researcher competence, institutional resistance, ethical complexity, publication and dissemination barriers, and scalability limitations. Rather than functioning independently, these challenges reinforce one another, creating methodological, institutional, and practical constraints that influence the quality, implementation, and impact of DDR in applied linguistics.

Despite these constraints, the analysis demonstrates that DDR remains a highly valuable methodological approach for narrowing the long-standing gap between educational research and classroom practice. Its iterative design process, emphasis on authentic learning environments, and integration of theory with practice enable researchers to develop innovative, context-sensitive instructional solutions while simultaneously generating theoretical contributions. Consequently, the future success of DDR depends not only on methodological rigor but also on stronger institutional support, improved researcher preparation, more appropriate publication opportunities, and funding mechanisms that recognize the distinctive nature of design-oriented research.

Future research should move beyond conceptual analyses by conducting empirical investigations that evaluate how the identified challenges influence the implementation and outcomes of DDR projects across different educational contexts. Comparative studies involving multiple universities and countries would enhance understanding of the transferability and scalability of DDR interventions. Researchers

should also examine effective models of collaborative and interdisciplinary DDR teams, investigate ethical frameworks specifically designed for iterative educational research, and evaluate institutional policies that support design-based scholarship. Furthermore, longitudinal studies are needed to determine the sustainability

and long-term educational impact of DDR-generated innovations in ESL teaching. Such investigations will contribute to strengthening the methodological foundation of DDR and expanding its role in advancing evidence-based language education.

## References

- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences*, 13(1), 1–14.
- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences*, 2(2), 141–178.
- Brown, J. D., & Lee, H. K. (2019). Design-based research and the doctoral dissertation: Challenges and opportunities. *TESOL Quarterly*, 53(4), 1123–1135.
- Chapelle, C. A. (2003). *English language learning and technology: Lectures on applied linguistics in the age of information and communication technology*. John Benjamins.
- Chapelle, C. A. (2020). Replication and design-based research in CALL. *Language Learning & Technology*, 24(2), 12–24.
- Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15–22). Springer.
- Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5–8.
- Dewey, J. (2008). *Democracy and education*. Wilder Publications. (Original work published 1916)
- Egbert, J., & Hanson-Smith, L. (2021). Researcher competence in design-based research: A study of CALL doctoral dissertations. *CALICO Journal*, 38(3), 312–331.
- Ellis, R. (2005). *Instructed second language acquisition: A literature review*. Ministry of Education.
- Kessler, G. (2018). Technology and the future of language teaching. *Foreign Language Annals*, 51(1), 205–218.
- Kukulska-Hulme, A. (2019). Mobile language learning innovation inspired by migrants. *Journal of Learning for Development*, 6(2), 116–129.
- Levy, M. (2019). Design-based research in CALL: A methodological framework. *Language Learning & Technology*, 23(3), 45–62.
- Lightbown, P. M. (2019). Perfecting the imperfect: Classroom research over multiple decades. *Annual Review of Applied Linguistics*, 39, 1–14.
- Plonsky, L. (2022). Publication challenges in applied linguistics: A survey of journal editors. *Studies in Second Language Acquisition*, 44(3), 812–834.
- Plonsky, L., & Ziegler, N. (2016). The CALL SIG survey: Defining, describing, and charting a research agenda. In C. A. Chapelle & S. Sauro (Eds.), *The handbook of technology and second language teaching and learning* (pp. 457–472). Wiley-Blackwell.
- Reinders, H., & Wattana, S. (2015). Affect and willingness to communicate in digital game-based learning. *ReCALL*, 27(1), 37–57.
- Reinking, D., & Bradley, B. A. (2022). Design-based research and the problem of scaling up. *Reading Research Quarterly*, 57(4), 1125–1142.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339.

## Acknowledgements (Optional)

The authors received no financial support for the research, authorship, and/or publication of this article.

## Author Declaration

### Declaration of Originality

The authors declare that this manuscript is an original work that has not been published previously and is not under consideration for publication elsewhere, either in whole or in part. All sources have been appropriately acknowledged and cited in accordance with academic standards.

### Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.