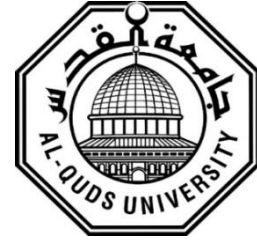


Deanship of Graduate Studies

Al-Quds University



**The Effectiveness of a Proposed Self- Directed Learning
Model Employing Generative Artificial Intelligence on
Eleventh Graders' English Writing and Speaking Skills
and their Perceptions of Using the Model**

Student name

Shaima Abdulkareem Ali

A Doctoral Dissertation

Jerusalem – Palestine

1448 / 2026

**The Effectiveness of a Proposed Self- Directed Learning
Model Employing Generative Artificial Intelligence on
Eleventh Graders' English Writing and Speaking Skills
and their Perceptions of Using the Model**

Prepared by:

Shaima Abdulkareem Ali

Supervisor:

Prof. Khaled Abdel Jaleel Dweikat

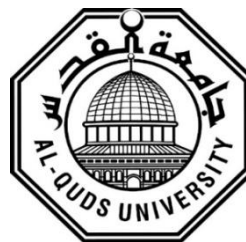
**A Doctoral Dissertation submitted in Partial Fulfillment of
the Requirements for the Degree of Doctor of Philosophy
(Ph.D.) in Curricula and Instruction/ Al-Quds University.**

1448 / 2026

Al-Quds University

Deanship of Graduate Studies

Program



Dissertation Approval

**The Effectiveness of a Proposed Self- Directed Learning Model Employing
Generative Artificial Intelligence on Eleventh Graders' English Writing and
Speaking Skills and their Perceptions of Using the Model**

Prepared by: Shaima Abdulkareem Ali

Registration No: 2311022

Supervisor: Prof. Khaled Abdel Jaleel Dweikat

Doctoral Dissertation submitted and approved date: 29 / 7 /2026

The names and signatures of the examining committee members are as follows:

- 1. Head of Committee:** Prof. Khaled Abdel Jaleel Dweikat
- 2. Internal examiner:** Prof. Buad ALKhales
- 3. External examiner:** Prof. Husam Qaddoumi
- 4. External examiner:** Dr. Zeyad Altaneh

Signature 
Signature 
Signature 
Signature 

Jerusalem-Palestine

1448/2026

Dedication

To my father and mother, who provided me with a blessing when you lit my way, you also provided me with faith and strength to continue my journey.

To my dear children, I do all this for you, and you are the joy at the end of every day. I hope this work will remind you that learning never ends, and that every dream can be attained.

To those who believed in me before I believed in myself, this is for you.

And to Palestine, I hope that our children will be able to capture every opportunity to learn, grow, and have the freedom to dream.

Declaration

I, the undersigned, Shaima Abdulkareem Ali, hereby declare that I have submitted this dissertation, entitled “The Effectiveness of a Proposed Self- Directed Learning Model Employing Generative Artificial Intelligence on Eleventh Graders’ English Writing and Speaking Skills and their Perceptions of Using the Model” to Al-Quds University for the degree of Doctorate. This dissertation is the result of my own research, except as indicated, and this dissertation or any part of it has not been previously submitted for any higher degree to any other university or institute.

Name: Shaima Abdulkareem Ali

The signature:

A handwritten signature in dark ink, appearing to be 'S. Ali', is placed next to the text 'The signature:'. The signature is written in a cursive, flowing style.

Date: 29th July 2026

Acknowledgements

My sincere appreciation goes to those dearest people who have helped me pursue my doctoral dissertation.

First of all, I would like to convey my deepest thanks to my supervisor, Prof. Khaled Abd Al-Jaleel Dweikat, whose expert guidance, scholarly insight, and support have accompanied me throughout this entire journey. His patience, constructive feedback, and true involvement in this work were invaluable.

I would also like to thank the faculty members at Al-Quds University who taught me during my doctoral program, providing knowledge and mentorship that shaped my development as both a researcher and an educator. Their commitment to the pursuit of academic excellence has had an impact on my dissertation and me- I will always be grateful for that.

I would like to give special thanks to Al-Quds University for providing me with the academic environment and resources that enabled me to complete my research, and for its continued commitment to the future of educational scholarship in Palestine.

Special thank you also goes to the members of the defense committee, Prof. Buad Alkales, Prof. Husam Qaddoumi, and Dr. Zeyad Altaneh.

Finally, I would like to express my sincere gratitude to the two cooperating teachers, who generously opened their classrooms and dedicated their time and effort to support the implementation of this study.

I am equally grateful to the students who participated in the ASIM project. Their honesty, active engagement, and genuine desire to learn were what truly brought ASIM to life, and without their willingness to embrace a new and demanding way of learning, this research would not have been possible.

Research Terms

The key terms of the study are defined as follows:

1. Generative Artificial Intelligence (GAI)

Conceptual Definition:

Generative AI is the name associated with a range of AI-based technology, most commonly represented by large language models (LLMs), such as ChatGPT, that generate new text, dialogue or responses using previously learned knowledge, creating opportunities for interactivity, personalization, and on-demand use of language in a way that has quickly become an important tool in language instruction (Li, Tan, Wang, & Lowell, 2025).

Operational Definition:

In this research, the term 'generative AI' is specifically defined as a student tool that uses ChatGPT in generating writing and speaking, receiving quick feedback on their language use, and interacting with one another through the generation of structured English during the three phases of the ASIM as it relates to language learning.

2. Self-Directed Learning (SDL)

Conceptual Definition:

Self-directed learning is the ability of people to manage their own education by autonomously determining their educational needs; setting goals for what they wish to achieve; finding out what resources are available to help them accomplish those goals; using appropriate learning techniques; and evaluating how well they have learned (Knowles, 1975).

Operational Definition:

For this study, students' self-directed learning (SDL) includes identifying their need for improving academic writing and speaking in English, establishing goals that are measurable and specific to improving those two skills, selecting and configuring appropriate educational tools, setting up a way to measure progress towards meeting the established goals independently during all stages of ASIM; and evaluating, using reflective journals and structured observation checklists maintained by the researcher and the cooperating teachers, students' SDL behaviors through the three phases of ASIM during the course of the intervention project.

3. Academic Writing Skills

Conceptual Definition:

Academic writing is a complex productive skill that requires learners to organize ideas logically, employ appropriate academic language, and adhere to specific standards of structure and textual organization (Hyland, 2019).

Operational Definition:

To examine academic writing skills, the researcher developed and validated two parallel forms (pre-test and post-test) to measure students' writing performance according to five criteria:

topic sentence construction, supporting sentences, concluding sentences, grammar and sentence structure, and vocabulary use. Each performance was scored out of 50 points using a standardized rubric, applied by the researcher and two cooperating teachers.

4. English Speaking Skills

Conceptual Definition:

Speaking skills involve the use of spoken English for communicative purposes, including fluent and spontaneous delivery, linguistic accuracy, pronunciation, and the appropriate use of language across communicative contexts. More broadly, oral language use requires the ability to engage in and respond to communicative acts within particular social situations, that is, to exercise communicative competence (Richards, 2008; Celce-Murcia, 2008).

Operational Definition:

Speaking skills were assessed using a parallel-forms oral pre-test and post-test consisting of two tasks. Task One (Responding to a Prompt) required students to speak for three minutes on an assigned topic and was scored across four criteria: fluency and pronunciation, content coverage, vocabulary range, and grammatical accuracy. Task Two (Interactive Conversation) required students to respond to five spoken questions and was scored using a single holistic performance descriptor per question. As with the writing assessment, scoring was conducted by the researcher and cooperating teachers using standardized rubrics, with high inter-rater reliability.

5. The AI-Student Inversion Model (ASIM)

Conceptual Definition:

An instruction model is a highly structured method of teaching that has been developed based on research. It consists of a well-defined set of procedures and strategies to reach a defined set of educational goals and maximum student learning (Joyce, Weil, and Calhoun, 2015).

Operational Definition:

The ASIM Instructional Model is a theoretical model developed by the researcher to teach students using AI tools, specifically ChatGPT, while using self-directed learning principles that create a three-stage instructional process: AI-guided discovery, Human-AI co-creation, and Student-as-Teacher mastery. In the first stage, students will operate independently with the AI in order to create the necessary foundational knowledge prior to participation in the classroom. In the second stage, students will produce authentic writing and speaking products in conjunction with the AI and the teacher, both of which will provide dual scaffolding support for each student. In the third stage, students will teach, transfer knowledge, correct errors, and evaluate the product of the AI from a teacher perspective, thereby inverting the traditional role of instructor and achieving the highest level of Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001). A more detailed explanation of ASIM can be found in Chapter three and in Appendix 1 of this dissertation.

6. Effectiveness

Conceptual Definition:

In effectiveness, the focus is on the impact or utility of the work done by teachers and schools on student learning outcomes informed by evidence and theory to enhance quality in education (Creemers & Kyriakides, 2015).

Operational Definition:

‘Effectiveness’ in this study is defined as the use of ASIM resulted in demonstrably significant (statistically and practically) improvement of eleventh grade student’s academic writing and English speaking skills approached by: differences between students’ mean pre- and post-test scores on the academic writing skills test and English speaking and conversation skills test; the significance of this difference established using paired-samples t-tests and an independent samples t-test to test for between-groups significance.

Abstract

This study aimed to design, implement, and evaluate the effectiveness of the Student Reflection and Artificial Intelligence (ASIM) model in developing academic writing and speaking skills in English among eleventh-graders in Palestine. The study followed the experimental method with a quasi-experimental design and a mixed approach (quantitative, qualitative). The sample consisted of 50 students (28 female and 22 male) from two public schools in Ramallah and Al-Bireh Governorate during the 2025-2026 academic year. To achieve the study's objectives, two academic writing and speaking tests were developed (pre-test and post-test), along with an observation checklist, a semi-structured interview guide, and reflective journals for the researcher and two cooperating teachers. The validity and reliability of these instruments were verified. The results showed statistically significant differences ($\alpha \leq 0.05$) between pre-test and post-test scores in academic writing and speaking skills, in favor of the post-test, for both the female and male groups who received the ASIM intervention. Furthermore, female students outperformed male students in the post-test for both skills. Observational data revealed a steady increase in self-learning behaviors across the six sessions, with the average rising from (31.67) in the first session to (46.50) in the sixth. Qualitative findings from semi-structured interviews and reflective journals revealed that students perceived the ASIM model as a notably distinct and more engaging approach to learning, one that appeared to support the development of their self-learning behaviors, cognitive awareness, and confidence in academic writing, despite initial challenges related to the technical limitations and cognitive requirements of the third phase of the model. The study recommended adopting the ASIM model as a structured framework for teaching English writing and speaking skills across all secondary school levels.

Keywords: Generative Artificial Intelligence (GAI), Self-directed learning (SDL), Academic Writing Skills, English Speaking Skills, The AI-Student Inversion Model (ASIM), and Effectiveness.

Table of Contents

Declaration	i
Acknowledgements	ii
Research Terms.....	iii
Abstract	vi
List of tables	xii
List of Appendices	xv
List of Definitions	xvi
Chapter One.....	1
1.1Background of the Study.....	1
1.2 Research Problem	3
1.3 Research Questions.....	6
1.4 Research Hypotheses	7
1.5 Research Objectives.....	7
1.6 Significance of the Research	8
1.7 Study Delimitations.....	10
1.8 Study Limitations.....	11
Chapter Two	13
Theoretical Framework and Previous Studies.....	13
2.1 Introduction.....	13
2.1.1 The Structural Logic of the Chapter	14
2.1.2 The Theoretical Framework as an Integrated Whole.....	17
2.2 Constructivism: The Epistemological Foundation of ASIM.....	17
2.2.1 The Problem Constructivism Is Answering.....	17
2.2.2 Piaget: Active Construction of Knowledge	18
2.2.3 Vygotsky: The Social Conditions of Constructive Learning.....	21
2.2.4 The Constructivist Core of ASIM's Three Phases	23
2.3 Self-Directed Learning Theory.....	25
2.3.1 Knowles' Andragogy and Self-Directed Learning	25
2.3.2 Learner Autonomy.....	27

2.3.3 Self-Regulated Learning	28
2.3.4 Role of Technology in Self-Directed Learning	29
2.3.5 Synthesis: SDL as ASIM's Foundation	29
2.4 Learning by Teaching: The Theoretical Rationale for ASIM's Phase 3 Role Inversion	30
2.4.1 The Problem That Learning by Teaching Addresses	30
2.4.2 Teaching as the Most Demanding — and Most Effective — Learning Strategy	31
2.4.3 The Six Mechanisms: Why Teaching Produces the Deepest Learning	31
2.4.4 ASIM's Phase 3 as a Realization of All Six Mechanisms.....	34
2.4.5 Why the AI Is an Effective Teaching Target	35
2.4.6 Learning by Teaching and Language Skill Development	36
2.4.7 The Relationship between LbT and the Constructivist Foundation	37
2.5 Cognitive Apprenticeship: How Expertise Transfers Across ASIM's Three Phases	38
2.5.1 The Problem: Invisible Expertise in Palestinian EFL Classrooms	38
2.5.2 The Cognitive Apprenticeship Framework	39
2.5.3 The CA Sequence across ASIM's Three Phases	40
2.5.4 The Problem of Invisible Cognition and the AI Solution	41
2.6 TPACK and S-TPACK: Relocating Technological-Pedagogical Integration from Teacher to Student.....	42
2.6.1 The Problem and the Framework	42
2.6.2 S-TPACK: From Teacher Integration to Student Agency	43
2.7 Higher-Order Thinking and Metacognition: The Cognitive Progression Built into ASIM	45
2.7.1 The Problem: A Cognitive Ceiling in Palestinian EFL Instruction.....	45
2.7.2 Bloom's Revised Taxonomy as a Design Framework	46
2.7.3 Metacognition: The Thread That Runs Through All Three Phases	47
2.7.4 The Cognitive Progression across ASIM's Phases	48
2.8 Artificial Intelligence in Education	48
2.8.1 Generative AI: Capabilities and Limitations	48
2.8.2 AI as Tutor, Tool, and Tutee: ASIM's Triadic Framework	50

2.8.3 AI Literacy and Prompt Engineering.....	51
2.8.4 Ethical Considerations and Pedagogical Concerns.....	51
2.8.5 Empirical Evidence: GAI in Language Learning	53
2.8.6 Distinctive Contributions of ASIM.....	55
2.9 English Language Skills Development: Writing and Speaking in the Palestinian EFL Context	55
2.9.1 Academic Writing Development in EFL Contexts.....	57
2.9.2 Speaking English Skills Development	60
2.9.3 Integration of Writing and Speaking in ASIM.....	63
2.9.4 Summary: How ASIM Addresses Writing and Speaking Development.....	64
2.10 Previous Studies	65
2.10.1 Theoretical Frameworks for Artificial Intelligence in Education	66
2.10.2 Generative Artificial Intelligence and EFL Writing Skills.....	67
2.10.3 Generative Artificial Intelligence and EFL Speaking Skills	69
2.10.4 Self-Directed Learning and Generative Artificial Intelligence	71
2.10.5 Critical Thinking, Motivation, and Engagement in AI-Enhanced EFL Learning.....	73
2.10.6 AI and Technology Integration in the Palestinian Context.....	74
2.10.7 Summary and Research Gap	77
Chapter Three.....	79
Methodology.....	79
3.1 Research Methodology	79
3.2 Research Design	80
3.2.1 Justifications for Selecting this Design	81
3.3 The AI-Student Inversion Model (ASIM)	81
3.3.1 Model Overview and Conceptual Origins.....	81
3.3.2 Theoretical Foundations	82
3.3.3 Model Structure: The Three Phases.....	82
3.3.4 ASIM in the Current Study.....	85
3.4 Study Population and Sample:	86
3.5 Study Sample.....	86

3.5.1 Sample Selection Criteria	87
3.5.2 Sample Size Justification	87
3.5.3 Sample Distribution	88
3.5.4 Statistical Analysis Procedure.....	88
3.6 Data Collection Procedures and Implementation Context	88
3.6.1 Pre-Intervention Phase	89
3.6.2 Intervention Phase.....	89
3.6.3 Post-Intervention Phase.....	90
3.6.4 Qualitative Data Collection Phase	90
3.7 Research Instruments.....	91
3.7.1 Quantitative Measurement Instruments.....	91
3.7.2 Control Variable Instrument	93
3.7.3 Qualitative Data Collection Instruments	94
3.8 Study Variables	96
3.8.1 Independent Variable	96
3.8.2 Dependent Variables	96
3.8.3 Moderating and Control Variables	97
3.9 Validity and Reliability	98
3.9.1 Validity	98
3.9.2 Reliability	100
3.10 Ethical Considerations	101
3.10.1 Institutional and Official Permissions	102
3.10.2 Informed Consent.....	102
3.10.3 Confidentiality and Anonymity	103
3.10.4 Voluntary Participation and Right to Withdrawal.....	103
3.10.5 Protection from Harm	103
3.10.6 Data Use and Research Integrity	103
3.10.7 Ethical Considerations in the Palestinian Context.....	104
Chapter Four	105
The Results of the study	105
4.1 Introduction.....	105

4.3 Results related to the first research question	106
4.4 Results related to the second research question.....	107
4.5 Results related to the third research question.....	107
4.6 Results related to the fourth research question	109
4.7 Results related to the fifth research question	110
4.7.1 Quantitative Findings: Self-Directed Learning across Sessions	110
4.7.2 Qualitative Findings: Structured Observation Notes	113
4.7.3 Qualitative Findings: Semi-Structured Interviews	114
4.8 Results related to the sixth research question	118
4.8.1 Reflective Journals —Researcher and Cooperating Teachers	118
Chapter Five.....	122
Discussion of the Results	122
5.1 Introduction.....	122
5.2 Discussions	122
5.3 Recommendations	137
References.....	140
Appendices	150
الملخص	213

List of tables

No	The table's No	Title	Page
1	Table (2.1.A)	Theoretical Frameworks, Principal Contributions, and Alignment with ASIM's Three-Phase Architecture	15
2	Table (2.1.B)	Theoretical Frameworks, Principal Contributions, and Alignment with ASIM's Three-Phase Architecture	16
3	Table (2.2.A)	Piagetian Principles: The Problem in Conventional EFL and ASIM's Architectural Response	20
4	Table (2.2.B)	Piagetian Principles: The Problem in Conventional EFL and ASIM's Architectural Response	21
5	Table (2.3.A)	Vygotsky's (1978) Framework: Structural Constraints in Palestinian EFL and ASIM's Responses	59
6	Table (2.3.B)	Vygotsky's (1978) Framework: Structural Constraints in Palestinian EFL and ASIM's Responses	60
7	Table (2.4.A)	Writing Challenges and ASIM Solutions	64
8	Table (2.4.B)	Writing Challenges and ASIM Solutions	65
10	Table (3.1)	ASIM Instructional Cycle: Phase Structure and Cognitive Alignment	83

11	Table (4.1)	Paired Samples t-test for Students' Pre- and Post-Test Scores in Academic Writing Skills	106
12	Table (4.2)	Independent Samples t-Test for Gender Differences in Post-Test Academic Writing Scores	107
13	Table (4.3)	Paired Samples t-test for Students' Pre- and Post-Test Scores in English-speaking Skills	108
14	Table (4.4)	Independent Samples t-Test Results for Gender Differences in Students' Post-Test English Speaking Skills	109
15	Table (4.5)	Descriptive Statistics for Self-Directed Learning Behaviors across Intervention Sessions	110
16	Table (4.6)	Tests of Within-Subjects Effects for Session Improvement	111
17	Table (4.7)	Paired Samples t-Test Results for Session 1 vs Session 6 across Dimensions	112

List of figures

No	The Figure's No	Title	Page
1	Figure 2.1	Learning by Teaching: Six Mechanisms and Their Realization in ASIM Phase 3	35
2	Figure 3.1	The AI-Student Inversion Model (ASIM): Three-phase instructional cycle showing role distribution, Bloom's cognitive levels, and theoretical foundations	83
3	Figure 4.1	Thematic map: Semi-structured interview findings	115
	Figure 5.1	Integrated Synthesis of Quantitative and Qualitative Findings across ASIM's Three Phases	137

List of Appendices

No	Appendix	Title	Page
1	1	The AI-Student Inversion Model (ASIM) A Researcher Created, Theoretical-Based Instructional Model For the use of GAI in EFL Education	150
2	2	List of Expert Judges for Content Validity of Proposed Model	171
3	3	University Letter of Facilitation for Research Implementation	173
4	4	Pre- and Post-Test Assessments with Scoring Rubrics English Language Writing and Speaking Skills — Grade 11	174
5	5	List of Expert Judges for Content Validity of Study Instruments	192
6	6	Pre-Intervention Technology Familiarity Questionnaire AI-Student Inversion Model (ASIM) Study (2025-2026)	194
7	7	Semi Structured Interviews After the implementation stage Eleventh Graders' Perceptions of the AI-Based Self- Learning Model	198
8	8	Student Observation Checklist Self-Directed Learning with (GAI) — ASIM Implementation	200
9	9	Researcher's Reflective Journal — Weekly Entry	210

List of Definitions

Abbreviation	Full Form
AI	Artificial Intelligence
ASIM	AI-Student Inversion Model
AWE	Automated Writing Evaluation
CA	Cognitive Apprenticeship
CK	Content Knowledge
EFL	English as a Foreign Language
ESL	English as a Second Language
GAI	Generative Artificial Intelligence
HOTS	Higher-Order Thinking Skills
ICT	Information and Communication Technology
LbT	Learning by Teaching
LLM	Large Language Model
MKO	More Knowledgeable Other
PA	Palestinian Authority
PCK	Pedagogical Content Knowledge
PK	Pedagogical Knowledge
SDL	Self-Directed Learning
SPSS	Statistical Package for the Social Sciences
SRL	Self-Regulated Learning
S-TPACK	Student Technological Pedagogical Content Knowledge
TK	Technological Knowledge
TPACK	Technological Pedagogical Content Knowledge
ZPD	Zone of Proximal Development

Chapter One

1.1 Background of the Study

Language teaching has always evolved in response to the tools available to educators and learners alike. From the printed textbook to the language laboratory, and from audio cassettes to interactive multimedia software, each technological advancement has reshaped how languages are taught, practiced, and acquired. Yet few of these developments have arrived with the speed and disruptive force of artificial intelligence. Within the span of only a few years, AI has moved from a peripheral curiosity in language classrooms to a tool that learners encounter, often independently of their teachers, on a daily basis. This rapid arrival raises a fundamental pedagogical question: not whether AI belongs in language education, but how it should be structured so that it serves genuine learning rather than passive consumption.

The increase in Artificial Intelligence (AI) tools, such as Generative AI (GAI), including ChatGPT, Gemini, and Perplexity AI, has transformed our world today (Xiao et al., 2025). GAI, in particular, has transformed the way education is conveyed, particularly foreign language with gamified, rich, interactive environments that imitate actual communication with speakers of the language and aid learners to become autonomous (Song & Song, 2023). In a systematic review of 144 peer-reviewed empirical studies published between 2023 and 2024, Li et al (2025) documented an explosion of GAI language education research, and a move of inquiry methods from most being exploratory in nature to being more systematic and evidence-based - showing how quickly AI has moved from the sidelines of novelty to become a preferred pedagogical tool for teaching English as a Foreign Language (EFL) globally.

Mobile technologies in combination with self-directed learning were seen to promote the autonomy and motivation of learners involved in language learning (Alfares, 2025). This trend is supported by Roy and Gandhimathi's (2025) systematic review that identifies mobile-assisted language learning, and self-directed learning principles as offering language learners multiple flexible, context-relevant, and self-paced opportunities to learn language outside the classroom, so promoting sustainable language acquisition. The application of self-directed learning techniques to generative AI practices is supposed to provide learners with a

personalized and flexible environment where learners can create their own environment based on their requirements, and which will therefore allow learners to acquire their language skills in a sustainable way (Alfares, 2025).

High school students require two key components for college and job pathways: writing academically and conversing across cultures using English. Recent findings (Barrot, 2023) show that AI, specifically ChatGPT, positively affects both academic writing performance and motivation through immediate feedback and interaction; however, traditional teaching methods do not use ongoing interactions or adaptive supports to nurture autonomous learners. The deficiencies of traditional teaching methods are particularly prevalent within the Palestinian EFL secondary school system due to overcrowded classrooms, examination-driven curricula, and outdated facilities. For these reasons, EFL students face the problem of building up their cognitive abilities as required and as defined by century-old cognitive conceptualizations. For these reasons, EFL students face difficulty developing the cognitive abilities required by contemporary language learning demands. Instead, many students remain passive, waiting for explicit direction from the teacher rather than initiating their own learning process (Qaddumi, Bartram, & Qashmar, 2021).

The purpose of this study is to develop, implement, and evaluate a new instructional model for integrating AI into EFL teaching, one that fundamentally departs from prevailing approaches in which learners are positioned merely as recipients of AI-generated content, feedback, and correction. Instead, the proposed model repositions learners as active agents who progressively assume greater cognitive authority over their interaction with AI throughout the learning process. This change in thinking does not simply represent a change of methodology but is indicative of a greater epistemological change in the conceptualization of technology literacy (Ouyang & Jiao, 2021). The dichotomous approach to AI use in EFL instruction has prevailed in the educational community, with very few examples of how learners can direct AI (Ouyang & Jiao, 2021). The objective of this study is therefore to provide an example of the role-reversal approach to AI and how it can be implemented within the Palestinian secondary EFL context, where the need for creative and agency-based instructional models is both pressing and underexplored (Qaddumi et al., 2021; Hamamra, Mayaleh, & Khalif, 2024; Hamdan et al., 2025).

The significance of both Academic Writing and Speaking is considerably elevated in the secondary EFL curricula in Palestine; they are frequently beyond curricular concern due to the significant role they play academically. For students in the public school system, academic writing skills will be the only point of linguistic access to university and professional life; thus, they are assessed directly through the General Secondary Certificate Examination (Tawjihi) taken by students wishing to attend university. Similarly, oral communication skills will continue to be significant in preparing students for both academic and civic engagement and for future professional lives in a globalized and increasingly digitally connected world.

Ironically, these two skills have generally been very poorly developed (or at least developed with a lesser degree of emphasis) in classrooms for (EFL) in Palestine. While students may not lack the motivation or abilities to develop their writing and speaking skills, the model of instruction generally used to develop those skills is group instruction geared towards preparation for exams rather than to any true personal development of those students, and EFL exam-pedagogies are quite prevalent throughout the Palestinian educational system (Qaddumi et al., 2021; Yaseen & Dweikat, 2025). Writing and speaking both require the type of tailored, individualized support that students receive in the form of continuous, meaningful feedback, in order for students to truly develop into proficient writers and speakers. The relationship of generative AI to the teaching of these skills is thus no simple upgrade in technology, but a structural response to ingrained teaching deficiencies.

It is against this backdrop that the current study emerged, was conceived, and was designed. Responding to the intersection of three underexplored gaps in the existing literature - the almost complete lack of AI-embedded instruction/learning research at the secondary school level, the predominance of student-as-recipient models in AI-assisted EFL research, and the lack of empirically based AI research in the Palestinian context - the researcher presents, implements, and evaluates ASIM: a theoretically grounded, researcher-designed instructional model that consciously repositions the student as an active cognitive agent in the human-AI relationship, rather than a passive recipient of AI-generated output. The model was applied to writing and speaking activities within the eleventh-grade EFL curriculum in Palestinian public schools, using a mixed-methods approach that integrates quantitative and qualitative evidence to examine both the model's effectiveness and students' lived experiences within it — thus contributing, in this small endeavor, to that growing if still fledgling body of work wondering what might become possible in EFL education when it is predicated on placing the student, not the technology, in the middle of the learning process.

1.2 Research Problem

This study is motivated by a persistent and interlocking set of deficits in the English language competencies of Palestinian secondary students, specifically in academic writing, oral communication, and self-directed learning. These deficits weaken students' academic performance at the secondary level and constrain their preparedness for higher education and, ultimately, for professional participation in a globalized economy. The researcher's conviction that these problems are both real and persistent rests on two converging sources of evidence: the empirical literature reviewed below, and nineteen years of direct classroom experience teaching English in Palestinian public schools, during which the same patterns of difficulty have recurred across successive cohorts of students. This dual grounding, in published research and in sustained professional observation, informs the analysis that follows and is addressed more fully toward the end of this section.

English in Palestine is taught and learned as a foreign language (EFL) rather than a second language (ESL), a distinction with significant pedagogical consequences. Unlike ESL learners, who are immersed in a society that reinforces classroom instruction through everyday communicative contact with the target language, Palestinian students have minimal exposure to English outside the classroom and virtually no naturally occurring opportunities to use it for authentic communication (British Council & LearnJam, 2025; Murad & Jalambo, 2019). This structural isolation from the language undermines the development of communicative competence, productive fluency, and autonomous learning habits, even though the Ministry of Education and Higher Education mandates five periods of English instruction per week from Grade 8 through the end of secondary school. These constraints are compounded by severely overcrowded classrooms: a 2025 evaluation by the British Council and LearnJam found that 32% of students in Palestinian public and UNRWA schools are taught in classes of 30 or more, with some classes reaching 45 students — conditions under which communicative teaching, individualized feedback, and meaningful skills development become nearly impossible. The consequences are reflected in outcomes: a national assessment of Grade 6 students found that only 9% achieved advanced proficiency in English, while more than 75% scored below 50 out of 100 (British Council & LearnJam, 2025), a pattern consistent with Shehadeh and Dwaik's (2013) earlier finding that large class sizes, insufficient instructional time, and the absence of digital learning platforms prevent Palestinian EFL teachers from realizing the communicative objectives of the national curriculum.

The assessment structure compounds these difficulties. The Tawjihi end-of-secondary examination, which determines what is taught and how, assesses neither listening nor speaking, producing a well-documented washback effect that systematically deprioritizes oral communication in EFL classrooms (British Council & LearnJam, 2025). Teachers consequently allocate classroom time to grammar, vocabulary, and reading comprehension — the skills the exam rewards — even where speaking and listening activities are explicitly required by the curriculum. As a result, students may complete twelve years of English instruction and pass the national examination without ever having a meaningful opportunity to demonstrate oral communicative ability.

These pedagogical constraints are further compounded by the political, socioeconomic, and humanitarian conditions of the Palestinian context. Military occupation, through roadblocks, checkpoints, and recurrent armed conflict, repeatedly disrupts the continuity of schooling and exposes students to chronic stress and trauma that impair cognitive engagement with learning (British Council & LearnJam, 2025). Teachers and supervisors continue to report the absence of clear institutional policy for addressing the resulting learning loss (British Council & LearnJam, 2025). Overcrowded classrooms, exam-driven pedagogy, limited authentic language input, and chronic instability thus combine to obstruct both the acquisition of

English-as-a-foreign-language skills and the development of the self-directed learning attitudes and behaviors necessary to sustain proficiency over time.

Within this broader context, three specific and interrelated deficits define the problem addressed by this study. First, academic writing remains an area of persistent difficulty at both secondary and tertiary levels. Palestinian EFL learners commonly exhibit insufficient linguistic competence, heavy reliance on literal translation from Arabic, weak cohesion and coherence, and limited command of academic register (Hammad, 2016), problems rooted largely in the transfer of Arabic rhetorical and organizational conventions into English writing, which affects syntax, paragraph development, and argumentation (Hammad, 2016). Second, oral communication is similarly constrained: Abu Alyan (2013) identifies poor pronunciation, limited vocabulary, insufficient exposure to the target language, first-language interference, and low self-confidence as principal barriers facing Palestinian EFL learners, compounded by a fear of error and negative judgment that leads students to favor silence over risk-taking in the classroom (Abu Alyan, 2013), and by minimal opportunity to use English outside formal instruction (Murad & Jalambo, 2019). Third, self-directed learning capacity remains underdeveloped under the teacher-centered pedagogy that predominates in Palestinian classrooms, where knowledge flows unidirectionally from teacher to student and learners have little opportunity to develop the metacognitive skills, strategic awareness, and self-regulation that autonomous language learning requires; the 2025 British Council review similarly found that Palestinian teachers, constrained by overcrowded classrooms, examination pressure, and limited professional support, rely predominantly on lecture-based, memorization-focused instruction that does not cultivate learner autonomy (British Council & LearnJam, 2025).

These findings are consistent with the researcher's own professional observations, accumulated over nineteen years of teaching English in Palestinian public schools. Across successive cohorts, students have consistently approached writing tasks as exercises in translation rather than as acts of independent composition, have shown a persistent gap between curricular expectations and actual output, and have demonstrated a habitual reliance on the teacher as the sole source of information, correction, and evaluation, with little capacity to initiate, monitor, or evaluate their own learning. This concern is particularly acute at Grade 11, the focus of this study, since students at this stage stand on the threshold of higher education, a context in which self-directed learning is not merely valued but required, and in which those who arrive without it are at heightened risk of both academic underperformance and continued underdevelopment of English proficiency.

The structural problem outlined above, persistent deficits in academic writing, oral communication, and self-directed learning, compounded by overcrowded classrooms, exam-driven pedagogy, and chronic political instability, has not been resolved by traditional EFL instruction, nor is it likely to be resolved by AI tools deployed within the prevailing student-as-recipient paradigm, which does not require learners to exercise the cognitive agency, self-

monitoring, and evaluative judgment that genuine language competence demands. This gap between what generative AI could offer and how it is currently used in Palestinian EFL classrooms motivates the central intervention of this study: the design and empirical evaluation of the AI-Student Inversion Model (ASIM), an instructional model that repositions the learner as director of AI-mediated learning rather than its passive recipient.

The intervention targets eleventh-grade learners specifically because this stage represents the final opportunity to develop self-directed learning capacities before the onset of Tawjihi examination pressures and before students enter higher education contexts in which such capacities become indispensable. Academic writing and speaking were selected as the target skills because they are the productive competencies identified above as most severely underdeveloped, and because they carry the greatest structural consequence for Palestinian students, university admission, professional advancement, and participation in a globalized economy all depend on them.

Evaluating the effectiveness of a new instructional model of this kind requires more than a measure of whether change occurs; it requires an understanding of how and why that change occurs, particularly given that gender-differentiated patterns of engagement, gradual shifts in learner autonomy, and the lived difficulty of adapting to an unfamiliar pedagogical role cannot be captured through test scores alone. For this reason, the study adopts an explanatory sequential mixed-methods design, in which quantitative outcomes are subsequently interrogated and contextualized through qualitative inquiry into students' perceptions and lived experiences of the model. Accordingly, the research questions that follow address both what ASIM changes and how students experience that change, and the methodology detailed in Chapter Three operationalizes this dual concern in full.

1.3 Research Questions

This study seeks to answer one main question and six sub-questions, divided into quantitative and qualitative sub-questions, as follows:

Main Question:

What is the effectiveness of the AI-Student Inversion Model (ASIM) — a self-directed learning model utilizing generative artificial intelligence — in developing eleventh-grade students' English academic writing and speaking skills?

Sub-questions:

1. What is the effectiveness of ASIM in developing eleventh-grade students' English academic writing skills?

2. Are there statistically significant differences between eleventh graders' mean scores on the post-test of English Academic Writing Skills attributable to gender?
3. What is the effectiveness of ASIM in developing eleventh-grade students' English-speaking skills?
4. Are there statistically significant differences between eleventh graders' mean scores on the post-test of English-speaking skills attributable to gender?
5. What are eleventh-grade students' experiences and perceptions of engaging with the three phases of ASIM, including their development of self-directed learning behaviors (self-monitoring and autonomous AI-directed learning)?
6. How do the researcher and cooperating teachers perceive the implementation of ASIM, in terms of students' engagement patterns and the contextual challenges encountered during the intervention?

1.4 Research Hypotheses

Null Hypotheses:

Hypothesis 1 (H_{01}): There are no statistically significant differences at ($\alpha \leq 0.05$) between eleventh graders' mean scores on the pre- and post-tests of English academic writing skills attributable to the implementation of the ASIM model.

Hypothesis 2 (H_{02}): There are no statistically significant differences at ($\alpha \leq 0.05$) between eleventh graders' mean scores on the post-test of English academic writing skills attributable to gender.

Hypothesis 3 (H_{03}): There are no statistically significant differences at ($\alpha \leq 0.05$) between eleventh graders' mean scores on the pre- and post-tests of English-speaking skills attributable to the implementation of the ASIM model.

Hypothesis 4 (H_{04}): There are no statistically significant differences at ($\alpha \leq 0.05$) between eleventh graders' mean scores on the post-test of English-speaking skills attributable to gender.

1.5 Research Objectives

The goal of this study is to develop, evaluate and assess a new model called the AI Student Inversion Model (ASIM) which seeks to improve the academic writing and spoken English capability of eleventh graders in the public school system in Palestine with generative AI tools

while learning about student-centeredness through self-directed learning. To achieve this goal, this research focuses on the following:

- To design and operationalize ASIM as a self-directed, AI-mediated instructional model responsive to individual students' learning needs and pace, grounded in the theoretical framework outlined in Chapter Two.
- To evaluate the effectiveness of ASIM in developing eleventh-grade students' English academic writing skills across the dimensions of content, organization and coherence, grammatical accuracy, word choice, and mechanics.
- To evaluate the effectiveness of ASIM in developing eleventh-grade students' English-speaking skills, as measured by the dimensions defined in the study's speaking assessment instrument (see Chapter Three, Instrumentation).
- To examine, through classroom observation and behavioral evidence, how ASIM supports the development of students' self-directed learning behaviors, including self-assessment, self-monitoring, and autonomous AI-directed learning.
- To compare the writing and speaking outcomes of male and female students in order to determine whether ASIM is equally effective in developing these skills across genders.
- To explore eleventh-grade students' perceptions and lived experiences of engaging with ASIM, including the opportunities it presents, the challenges of applying self-directed learning strategies with AI tools, and students' own perceptions of their language development.
- To examine the researcher's and cooperating teachers' perspectives on the implementation of ASIM, including students' engagement patterns and the contextual challenges encountered during the intervention.
- To contribute empirical evidence and theoretical insight to the growing body of literature on the integration of generative AI into EFL instruction at the secondary level.

1.6 Significance of the Research

This study is situated within the broader context of rapid advances in generative AI and increasing calls within educational systems to cultivate learners' capacity for self-directed learning. By proposing, implementing, and evaluating the AI-Student Inversion Model (ASIM) within Palestinian secondary EFL education, this research contributes to an emerging body of scholarship on AI-based instructional models that are theoretically grounded and empirically tested, and that position the student, rather than the technology, as the central agent in the learning process.

Theoretical Significance

This study contributes to the educational research literature by proposing an integrative theoretical framework linking classical self-directed learning theory (Knowles, 1975) to contemporary applications of generative AI in EFL education. ASIM draws on six related theoretical strands, namely Constructivism, Self-Directed Learning, Mastery Learning, Learning by Teaching, Cognitive Apprenticeship, and an extended S-TPACK model proposed by the researcher, which converge on a shared premise: that meaningful learning with AI requires the learner to actively direct, question, and evaluate AI-generated input rather than passively receive it. Constructivism and Cognitive Apprenticeship inform how students construct and internalize language knowledge through guided practice; Mastery Learning and Self-Directed Learning theory inform the model's emphasis on individualized pacing and learner autonomy; Learning by Teaching informs the reflective, output-oriented tasks embedded in the model; and the extended S-TPACK component situates these processes within the specific pedagogical and technological demands of AI-mediated language instruction. Building on this integration, the study offers a conceptual reframing of the EFL learner's role in AI-mediated environments, from a user of AI-generated output to a director of the AI-mediated learning process, a distinction that speaks to a gap identified by Ouyang and Jiao (2021) in how learner agency is conceptualized in the AI-in-education literature.

This research also addresses an underrepresented area within existing empirical work on generative AI in language education. Li et al. (2025) report that the substantial majority of current research on generative AI in language education has been conducted in higher education settings, leaving secondary-level learners comparatively understudied. By examining eleventh-grade students in Palestinian public schools, a context shaped by political instability, infrastructural constraints, and a predominantly teacher-centered instructional culture, this study aims to extend the applicability of existing theoretical accounts to a context that has received limited direct empirical attention (Hamamra et al., 2024; British Council & LearnJam, 2025).

Methodological Significance

This research contributes a set of instruments developed for this study, including parallel writing and speaking assessments, a structured observation checklist, a semi-structured interview guide, and reflective journal protocols for the researcher and cooperating teachers, designed to jointly capture language outcomes, self-directed learning behaviors, and students' lived experiences within a single evaluative framework. These instruments were developed and piloted specifically for this study; their psychometric properties are reported in Chapter Three, and their broader applicability to other EFL contexts remains a question for future validation rather than a claim established by this study alone.

Practical Significance

Findings from this study are intended to be of practical relevance to three groups. For students, ASIM is designed to offer a more individualized learning environment in which academic writing and speaking skills can be developed alongside self-regulatory capacities such as goal-setting, self-monitoring, and self-assessment, capacities identified in Section 1.2 as underdeveloped among Palestinian EFL learners. For teachers, the model offers a framework in which AI supports routine linguistic feedback, potentially allowing instructional time to be redirected toward higher-order concerns such as argumentation and rhetorical organization; structured observation tools may further support more systematic tracking of individual student progress than is typical in large Palestinian classrooms. For educational institutions and curriculum developers, the study offers an empirically evaluated model that may inform curriculum reform or serve as a supplementary resource within the existing eleventh-grade EFL curriculum, and its instruments and findings may be relevant to teacher preparation and professional development programs as AI-related competencies become an increasing consideration in Palestinian EFL education. These potential benefits are proposed as reasoned expectations to be examined through the study's findings, rather than as outcomes guaranteed in advance.

1.7 Study Delimitations

This study was designed and conducted according to the following delimitations, or intentional boundaries of the study.

Thematic Delimitations:

This study aims to produce two productive language skills (i.e. academic writing and speaking English) as the desired outcomes of the ASIM. Although reading and listening are also considered to be essential components of English language skills, they will not be considered as part of this study. As well, only one generative AI tool (ChatGPT) has been chosen for this study as it is easy to access, frequently used within EFL research, and suitable for writing and speaking tasks involved in all three phases of ASIM. No other AI tools or language learning applications were included.

Population Delimitations:

The participants in the study were eleventh-grade students in public secondary schools in Palestine. The target population was divided into two sub-groups: 28 female students from a girls' public school and 22 male students from a boys' public school, both located in the Governorate of Ramallah and Al-Bireh. Neither private school students nor students at other grade levels were included in the target population. Rather, the study focused on public-sector students in Palestine's education system where the major structural challenges identified in the theoretical framework (large class sizes, assessment-driven instruction and limited access to individualized feedback) are most prominent.

Geographical Delimitations:

This study was performed within Ramallah and Al-Bireh in the West Bank, Palestine. The geographical specification of this study was determined by the researcher's affiliation with the institution where data collection will be performed as well as the researcher's ability to practically access the participant schools. Despite this geographical limitation, the findings of the current study may be applicable to similar Palestinian EFL settings in other governorates.

Temporal Delimitations:

The ASIM program was implemented over a span of eight to twelve weeks during the 2025-2026 school year. Therefore, the results of this study demonstrate the short and medium-term impact of ASIM on students' perceptions and proficiencies in writing and speaking, but do not guarantee any continued language development beyond the time period of this research.

Methodological Delimitations:

This research used a mixed-methods explanatory sequential design to gather and analyze quantitative data via pre- and post-test scores for academic writing and speaking skills, plus structured observation data through a checklist, which were analyzed quantitatively across six sessions. The second phase of data collection for this study was qualitative in nature and made up of semi-structured interviews and reflective journals (used to contextualize and interpret the quantitative results) completed by the researcher and cooperating teachers. Because of this design, findings from ASIM's effectiveness and how students experienced it can only be generalized to provide rich contextual data grounded in theory and not statistically representative results reflective of the Palestinian EFL population.

1.8 Study Limitations

Several limitations should be considered when interpreting the results of this study and their generalizability.

Sample-Related Limitations

The sample comprised fifty students (22 males, 28 females) from two public schools. While this sample was appropriate for the explanatory sequential mixed-methods design, yielding valid quantitative data and rich qualitative data, its size limits the generalizability of the findings. The results should therefore be interpreted as context-bound insights into students' experiences with ASIM rather than as statistically representative conclusions for the broader population of eleventh-grade EFL students in Palestine.

Technical Limitations

The intervention required consistent internet access and functioning smart devices, resources not uniformly available given the infrastructural constraints common to Palestinian public schools. In addition, reliance on the free version of ChatGPT constrained students' engagement

with more advanced features, such as the deeper grammar analysis, extended interaction history, and richer feedback options available through ChatGPT Plus, which may have limited the depth of some students' interaction with ASIM.

Linguistic and Cultural Limitations

The study was conducted under difficult political and socioeconomic conditions in Palestine. Economic pressures on the education sector reduced the school week to three days, and intermittent closures and restrictions on movement between localities limited the researcher's ability to expand the sample, extend the study's geographical scope, and maintain consistent in-person follow-up. A fifty-day period of remote learning further required parts of the intervention to be delivered via Zoom and community-based sessions. These adaptations were managed to the extent possible but introduced disruptions to the implementation timeline that were beyond the researcher's control and may have affected the continuity of students' engagement with ASIM. Similarly, consistent with standard practice for classroom-based research in Palestinian public schools, parental consent was obtained verbally, through in-person meetings at each school and follow-up telephone calls, rather than through signed written forms; while this procedure was administratively supported by the schools' principals, the absence of documented written consent is acknowledged as a limitation of the study's ethical procedures.

Contextual and Political Limitations

The study was conducted under difficult political and socioeconomic conditions in Palestine. Economic pressures on the education sector reduced the school week to three days, and intermittent closures and restrictions on movement between localities limited the researcher's ability to expand the sample, extend the study's geographical scope, and maintain consistent in-person follow-up. A fifty-day period of remote learning further required parts of the intervention to be delivered via Zoom and community-based sessions. These adaptations were managed to the extent possible but introduced disruptions to the implementation timeline that were beyond the researcher's control and may have affected the continuity of students' engagement with ASIM.

Chapter Two

Theoretical Framework and Previous Studies

2.1 Introduction

Instructional models are not just a means of delivering material but rather structured formats that dictate how learning will be expected to take place. Delivery methods need to change depending on educational context, learner types, and curriculum objectives. For example, take the history of language instruction: from grammar/translation through audio-lingual model to communicative language instruction through task-based instruction, there has been a constant pattern of methodological change in response to both perceived shortcomings in current methods and an increasing understanding of how language acquisition occurs (Richards & Rodgers, 2014). The increasing number and diversity of instructional models across all types of educational settings do not reflect a lack of consistency among theories; rather, they reflect the active development of responsive approaches to the complicated nature of learning; of approaches that can be assessed, improved, and justified based upon sound principles.

Implicit in all instructional models are assumptions regarding the process of learning. For instance, a teacher-centered instructional model implicitly claims that one person can put knowledge into another's mind (Freire, 1970; Biesta, 2009). A model that incorporates cooperative learning assumes that students are actually learning more through their own social interactions with other students than they would if they studied independently (Vygotsky, 1978; Johnson & Johnson, 1989). A model that assigns homework to students prior to instruction supports the premise that preparing for instruction occurs prior to receiving direct instruction (Bergmann & Sams, 2012). Because no explicit theoretical basis exists in the above models, there is no foundation for making principled judgments about the models; therefore, there can be no basis for improving, defending, or knowing the conditions of failure neither for each instructional model nor for modifying an instructional model based on documented evidence (Schön, 1983; Carr & Kemmis, 1986). The necessity of theoretical

explicitness is much more than providing the theoretical framework; it provides a foundation for implementing pedagogical practices grounded in principles.

The ASIM rests on several foundational premises about learning that diverge from traditional assumptions of instruction. The first premise holds that learning is enhanced when a student directs artificial intelligence to assist their instruction prior to formal classroom teaching, rather than encountering the material for the first time through direct teacher-led instruction. The second premise holds that a student who has taught a concept to an AI develops a deeper understanding of that concept than would result from an equivalent period of conventional study. The third premise holds that the most valuable role a teacher can play is not the transmission of content but higher-order cognitive coaching, and that AI can assume many of the functions that currently consume the majority of a teacher's instructional time.

These premises are grounded in established learning theory, and the purpose of this chapter is to justify them theoretically, to demonstrate, given what is known about how people learn, why an instructional model built on these premises should be expected to produce the outcomes it targets. The empirical evaluation of ASIM's effectiveness, reported in Chapters Four and Five, tests the model's overall impact on students' writing and speaking outcomes; it does not constitute a direct test of each premise in isolation.

The justification of the model's theoretical foundation is based on seven different theoretical frameworks, which aim to add unique and non-duplicated components to the model's foundation. Each framework does not stand alone, as they address the different areas of the pedagogical problem. Constructivism, for instance, deals with the epistemological aspect of knowledge; it asks how knowledge is defined and how this cognitive process produces a comprehensive understanding. Self-Directed Learning (SDL) addresses the structural aspects of the learning context; how must the environment of learning be designed for the cognitive process to be activated? Cognitive Apprenticeship addresses the instructional sequence aspects of how to transfer expertise from one individual to another; what methods will transfer cognitive knowledge/expertise? Learning by Teaching addresses the deepest part of how to relate knowledge acquisition to creating knowledge that is applied or interpreted. TPACK ending with S-TPACK addresses technological integration; how does the mastery of ASIM's use of AI-mediated instruction count as knowledge? Higher-Order Thinking Skills (HOT) frameworks address cognitive progression; how does ASIM's three-phase structure connect with ascending cognitive levels? The literature of language skills development addresses the context of why all these areas of theory have particular urgency regarding the Palestinian EFL context that ASIM was developed for.

2.1.1 The Structural Logic of the Chapter

Section 2.2 establishes Constructivism as the philosophical foundation of ASIM, the epistemological position without which none of the model's subsequent commitments makes sense. If knowledge can be transmitted, then the entire rationale for ASIM's learner-active architecture collapses; it is only because knowledge must be constructed by each individual learner since the model emphasizes learner's agency, productive challenge, and cognitive accountability which are essential components rather than desirable factors. Section 2.3 examines how Self-Directed Learning theory (SDL) operationalizes this constructivist commitment within ASIM's Phase 1 architecture. Section 2.4 develops the theoretical rationale for Learning by Teaching — the framework that underpins perhaps ASIM's most innovative feature: the Phase 3 role inversion in which the student becomes the teaching agent. Section 2.5 explains the instructional sequence logic that governs Cognitive Apprenticeship across all three phases. Section 2.6 introduces TPACK and its ASIM-specific extension, S-TPACK. Section 2.7 examines Higher-Order Thinking and Metacognition. Section 2.8 surveys the AI in education research base. Section 2.9 addresses language skills development in the Palestinian EFL context. Finally, Section 2.10 synthesizes all seven frameworks into a unified theoretical map.

As a preliminary observation concerning the chapter's organization, the arrangement of frameworks does not demonstrate ascending or descending ranking according to degree of importance; rather, their sequence is based upon the degree of theoretical generality from the most philosophically fundamental (i.e., Constructivism) to the most contextually specific (i.e., Palestinian EFL language development). The framework sequence ultimately reflects the logical dependency on each framework from one to the next; specifically, each subsequent framework builds upon the epistemological and instructional bases established by the previous framework and cannot be comprehended fully in isolation from the foundational framework that constitutes the basis for the whole framework. The next is (Table 2.1) provides an overview of all seven frameworks, their key contributors, their key contributions toward ASIM's design, and the phases that each directly contributes to defining, thus creating a reference guide to the reader which may be consulted throughout the chapter.

Table 2.1.A: Theoretical Frameworks, Principal Contributions, and Alignment with ASIM's Three-Phase Architecture.

Theoretical Framework	Principal Theorists	Core Contribution to ASIM	Primary ASIM Phase
Constructivism	Piaget (1970); Vygotsky (1978)	Knowledge is built through active engagement, not passive reception; the Zone of Proximal Development) ZPD) principle justifies using AI as individualized "More Knowledgeable Other (MKO)" for all learners at one'.	Foundation across all three phases

Table 2.1.B: Theoretical Frameworks, Principal Contributions, and Alignment with ASIM's Three-Phase Architecture.

Self-Directed Learning	Knowles (1975); Zimmerman (2002)	Learner determines need, sets goal, chooses resources, and evaluates outcome; teacher acts as facilitator not as transmitter'.	Phase 1 — student-directed pre-class AI inquiry
Cognitive Apprenticeship	Collins, Brown & Newman (1989)	Expertise for cognitive processes visible through modeling-coaching/scaffolding/fading Responsibility on to learner.	Phases 1–2 (model/coach/scaffold); Phase 3 (fading/articulation)
Learning by Teaching	Duran & Topping (2017); Fiorella & Mayer (2015)	Teaching will cement student understanding more than equivalent passive study using six different cognitive mechanisms.	Phase 3 — student as teaching agent who directs and evaluates the AI
S-TPACK	Mishra & Koehler (2006); ASIM extension	ASIM extension ASIM reconceptualises integration of Content Knowledge, Pedagogical Knowledge, and Technological Knowledge as learner competency demonstrated via prompt engineering.	Phase 3 — student-teacher integrates CK + PK + TK
Higher-Order Thinking & Metacognition	Anderson & Krathwohl (2001); Flavell (1979)	Three ASIM phases align with Bloom's progression; metacognition monitored continually across all three phases.	Progressive: Phase 1 (remember/understand), Phase 2 (apply/analyze), Phase 3 (evaluate/create)
EFL Language Skills Development	Hyland (2019); Canale & Swain (1980)	For writing and speaking in EFL, authentic, sustained, individualized practice required but not possible due to structure of large-class Palestinian instruction.	Phases 2–3 — AI-mediated practice with real-time feedback compensates for structural EFL constraints

2.1.2 The Theoretical Framework as an Integrated Whole

Before proceeding to the individual frameworks, it is important to articulate the single pedagogical issue that each of the seven pedagogically based frameworks is responding to in their own way. The issue can be stated very simply: conventional EFL teaching in Palestinian secondary schools has been designed to create passive language learners who were told about English and who have not been taught to be active participants in using English. The predominant form of delivery for language instruction has been teacher-centered, uniform and assessment-driven (Al-Ureimi, 2021; Aqel, 2002). A teacher explains the grammar rules for the students to copy and practice, then the lesson is completed and the same cycle is repeated. During this cycle, the learner is an unchanging passive recipient. The student is the recipient of an explanation, the completion of a controlled exercise, and the demonstration of their retention of knowledge through test measurement that represents only their ability to reproduce received knowledge, and not their ability to use their internalized language competence. This description is not purely a theoretical abstraction - it is a reflection of the researcher's professional experience over 15 years of teaching EFL in Palestinian secondary schools, and it has been confirmed by the researcher through providing university level instruction in EFL at a Palestinian institution of higher education, as well as being consistent with the patterns documented in existing EFL research.

The sections that follow develop each framework in the depth that this multi-dimensional problem requires, beginning with the epistemological foundation on which all subsequent theoretical commitments rest.

2.2 Constructivism: The Epistemological Foundation of ASIM

2.2.1 The Problem Constructivism Is Answering

To understand why Constructivism is the appropriate philosophical foundation for ASIM, it is necessary to begin not with the theory but with the problem that motivates it. The problem is this: classrooms around the world, and Palestinian EFL classrooms with particular intensity, are structured around the assumption that a teacher who explains something clearly and completely has done the essential work of instruction, and that a student who listens attentively and takes notes has done the essential work of learning. The flow of expertise is from the teacher's verbal and written explanation into the learner's head, similar to the way in which a signal travels through a wire from a transmitter to a receiver. Accordingly, teaching is a process of transmission, and learning is a process of reception.

Regardless, this model is appealing and easy to work with logistically; nevertheless, overwhelming evidence reveals that it is not how real understanding comes about. Classrooms everywhere produce students who are able to reproduce the answers they are taught on the tests they are given and not able to apply those answers in situations where they have not seen

that same situation before, explain why they do what they do in the procedures they have memorized, or transfer their ‘knowledge’ to a situation that bears a superficial resemblance to that which they studied (Bransford et al., 2000; Biggs & Tang, 2011). Palestinian EFL instruction exemplifies this pattern with particular clarity: students who have studied English for a decade cannot hold a spontaneous conversation, write an unprompted text, or identify what is wrong with a grammatically deviant sentence they have not previously encountered (Al-Ureimi, 2021; Aqel, 2002). They have been told about English, systematically and for years, without developing the capacity to use it. The gap between what transmission pedagogy produces and what genuine language competence requires is not a matter of insufficient effort or teaching skill; it is a structural consequence of a fundamentally mistaken theory of what learning is (Freire, 1970; Biesta, 2009).

According to constructivism, knowledge is not entity that can be given to you by someone else, but instead, it is a collection of cognitive structures that each individual develops through their own experience of the world (Piaget, 1970; Von Glasersfeld, 1995). This is not a claim that teachers are useless or that explanation is valueless; it is a claim that explanation, even when it is clear and comprehensive, is only a stimulus — an occasion for constructive cognitive activity by the learner — and that the construction itself cannot be performed by anyone other than the learner (Dewey, 1938; Von Glasersfeld, 1995). For example, when a teacher gives a student an explanation about how to use the present perfect tense, the student will then have an opportunity to build a conceptual understanding of the present perfect tense, but whether or not the student is able to construct this understanding will depend on how much work the student does with that explanation (Ellis, 1997; Vygotsky, 1978). An explanation does not construct anything.

2.2.2 Piaget: Active Construction of Knowledge

Piaget’s conception of knowledge comes from his background in biology, not necessarily from his experiences as a teacher. Piaget derived his notion of knowledge through his educational background in biology (rather than through direct experience as an educator). According to Piaget, intelligence is interpreted as an organismic system; thus, intellectually, an individual is viewed as an adaptive being similar to any other living thing. Just like the physical body changes physically in order to survive in its environment, so does the mind change, by changing its schemas (i.e., “knowledge of the world”). If an individual has an experience that “fits” within an existing schema, then they will absorb that experience into their existing knowledge structure; this is assimilation. If they experience or come upon something that does not fit into their existing knowledge structure or schema, then they will need to change that schema to make sense of it; this is accommodation. This state, in which the individual seeks to achieve, is known as equilibration, where the schemas used to explain a person’s experiences continue to balance out and not clash against them, thereby creating discomfort for the individual (Piaget, 1952).

The educational implications of this framework are clarifying and, for conventional pedagogy, genuinely subversive. A Palestinian EFL student who has never been told that English has a tense called "present perfect" has no schema for this concept. What they do have are the schemas that they have constructed in Arabic for years which address temporal reference differently and do not have direct English Language-equivalents (Ellis, 1997). Therefore, even when their teacher gives them a clear explanation of the present perfect, there is no way for them to have developed a cognitive level of understanding of this concept because they have no cognitive structure to store this concept except through purely phonological means, they can repeat the explanation, write down the examples on their worksheet, etc., but they do not yet have an understanding of it on an operational level because an understanding of a concept requires developing a schema and developing a schema can never occur by just hearing an explanation. A schema can only be developed from instances of the concept; it requires the active engagement of the student with instances of the concept followed by the assimilation/accommodation process from which the schema will be positively reconstructed (Piaget, 1952).

Piaget's (1952, 1970) concept of equilibration offers a general account of cognitive disequilibrium as the engine of learning, the state created when existing schemas fail to accommodate a new experience, prompting the cognitive effort through which new schemas are constructed. This mechanism is not specific to any particular instructional medium; it operates whenever a learner, working with a teacher, a text, a peer, or any other source of input, encounters material that existing understanding cannot fully explain. Piaget's critique is directed instead at instructional environments, regardless of the medium employed, that are designed to eliminate confusion and produce seamless, errorless performance, which he argues are optimally structured to prevent learning rather than support it.

This principle informs ASIM's design in that each of its three phases was deliberately structured to preserve, rather than eliminate, moments of productive disequilibrium: Phase 1 exposes the student to unfamiliar language for independent discovery; Phase 2 requires the evaluation of AI-generated content for error or limitation; and Phase 3 requires judgment against standards not yet fully internalized. The discomfort of not understanding is treated, in accordance with Piaget's framework, as a necessary condition for learning rather than an obstacle to be removed, a design choice that could, in principle, be realized through other instructional media, but which the present study operationalizes specifically through AI-mediated interaction for the practical and structural reasons discussed in Section 2.9.

Piaget's (1970) work extends beyond a simple distinction between verbal and operational knowledge to a deeper claim about the relationship between language and thought: genuine linguistic competence requires the learner to think through language, to actively construct meaning from what is read or produced, rather than merely reproduce its surface form. What Piaget termed 'verbal learning' describes precisely the failure of this process: words or

formulas are learned and reproduced without the underlying cognitive engagement that would make reading, writing, or speaking a conscious, meaning-making act rather than a mechanical one.

This distinction illuminates a pattern well documented among Palestinian EFL learners, who frequently demonstrate strong formal knowledge of grammar and vocabulary yet struggle to deploy that knowledge spontaneously to construct or interpret meaning in authentic communicative situations (Al-Ureimi, 2021; Aqel, 2002). Read through Piaget's framework, this is not a knowledge gap but a thinking gap: students have acquired linguistic forms without the operational, meaning-directed engagement that would make those forms usable.

Closing this gap requires tasks that compel the learner to think about meaning, not merely recall form, and it is on this specific requirement, rather than on any inherent property of AI itself, that ASIM's design rests. When a student must formulate a query to obtain a specific meaning, or must judge whether an AI-generated text achieves a particular communicative purpose, the task demands exactly the kind of active, meaning-directed cognitive engagement that Piaget identified as the condition for genuine language competence. The same demand could, in principle, be created through other tasks, a teacher's Socratic questioning, a carefully designed writing prompt, a peer-evaluation exercise, but ASIM operationalizes it specifically through AI-mediated query and evaluation because doing so allows this kind of meaning-directed engagement to occur repeatedly, individually, and on demand, at a scale that a single teacher cannot provide to a class of 30–40 students (see Section 2.9). Table 2.2 illustrates how Piaget's key theoretical principles relate to the specific problem each principle identifies with regard to traditional EFL teaching and ASIM's architectural solution to the issue at hand.

Table 2.2.A: Piagetian Principles, the Problem in Conventional EFL and ASIM's Architectural Response.

Piagetian Principle	Problem in Conventional EFL	ASIM's Architectural Response
Active construction over passive reception	Students listen, copy, and complete gap-fills; the teacher performs the cognitive work of organizing and presenting knowledge while the student reproduces it	All three phases structurally require the student to act: formulate queries, evaluate responses, direct AI output, and reflect on learning. Passive reception is architecturally impossible.
Assimilation and accommodation	Uniform instruction presents the same input to all learners regardless of their existing schemas, maximizing assimilation without creating productive accommodation	Each student's AI query is formulated against their own prior knowledge, ensuring responses challenge rather than merely confirm existing understanding and trigger genuine accommodation

Table 2.2.B: Piagetian Principles, the Problem in Conventional EFL and ASIM's Architectural Response.

Equilibration: disequilibrium as driver	Instruction optimized for clarity and correct-answer production systematically eliminates the cognitive friction that equilibration theory identifies as learning's primary engine	Phase 1 gap-identification, Phase 2 critical AI evaluation, and Phase 3 independent quality judgment each place the student in productive disequilibrium from which new schemas must be actively constructed
Concrete operations before formal abstraction	Grammar rules and vocabulary definitions are presented abstractly before the learner has encountered the linguistic phenomena they describe in authentic communicative use	Phase 1 places the student in authentic AI-mediated language interaction before any formal classroom instruction, providing the operational experience that Piaget identifies as prerequisite to genuine abstract understanding
Rejection of verbal learning	Test performance based on reproduced formulas creates the illusion of understanding; students can produce correct answers on tests without possessing the operational knowledge those answers purport to demonstrate	Phase 3 requires the student to demonstrate operational knowledge — to evaluate, explain, and judge AI-generated language quality — an activity that cannot be completed through formula reproduction

2.2.3 Vygotsky: The Social Conditions of Constructive Learning

Piaget's constructivism provides a powerful account of the internal mechanisms by which knowledge is built — the assimilation-accommodation cycle, the role of disequilibrium, the progressive construction of operational schemas (Piaget, 1952). However, very little is said about the social conditions to support the greatest current productivity of internal processes above, which is compensated by the social constructivism of Vygotsky (Vygotsky, 1978). While Piaget explains the internal workings of the learner and knowledge built by that learner, Vygotsky focuses on the relational and cultural factors that will best assist in building such knowledge. In regard to instructional design, these two frameworks complement each other when they provide a complete understanding of the instructional needs, both at the cognitive

level (Piaget) as well as the conditions that must be present in the surrounding environment (Vygotsky) to promote cognitive functions; creating a synthesis of ideas that is the foundation of current work on effective instruction (Bransford et al., 2000).

The most educationally influential concept in Vygotsky's framework is the Zone of Proximal Development (ZPD), defined as the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers (Vygotsky, 1978). The ZPD is not merely a description of what the learner cannot yet do alone; it is a positive characterization of the developmental frontier — the cognitive ground that is most responsive to instructional support, most generative of genuine development, and most likely produce the kind of durable understanding that is transferred into new contexts (Bransford et al., 2000). Instruction pitched below the ZPD confirms what the learner already knows. Instruction pitched above the ZPD produces frustration without development. Only instruction pitched within the ZPD — challenging enough to require support, manageable enough to be achievable with appropriate assistance — produces genuine cognitive growth (Vygotsky, 1978; Bransford et al., 2000).

The consequences of the structural implications of the ZPD for Palestinian English as a foreign language (EFL) instruction are far-reaching. They represent a fundamental limitation to traditional conversations occurring in a classroom setting; there are no skills or times that a teacher can employ to empower them to work effectively with the needs of their students. Specifically, a teacher working with 30 - 40 students with many differing abilities will not be able to diagnose which students are learning in their ZPD during a forty-minute period (Thaher, 2005). This presents a physical and logistical impossibility for that teacher. Whole group instruction for large educational systems generally limits teachers' ability to meet the individual ongoing development needs of their students. A Vygotskian perspective points out that each student operates within their own zone of proximal development ZPD, and, consequently, teaching and learning using one set of standards may not serve to support all students at the same time. In this regard, the Palestinian context reflects a broader structural issue of large class size rather than addressing deficiencies in teachers' competencies (Thaher, 2005; Bransford et al., 2000).

Understanding how AI might address the ZPD limitations of large-class instruction requires revisiting Vygotsky's (1978) concept of the More Knowledgeable Other (MKO): the person or agent whose greater competence enables the learner to operate within their ZPD. Vygotsky's own definition of the MKO was deliberately broad — the role could be filled by a teacher, a more capable peer, or, as later scholars have extended the concept, any cultural tool or agent capable of providing responsive support (Vygotsky, 1978; Wertsch, 1985); Vygotsky did not, of course, anticipate generative AI as a possible instantiation of this role, as the technology did not exist in his lifetime.

Building on this broad definition, the present researcher argues that generative AI represents a structurally distinct instantiation of the MKO, one whose defining advantage lies in its capacity for individualized, real-time responsiveness, a capacity that a single teacher managing a class of 30–40 students cannot replicate for each learner simultaneously (Thaher, 2005). Where a teacher must calibrate explanation to an approximation of the class's collective level, AI can, in principle, tailor its response to what an individual student's specific question reveals about that student's current understanding, offering differentiated support to a struggling and an advanced learner asking about the same grammatical feature.

It is this reasoning, not any claim originating in Vygotsky's own work that motivates ASIM's use of AI as an MKO calibrated to each learner's ZPD: a design choice the researcher considers a principled structural response to a pedagogical constraint that conventional Palestinian EFL instruction, given prevailing class sizes, cannot resolve through a human teacher alone.

The theory of psychological tools proposed by Vygotsky (1978), contributes to a more comprehensive understanding of how ASIM's use of AI in language learning creates change. According to Vygotsky's framework, a psychological tool is a sign or device that has been produced within a meaningful culture that serves to mediate and change the thought processes that the psychological tool is used for, i.e. a psychological tool does not simply assist in performing existing thought processes but instead restructures them. Vygotsky's perspective places language as the primary example of a psychological tool, in that it not only conveys a person's ideas but it also provides a means by which they may create and organize their ideas. Therefore, when AI is used by ASIM as an intermediary tool for the purpose of language learning, AI not only provides language information but acts as a responsive dynamic channel for qualitatively transforming the way a student engages cognitively with English. A student who composes a natural-language question to an AI, processes the AI's response, rewrites the natural-language question, and evaluates the response to the rewritten question is engaging in a type of cognitive engagement with a target language that is not typical of a traditional language-learning exercise. The student is conducting active, dialogic cognitive work in their target language as a means of thinking with the language rather than simply describing it. According to Vygotsky's framework, this type of linguistic-cognitive accomplishment is the most meaningful and productive way in which language and cognition can be developed.

2.2.4 The Constructivist Core of ASIM's Three Phases

The theoretical underpinnings provided by both Piagetian & Vygotskian constructivism allow an unambiguous articulation of the rationale/commonality between the phases of ASIM (three-phase model). The analysis further demonstrates that there is a systematically rather than randomly distinct form of progression through the three-phase model representing cognitive

engagements that build upon one another to achieve progressively higher levels of cognitive participation and meaning-making.

Phase 1 instantiates the Piagetian requirement for active, self-directed engagement with material at the learner's own developmental frontier (Piaget, 1952). The student's task of formulating an AI query requires them to perform the cognitive work of identifying what they do not yet understand — the kind of self-diagnosis that triggers the disequilibrium Piaget identifies as learning's engine (Piaget, 1952). Subsequently, the AI's response during Phase 1 then consists of the ZPD-calibrated MKO support that Vygotsky describes as the condition necessary for development at the learner's edge; responding at the same developmental level as the learner's query and adjusting in real-time to the learner's follow-up questions, reformulations and other forms of learner-generated input (Vygotsky, 1978).

Phase 2 distributes MKO support between two sources, the AI and the teacher, each addressing a different dimension of the learner's need. The researcher's rationale for this distribution is that a teacher no longer required to correct every grammatical error individually is freed to focus MKO support specifically on the dimensions AI cannot reliably assess, argumentation, rhetorical organization, and content development, while AI addresses linguistic accuracy at a pace and scale no single teacher could sustain across 30–40 students. Whether two coordinated sources of MKO support produce a stronger learning outcome than a single well-calibrated source is not something Vygotsky's original theory specifies; it is a design hypothesis the researcher advances on the basis of the complementary, rather than overlapping, nature of the two forms of support, to be examined empirically through the effectiveness data reported in Chapter Four.

Phase 3, in the researcher's application of Piagetian theory, is designed to require the demonstration of knowledge through performance rather than through verbal reproduction (Piaget, 1952), a distinction Piaget drew in relation to learning generally, not in reference to any particular instructional model. A student who evaluates the quality of AI-generated language, identifies its strengths and weaknesses, and communicates that evaluation to a professional assessor is engaging in the kind of operational demonstration Piaget distinguished from simple recitation. The researcher interprets this as evidence, consistent with Piagetian theory, that Phases 1 and 2 are more likely to have produced genuine constructive learning rather than the illusion of learning that Biggs and Tang (2011) describe as a limitation of assessments relying on reproduced knowledge, though this interpretation, like the broader effectiveness of ASIM, remains a claim to be substantiated by the empirical findings presented later in this dissertation, not a conclusion established by theory alone.

The epistemological basis for ASIM is provided by constructivism (Piaget, 1970; Vygotsky, 1978), since constructivism provides a theoretical explanation of the conditions necessary for true learning (which differs from the rote memorization and verbal reproduction that are

traditionally dominant in Palestinian classes of English as a Foreign Language) to occur. The work of Piaget makes clear that ASIM must be formulated to require active engagement of learners' cognitive processes at all three levels; that productive disequilibrium (i.e., cognitive conflict) must be regarded as a design element of ASIM and not as a design flaw; and that operational, not verbal, evidence of understanding must be used to demonstrate true competency in Phase Three. The work of Vygotsky illustrates how the use of artificial intelligence in ASIM as a "Mediating Knowledge Object" calibrated to the individual Zone of Proximal Development of each learner overcomes one of the structural limitations associated with large classes of students and cannot be overcome with other forms of pedagogy; and that the use of artificial intelligence as an interactive media tool fosters higher levels of cognitive engagement with English than do mere exposure or controlled practice. Collectively, the work of Piaget and Vygotsky provides theoretical support for the assertion that the three-phase structure of ASIM does not represent an innovative new approach to teaching English, but that it represents a theoretical necessity by demonstrating what constructivism indicates are the conditions under which true learning of a language can occur.

2.3 Self-Directed Learning Theory

The basic orientation for all of ASIM is Self-Directed Learning (SDL) which is not simply a single strategy but also provides the theoretical grounding for teaching and learning languages with the use of GAI. By providing unique opportunities for students to engage in personalized asynchronous (on-demand) interactions with the AI system, GAI creates an environment where the learner can be truly self-directed in their language learning.

Self-Directed Learning (SDL) represents a major paradigm change from teacher-centered to learner-centered education. Effective learners act independently in diagnosing needs, developing goals, identifying resources, implementing strategies, and evaluating their outcomes, just as with other typologies of Learning (Knowles, 1975). For ASIM, SDL is more than a single "method of delivery"; it is the foundational orientation for ASIM as a whole. This theoretical framework will support the GAI-enhanced SDL process, thus enhancing the language education experience.

2.3.1 Knowles' Andragogy and Self-Directed Learning

According to Malcolm Knowles, "Self-directed learning (SDL)" is an approach in which individuals take the initiative with or without assistance from others, to determine their educational needs, set their educational goals, find the appropriate resources to achieve those goals, determine how to achieve the goals and decide the extent to which they achieved their objectives (Knowles, 1975). This ability shows that learners have autonomy over their own education.

Additionally, Knowles explained the differences between adult learning (andragogy) and teaching children (pedagogy) when he demonstrated that adults can succeed in learning if they have the chance to control their own learning process. While Knowles originally focused on adult learners, more recent research shows that young people (11th graders) can also learn through self-directed methods with proper scaffolding and support from mentors (Guglielmino, 2013).

In Knowles' model of SDL, he has defined four interrelated stages are manifested:

1. Self-Diagnosis: This stage involves identifying what the learner's needs are.
2. Goal Setting: At this stage the learner will set a specific goal for themselves by writing down their goal.
3. Resource Identification: During stage three the learner will look for and obtain the necessary resources to achieve their goal from the previous stage.
4. Outcome Verification: The last stage includes evaluating how successful the processes were in carrying out the stated goal(s).

Application to ASIM

ASIM's three-phase structure, AI-Guided Discovery, Human-AI Co-Creation, and Student-as-Teacher Mastery, was designed as a direct operationalization of Knowles' (1975) four-stage SDL cycle: self-diagnosis, goal-setting, resource identification, and outcome verification. Rather than treating these four stages as a single undifferentiated competency to be developed all at once, ASIM distributes them deliberately across its three phases, on the theoretical grounds that each stage of Knowles' model places a distinct cognitive demand on the learner and therefore benefits from a dedicated instructional space.

Phase 1 operationalizes Knowles' self-diagnosis stage. Where conventional instruction positions the teacher as the one who identifies a learner's weaknesses through assessment, Phase 1 requires the learner to perform this diagnostic function independently, converting passive reception of content into active self-assessment — the theoretical prerequisite, in Knowles' model, for any subsequent autonomous goal-setting.

Phase 2 operationalizes Knowles' goal-setting and resource-identification stages jointly, on the theoretical grounds that these two stages are naturally interdependent: a learner cannot

meaningfully select resources and strategies without first having articulated a specific goal, and the act of pursuing a self-set goal through available resources is, in Knowles' framework, what distinguishes self-directed practice from teacher-assigned practice. The shift in the teacher's role from director to coach reflects this same theoretical premise; Knowles' model requires that the learner, not the teacher, retain ownership of the goal being pursued.

Phase 3 operationalizes Knowles' outcome-verification stage. Rather than relying on an external judgment issued by the teacher, the learner verifies their own mastery through performance, a design consistent with Knowles' emphasis on self-directed learners as those capable of evaluating the extent to which their own goals have been achieved, rather than remaining dependent on external evaluation for this determination.

A full account of the specific activities, instruments, and procedures through which each phase is implemented is provided in Chapter Three and Appendix 1; the purpose of the mapping above is to establish the theoretical correspondence between Knowles' four-stage model and ASIM's three-phase architecture, not to describe the operational details of implementation.

Solitary in comparison with Self-Directed Learning

In his work, Knowles (1975) points out that though SDL means being in control of your own learning, it does not mean that you are learning in isolation. Self-directed learners have a learning responsibility to take advantage of teachers, peers, resources such as print and electronic materials, and technology to help them learn. This difference is critical to ASIM, where students who are self-directed use their own choice regarding the timing of their consultation with AI, teacher assistance, and peer group cooperation, but they do so as part of a community; hence, they maintain a locus of control while participating in a collaborative learning environment.

The clarification from Candy (1991) on SDL is that it can be viewed in two ways, as a self-directed learner will experience SDL both as a process (what they do) and as a goal (what the educational system will help them achieve). ASIM represents both views, allowing learners to experience a formal process where they practice SDL and, concurrently, develop the skills and disposition necessary to continue their self-directed learning throughout their lives.

2.3.2 Learner Autonomy

According to Holec (1981), learner autonomy is defined as "the ability to take responsibility for one's own learning" (p. 3). He also states that learner autonomy can be seen as a capability and a right. Traditional education denies learners this right by creating an educational structure in which teachers effectively make all decisions regarding content, pace, and assessment.

Benson (2001) developed a model of autonomy based on Holec's work, which consists of three interrelated dimensions. These dimensions include technical autonomy (control learners over practical aspects of their education), psychological autonomy (self-efficacy beliefs, intrinsic motivation to achieve, metacognition, and willingness to take on new challenges), and social autonomy (the importance of building supportive social environments). In his revised publication (Benson, 2013), Benson (2013) stated that effective education must intentionally create opportunities for all three autonomy dimensions to develop together and that autonomy is not a fixed characteristic; rather, it is an evolving ability that is influenced by educational experiences.

Little (1991) remarks that autonomy is something to be acquired; to risk too much –too soon is to expose unskilled and unconfident learners to certain failure. This idea is reflected in the three-phase structure of ASIM, where responsibility is gradually shifted from teacher to learner, across guided, supported and independent stages.

As Benson (2013) points out, the notion of autonomy is particularly compelling in language education where "authentic practice, personal investment, and out-of-class exposure are crucial." Murase (2015) adds some evidence, that highly autonomous learners exhibited significantly greater gains in English proficiency over two years than less autonomous learners receiving the same instruction in class — and the difference is attributable to continued practice and resourcefulness out of class, but also to what she terms motivation catalysts.

2.3.3 Self-Regulated Learning

Autonomy supports an individual's ability and right to manage their own learning whereas self-regulated learning (SRL) refers to the cognitive/metacognitive processes that demonstrate that control. Zimmerman and Schunk (2011) define SRL as a process that is dynamic and cyclical, in which learners use their mental abilities to produce task-related skills.

The Cyclical Model of SRL created by Zimmerman (2000) contains three phases that are dependent on one another. The forethought phase is when learners develop specific goals, utilize metacognitive knowledge in strategy selection, and establish motivation (self-efficacy & task interest). The performance phase is the application of the previously developed strategies; monitoring of progress as it relates to the task being performed, and modifying strategy when the desired outcome is not achieved. The self-reflection phase is where learners evaluate their results of that task compared to the forethought goals, as well as their own self-reaction, which then feeds into the forethought phase.

Panadero (2017) examined different SRL models and found that metacognition, motivation, and action are common characteristics. Schunk and DiBenedetto (2020) showed that

purposeful SRL training improved achievement by 23% more than control groups, and that having structured opportunities as well as direct instruction in self-regulation is needed.

ASIM applies these theories in a progressive way: phase 1 develops forethought using AI based goal-setting and previewing; phase 2 develops performance monitoring during the performance phase via repeated use of feedback from peers and subsequent revisions based on that feedback; and phase 3 develops deep self-reflection through teaching, which helps students determine and express the limits of their current knowledge.

2.3.4 Role of Technology in Self-Directed Learning

Self-directed learning opportunities have greatly increased as technology has erased many of the traditional barriers, including time, location, and expert assistance (Luckin et al., 2016). Additionally, mobile devices have also expanded these opportunities; for example, Lai and Zheng (2018) demonstrated that mobile language learners had higher levels of autonomy because they utilized short, opportunistic moments to practice rather than relying on pre-scheduled study times.

GAI represents a qualitative leap forward via dialogic interactions, allowing learners to ask clarifying questions, request personalized support, and receive answers that are scaled to their individual levels of proficiency. Al-Atrash (2024) reported that Palestinian EFL learners using ChatGPT demonstrated increased goal-setting behaviors, greater levels of self-regulation, and lessened dependence on their instructors. Prompt engineering itself was identified as a metacognitive process (Ali et al., 2023) as students create effective prompts while simultaneously learning to assess learning needs and evaluate responses. Furthermore, Al-Twijri and Alghizzi (2024) found that learners participated in AI-enabled instruction reported lower levels of foreign language anxiety through increased motivation; both of which are necessary foundations for autonomous learning.

2.3.5 Synthesis: SDL as ASIM's Foundation

The foundational theory behind ASIM is SDL. ASIM involves moving learners from the role of passive receiver to that of an active participant. ASIM transfers cognitive and evaluative responsibility to the learner in three progressive phases. Phase 1 involves AI assistance and exploration; Phase 2 involves independent production with support; and Phase 3 involves demonstrating mastery through teaching.

Metacognitive development is made explicit and structured through things like Self-Diagnosis Reports, Learning Strategy Logs, and Teaching Reflections, rather than leaving that to chance. In addition, the teacher's role also changes from being a source of information to being a designer of challenges and coach in self-regulatory strategies.

This framework aligns with Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001), which correlates with the three progressive phases of ASIM; cognitive levels of Remembering & Understanding in Phase 1; Applying & Analyzing in Phase 2; Evaluating & Creating in Phase 3. A critique made by Biesta (2022) of the phenomenon referred to as "learnification" indicates authentically educative processes can only occur when learners are exposed to environments external to their immediate preference base. The teacher will facilitate their learners' experiences of the latter by designing structured challenges that would not occur with the sole application of autonomy.

2.4 Learning by Teaching: The Theoretical Rationale for ASIM's Phase 3 Role Inversion

2.4.1 The Problem That Learning by Teaching Addresses

Consider a Palestinian EFL learner who has worked through Phases 1 and 2 of an ASIM learning cycle. Through 'AI-supported inquiry' the learner has explored the present perfect tense and received 'scaffold assistance' to help move across the Zone of Proximal Development (ZPD) before taking part in classroom discussions to explore conceptual 'gaps' arising from inquiry. The student then drafts a letter in the target language before getting AI-generated in-language feedback about the letter they wrote and then revising the letter by incorporating their classroom teacher's assistance or conceptual coaching. This learner has been supported, engaged and active as they went through this process. They have done more in this specific situation than any traditional EFL lesson can provide. These are not theoretical problems; moreover, evidence provided by Bjork (1994) demonstrates that experiences that appear to produce knowledge — like receiving explicit instruction, performing the practice of receiving guidance, or getting instantaneous error correction — will often create the illusion of knowing. While a person may perform adequately on their demonstration day, their learning within this context will not transfer to other contexts once that context is removed. Roediger and Karpicke (2006) further demonstrated that, while being exposed repeatedly to the material and being actively involved in the learning process do not provide assurances of durable retention, what does provide assurance is the process of accessing and performing the "knowledge" while under true cognitive demand. The question is: how deeply have they learned? How long will that learning last? How readily will they be able to deploy it in new situations (such as completing an essay on a prompt that they did not expect, going through a job interview or talking with an individual that they did not anticipate talking to)? Furthermore, how will they know (prior to an examination) whether they have achieved durable learning and will not lose it?

Conventional education assessments cannot answer these questions until the test occurs, and the test reveals whether or not the student truly understood or was just superficially familiar with the material (Nestojko et al., 2014). However, there is another form of cognitive activity

that provides evidence regarding a student's understanding of something before they take a test; provides diagnostic information about what the student truly understands from what they are merely familiar with; and helps the learner create deeper and longer-lasting learning than will be gained by studying the same amount of time (Duran, 2017; Fiorella & Mayer, 2013). A learner who prepares to teach something and then teaches it engages in cognitive activity that no other learning activity provides; that provides opportunities to reorganize knowledge; that brings to light gaps in understanding; that requires evaluation of one's own understanding using evaluative processes (Roscoe & Chi, 2007); and that provides socially accountable preparation for teaching others rather than preparing only for oneself (Benware & Deci, 1984). This phenomenon can be explained by the Learning by Teaching Theory which is the primary reason for ASIM Phase 3 role reversal.

2.4.2 Teaching as the Most Demanding — and Most Effective — Learning Strategy

According to their research into generative learning strategies, Fiorella and Mayer (2015) provide the most thorough and empirically supported overview of Learning by Teaching as an educational phenomenon. Fiorella and Mayer use this extensive body of empirical evidence (drawn from a wide range of studies) from a number of different learner populations and across a variety of different content areas to conclude that Learning by Teaching is both the most cognitively demanding and most consistently effective of the eight different learning strategies they analyzed. They also concluded that the effects of Learning by Teaching cannot be explained by additional practice, increased motivation, or increased attention, but rather are a result of specific cognitive mechanisms that are uniquely activated by the act of teaching. Duran and Topping (2017) compliment Fiorella and Mayer's cognitive study by providing a systematic review of Learning by teaching research in a classroom setting, exploring how the motivational, relational, and metacognitive aspects of teaching are related to and enhance the cognitive mechanisms identified by Fiorella and Mayer.

Both frameworks share the main belief that the act of teaching, which is comprised of both preparing to teach and then teaching, will create better learning outcomes than if an equivalent amount of time and effort were spent studying in a traditional manner. The difference is more than marginal. As shown by Duran & Topping (2017), a review of experimental studies clearly demonstrated that students in the teaching role benefit greatly from peer-tutoring and learning by Teaching (LbT), regardless of age, subject area or context. Investigating why these benefits occur requires consideration of the cognitive processes involved with the teaching function — cognitive processes that are systematically ignored in traditional settings.

2.4.3 The Six Mechanisms: Why Teaching Produces the Deepest Learning

Synthesizing the accounts of Fiorella and Mayer (2015) and Duran and Topping (2017), six distinct cognitive and motivational mechanisms through which the teaching role produces

superior learning outcomes can be identified. They both synthesized their findings into six distinct mechanisms. Each mechanism works independently and cumulatively, and it is their combined effect that creates the superior outcomes consistently found in LbT literature.

Mechanism 1: Knowledge Reorganization for Explanation

Every student who has studied material has organized their knowledge from the perspective of a consumer: they know what they know, in the order they encountered it, structured around the features that were salient to them as learners. This organization is adequate for recognizing correct answers on tests but inadequate for teaching, which requires a fundamentally different cognitive structure, a structure organized around what the listener needs to understand, in what sequence, at what level of detail, and connected to what prior knowledge. The cognitive work of reorganizing personal knowledge into this teacher-oriented structure is demanding and unavoidable, and it produces a qualitatively different representation of the material than any amount of re-reading or additional study can create (Roscoe & Chi, 2007). It is also one of the primary mechanisms by which teaching produces learning benefits for the teacher rather than merely the learner.

Mechanism 2: Gap Identification through Teaching Demands

Fiorella & Mayer (2015) describe the problem of what they call the "illusion of knowing" as one of the most frequently occurring (and detrimental to education) elements of traditional study - where learners believe they are knowledgeable about a subject when they may only have superficial knowledge (e.g., test, recognizes or recalls (correct) answers). The "illusion of knowing" cannot continue once one attempts to provide an explanation of that principle to another person (e.g. answer questions that they had not anticipated, provide examples that illustrate the principle without merely repeating previously, connect the principle to the existing knowledge base of the listener) because the requirements of the cognitive demands required to provide explanations will eliminate any "illusion of knowing" (Nestojko et al., 2014). Therefore, the gaps in one's knowledge that is not apparent before teaching will become readily apparent once the cognitive demands are placed on both the teacher and the learner. These gaps are not subject failures; rather, they are, on the contrary, the most reliable or precise diagnostic indicators of actual levels of knowledge and understanding that may be presented in any educational context through teacher-centered pedagogical practices, thus, creating targeted disequilibrium that Piaget exemplifies as the driving force behind the learning process.

Mechanism 3: Deeper Encoding via Retrieval and Re-Expression

Cognitive scientists have identified the testing effect (or retrieval practice effect) through the years and decades of studies related to memory. These studies show that when you retrieve information from your long-term memory, you will encode that information far deeper than if you simply re-read or re-expose yourself to the same material (Roediger & Karpicke, 2006). Teaching a learner is the ultimate form of retrieval practice because it forces the learner to recall and reconstruct material from memory in an entirely new way (example, explanation, evaluation) rather than just recognize the material when you show it to the learners. The cognitive effort of reconstructing the material from memory, as documented by Duran & Topping (2017), is one of the most significant predictors of long-term retention in the literature related to Learning-By Teaching (LbT) and occurs regardless of whether you produce learning through your teaching.

Mechanism 4: Development of Evaluative Authority

In traditional EFL instruction, the teacher holds all evaluative authority. In this case, the teacher determines what is "correct" and what is "not correct;" provides feedback about the student's work; and assigns grades. In this structure, the student assumes the role of a recipient of evaluation, as opposed to acting as an independent agent of evaluation. Therefore, under this structure, the student does not establish the internal standards required to evaluate both their own work as well as the work of others (Benware & Deci, 1984). Through the construction of the teaching role, the reverse of this structure takes place — the learner is now not only getting evaluated and receiving feedback on their work, they are also creating the criteria that they will be measured against and for their own work than only to depend on outside evaluators. By developing evaluative authority through the Teaching process, the learner develops the most valuable trait of the truly competent language user: the ability to determine whether their production is sufficient to achieve the desired goal and to determine the changes necessary to improve their production. The phase of ASIM designed to implement this through the learner evaluating the quality of a language produced by AI (Phase 3) is directly tied to the establishment of the evaluative authority that is required for lifelong language development.

Mechanism 5: Social Accountability and Motivational Activation

According to Duran and Topping (2017), social accountability is one of the strongest motivators in the role of teacher. The student teaching a unit of material (i.e., explaining, defending, and anticipating questions from an evaluator) prepares for that role differently and more thoroughly than a student who studies the same material privately for a test (Benware &

Deci, 1984; Nestojko et al., 2014). The social aspect of the teaching role creates a social accountability that is greater than simply meeting personal performance expectations; the student is accountable not only for knowing the material, but also for being able to share it with another person, and therefore their failure to do so is a public failure versus a private failure. The result of this social accountability is not just increased amounts of study time, but, also, an increase in the quality of study time. Study time also, as a result of social accountability, involves much more effort, reflection, and evaluative engagement than studying in a private setting can achieve.

Mechanism 6: Metacognitive Activation

The final motivational mechanism is the student's ability to activate metacognitive monitoring that is unique to the role of teacher. When the learner studies for a personal purpose, the learner's metacognitive monitoring is primarily at the first - person level (i.e., do I understand?). The teacher role requires a third-person metacognitive monitoring perspective (i.e., do I have enough knowledge to be able to teach this to someone else?) Will my students be able to understand my explanation? Are there aspects of the topic that I will be unable to defend if my students have questions? (Roscoe & Chi; 2007). This third-person metacognitive monitoring perspective is more demanding, more visible, and closely resembles that of the more sophisticated language user who constantly monitors his or her production for adequacy, correctness, and effect (Flavell, 1979). In Phase 3 of ASIM, students will be required to assume the same metacognitive monitoring perspective that an effective language user would have toward AI generated text, and therefore, in completing this task, students will continue to develop their metacognitive monitoring abilities to a level that is comparable to that of an effective language user.

2.4.4 ASIM's Phase 3 as a Realization of All Six Mechanisms

The researcher's design rationale is that ASIM's Phase 3 role inversion, in which the student teaches the AI, corrects its intentional errors, questions it about complex issues, and submits their evaluations to the teacher, is structured to activate all six mechanisms simultaneously, rather than sequentially as in typical peer-tutoring contexts. Preparing for and completing Phase 3 is designed to require students to reorganize what they learned in Phases 1 and 2 from a consumer perspective to a teaching perspective (Mechanism 1). Through this process, students are expected to identify gaps in their knowledge that were not illuminated by the previous phases (Mechanism 2). The task is designed so that students retrieve and express session content in an alternate way, a process that retrieval practice research associates with deeper encoding (Mechanism 3). Phase 3 further requires students to develop and publicly apply criteria for evaluating AI output, thereby practicing the evaluative authority considered necessary for independent language development (Mechanism 4). Because students prepare and perform in front of a teacher, the professional evaluation involved is intended to create a

genuine sense of social accountability for their work (Mechanism 5). Finally, the phase is designed to position students in a third-person metacognitive perspective, that of an evaluator rather than a student, which is considered necessary for the expert use of language (Mechanism 6). Whether Phase 3 activates these six mechanisms as intended, and to what extent, is a question addressed empirically in Chapters Four and Five, rather than one resolved by theoretical design alone. The following diagram illustrates these six mechanisms collectively, as the theoretical foundation upon which the practices of this phase are built.

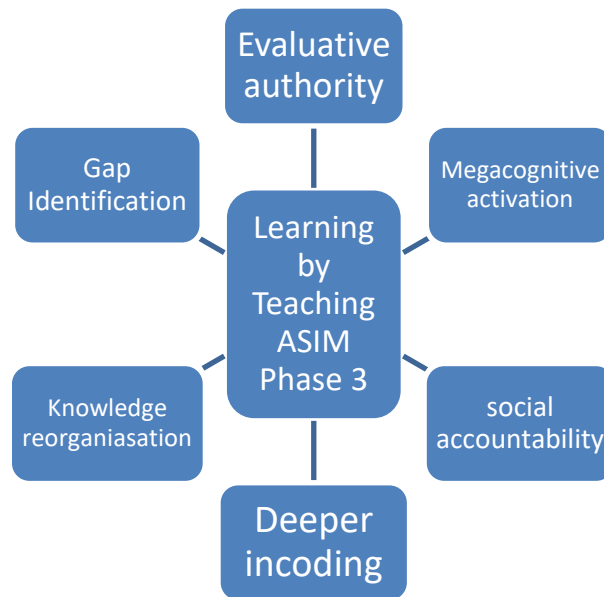


Figure 2.1: Learning by Teaching: Six Mechanisms and Their Realization in ASIM Phase 3.

2.4.5 Why the AI Is an Effective Teaching Target

The Phase 3 design of ASIM raises a key point to consider: if the effects of Learning by Teaching stem from the cognitive demand of explaining to another student who is in need of learning, do the same positive outcomes occur when another student is replaced with an AI? Based on the theoretical analysis of the six mechanisms discussed in the previous section, these intended benefits are considered to be possible and will provide additional theoretical insights into this issue beyond what the question itself would seem to ask.

According to Fiorella and Mayer (2015), the LbT mechanisms can be traced back to the teacher's cognitive demands of delivering effective lessons for their learners. However, these demands come from limitations of giving students the proper explanations that allow for evaluation; the teacher does not base these judgmental criteria on the characteristics of the

students (the audience) they are going to teach. All 6 mechanisms above can be used in a lesson when a student creates a prompt that transforms the AI into the equivalent of a naive learner, evaluates the output of the AI to ensure it is appropriately challenged as the confused learner, explains the appropriate description of why the AI is confused and provides an appropriate explanation to assist the AI in resolving its questions of confusion, and then evaluates the overall quality of the teaching that the student provided to the AI; even if the AI never truly learns anything. The AI is a more demanding and more revealing interlocutor because the responses generated by the AI are based on the quality of how the student directed the AI (specified by how well the student crafted their prompt) to generating challenges; the AI can represent full understanding when the student formulates an inappropriate prompt by generating an output that does not require improvement. The AI will represent having not fully understood when the student provides an appropriately crafted teaching prompt and generates an output that indicated need for another level of challenge and/or limits the quality of the explanation. Thus, by these measures, students who utilize an AI rather than a peer to teach will provide a more transparent and accurate measure of what the student needs will not occur, since the peer may (due to social courtesy) represent an understanding when, in fact, they may not have understood the lesson.

Furthermore, the actual teacher's evaluative process adds a layer of human evaluation that cannot be covered by AI interaction alone: the teacher observes, has the transcript reviewed, and assesses the student's ability to use and gauge the AI product — judging how well the student understood, directed and evaluated the AI product, as opposed to how well the AI product performed. This teacher evaluation serves as the basis for Mechanism 5 (social accountability) and provides the criteria used to determine whether or not the student's authority to assess something (Mechanism 4) is based on professional standards, as opposed to solely individual standards.

2.4.6 Learning by Teaching and Language Skill Development

The importance of Language-by-Teaching (LbT) for the development of English as a foreign language goes beyond the general cognitive mechanisms discussed so far and includes features that are specifically relevant to learning a language. Teaching English requires a level of metalinguistic awareness on behalf of the learner that traditional English as a foreign language instruction might not expect. When teaching grammar to an AI simulator as if it were an EFL learner, the student will need to know not just how to use a particular grammar rule, but also what that rule is, what the reason is for having that rule, how it compares to other types of grammatical categories, and what types of mistakes can result from misusing that particular rule. Having this knowledge about language rather than just being able to use it is one of the best indicators of future development of high-level language proficiency because this type of metalinguistic awareness allows the learner to better self-monitor, assess, and purposefully

develop their own linguistic output; whereas with strictly intuitive, implicit knowledge this is usually not possible (Roehr, 2007; Roehr-Brackin, 2018).

In the Palestinian EFL classroom after 12 years of instruction in English as a foreign language, students recognize but do not produce correct grammatical structures freely in spontaneous situations (Hammad, 2016a); therefore, the metalinguistic engagement required of teachers during Phase 3 is not an added feature of instruction but instead a corrective measure. During Phase 3, students who have been using language for 12 years as consumers of language instruction (applying rules when prompted and filling gaps when instructed) become producers of language instruction for the first time and are required to explain grammar rules to others as well as identify and evaluate the accuracy of other students' grammatically-based work, thus shifting their experience from being a consumer of language instruction to being a producer of language instruction (at the metalinguistic level), which is, according to LbT literature one of the most significant and most reliable agents of authentic language competence development (Swain, 1985; Roehr-Brackin, 2018).

2.4.7 The Relationship between LbT and the Constructivist Foundation

According to Piaget (1952), Phase 3 teaching creates more intense and clearly-defined disequilibrium than any other ASIM phases, since the learner has established an understanding of a principle that they cannot explain and has, therefore, the clearest examples of their existing schemas and the most motivation to develop new schemas. According to Vygotsky (1978), the teacher's evaluation of the student's instruction in Phase 3 is the most personalized and ZPD-specific of any instructional interaction in the model: In order to help each learner move towards their cognitive/instructional advancements through progress of skill/knowledge, the instructor gives tailored support based on what the learner is currently learning about (i.e. their teaching performance).

Accordingly, the LbT framework further develops the constructivist explanation by defining the circumstances that allow constructive cognitive activity to operate at its highest levels of complexity and most generative capacity. The first two phases provide the learner with opportunities to constructively engage with new information, while the third phase affords an opportunity for that engagement to be solidified, evaluated, and reshaped as a substantial, transferable, and evaluative rich form of knowing, thus producing authentic language ability. These three phases will provide a complete constructivist learning cycle, beginning with initial (active) engagement with new knowledge and continuing through scaffold growth in the learner's ZPD until the learner engages in the most difficult form of constructive activity (i.e. teaching) a cycle of experience not supported by mainstream EFL teaching methods, as they consistently disregard the learner's role as a cognitive agent.

2.5 Cognitive Apprenticeship: How Expertise Transfers Across ASIM's Three Phases

2.5.1 The Problem: Invisible Expertise in Palestinian EFL Classrooms

Collins, Brown & Newman (1989) identify a paradox central to traditional methods of teaching languages as a major factor in the failure of modern education: that language teachers are themselves expert users of language whose professional knowledge is both great and nearly completely invisible to their students in a typical classroom setting. A competent teacher knows from experience how an expert writer chooses between two or more words that all seem equally suitable, how the structure of a given sentence will carry an argument more powerfully than an alternative, and what qualities separate a convincing formal letter from one that adheres strictly to the form but is not persuasive. Schön (1983) calls this type of professional knowledge "knowing-in-practice" tacit, embedded, and largely unarticulated competence which practitioners demonstrate qualitatively but rarely manifest explicitly outside of their own minds.

The only visible aspect of an EFL lesson delivered following conventional methods is the result of the teacher's work: a corrected sentence written on the board; a model paragraph taken from the textbook; and a score given to the student's essay for such things as content, usage, spelling, and punctuation, regardless of how satisfactory to the teacher the student's actual paragraphs might be. The process, the reasoning used by the teacher to find an error, what was done in thinking about how to correct the error, and the criteria used to judge that a paragraph has been constructed in an orderly manner, remains invisible. Rogoff (1990) contrasts the invisibility of the process with what occurs in traditional forms of apprenticeship, wherein a novice can observe a competent practitioner performing work in great detail, be given progressively complex tasks to assist in performing for the practitioner, and develop their own ability through first-hand exposure to the actual work of competent practitioners, rather than only the products created by them.

The situational problem is complicated further in the EFL setting by the intrinsic nature of language itself. While the hands of a craftsman or the calculations of an engineer can be seen by others, the reasoning processes used by skilled language users, how to select the appropriate register for the situation or how to construct an argument for someone specifically, for example, occur only inside a user's mind and leave no concrete evidence outside the actual language produced. Lave and Wenger (1991) argue that true competence will only come from being a participant in authentic practice communities, where the full range of rationale used by competent practitioners is available for observation by novice practitioners; and, conversely, the EFL classroom in which students learn according to the conventional EFL instructional model offers students only evidence of the finished products of the rationale of the language users, but does not afford students the ability to recognize the rationale behind the results.

Cognitive Apprenticeship theory was formulated to help fill this gap, between the knowledge of competent language users and what students have the opportunity to observe.

2.5.2 The Cognitive Apprenticeship Framework

The Cognitive Apprenticeship (CA) model, introduced by Collins et al. (1989) as a solution to the issue of invisible expertise includes six instructional approaches to making cognitive processes explicit and transferring expert knowledge to novice learners via a systematic progression of gradually more autonomous performance. The six approaches include three primary approaches for facilitating transfer from expert to novice (Modeling, Coaching, and Scaffolding) and the three additional approaches used to foster the learner's development of independent competence (Fading, Articulation, and Reflection).

Modeling , as Collins et al. (1989) define it, is when an expert performs a target skill as demonstrated through the visible thought processes behind that action, thinking out loud, sharing your decisions with the learner as you arrive at them, and showing what you do and why you do it. In the context of second language acquisition modeling can also be seen as the externalization of cognition on how an expert uses the language, why they chose this structure rather than another, why they chose this register of vocabulary, and how they organize the text. Coaching occurs when an expert observes the performance of a learner and provides them with feedback related to their performance of the target skill. This feedback is specific to the performance of the learner and the gap between the current performance of the learner and the performance of the expert. Hattie and Timperley (2007) in their major meta-analysis on feedback have found that coaching is effective by the provision of feedback that focuses only on the gap between the current level of performance and the level of expert performance. This is exactly what coaching does throughout the CA process. Scaffolding, as introduced by Wood, Bruner, and Ross (1976) in their discussion of expert-novice interactions, involves structuring the learning environment so that the learner is able to do tasks that they cannot yet do independently and providing them with sufficient support so that the learner is able to perform beyond their current capacity, but with the support of the coach. The process of Fading, decreasing the support of both the scaffolding and the coaching, allows for greater responsibility on the learner and moves them towards independent performance (Collins et al., 1989). Articulation requires the learner to make their reasoning explicit. In other words, the learner has to be able to explain how and why they are doing what they are doing. Chi, Bassok, Lewis, Reimann, and Glaser (1989) demonstrated empirically that learners who generate self-explanations while working through examples develop significantly greater comprehension of an example than learners who do not. Reflection, as described by Schön (1983) in his analysis of professional practice, involves the learner comparing the way they performed against the way an expert performed, or developing a personal level of performance. Reflection develops the self-regulation required to separate competent from expert performers.

Cognitive Apprenticeship's methods as a design framework for ASIM are strong due to their organization and systematic performance of tasks. There is an identifiable sequence of Model, Coach, Scaffold, and Fade which demonstrates a cognitive transfer principle of responsibility from the expert to the apprentice, where the expert will model the task to the novice, coach them through it, and gradually remove the coaching until the novice can do the task without any input from the expert. In addition, the established sequence of performance is not simply a function of a pedagogical preference; it is rather reflective of the cognitive nature of the competencies that are created through the transfer of expertise within the human domain; that is, all domains in which high proficiency are attained by humans.

2.5.3 The CA Sequence across ASIM's Three Phases

The three-phase architecture of ASIM directly and precisely maps onto the Cognitive Apprenticeship (CA) sequence, not merely as analogy but rather as a structural implementation of the two sequences. Each phase comprises a distinct stage in the CA sequence, and the transition from phase to phase enacts the Fading identified by CA theory as the mechanism for true expertise transfer.

Phase 1: (AI-Guided Discovery) is the Modeling Phase. In this phase, the AI's role is the same as the modeling function described in CA theory; it makes expert language reasoning visible by showing not only what is correct but also why, providing worked examples with explicit reasoning, and identifying the principles behind the demonstrated reasoning processes. The AI also responds to each student's query, modeling the expert approach for that query. The AI performs this modeling for each individual student, calibrated to the context of the query that the student presents, so that each student receives modeling that addresses their particular learning needs rather than modeling directed at the entire class as a whole. This kind of modeling is prescribed by CA theory but is impossible for a single human instructor to provide to thirty-five students simultaneously.

Phase 2: Human-AI Co-Creation which is the coaching and scaffolding phase of the process. The teacher's role is primarily that of a coach as defined by CA. The teacher observes the student's performance, identifies where they fall short in their understanding or production against the standard, and gives feedback only on those particular points, rather than having to re-teach the entire lesson, thus, the teacher is coaching the student at that point only on their specific area of difficulty. In addition, the AI will continue to provide scaffolding for the student in the form of immediate, individualized, on-demand support for the grammatical, vocabulary, and fluency aspects of the task in relation to what the student is still unable to accomplish independently. Together, the coaching provided by the teacher to the student in Phase 2 for deeper conceptual understanding and the scaffolding of the AI to the student for development of how to express themselves using language create a richer and more precisely

defined framework for supporting the student's learning than could be created independently by either the teacher or the AI.

Phase 3: Student-as-Teacher Mastery is the fading, articulating, and reflection phase. The teacher and the AI have now withdrawn their support, and the student now assumes total cognitive responsibility for their learning. Specifically, the student will teach the AI, correct the AI's errors (intentional), and evaluate the quality of the performance of the AI. The student who can perform these tasks represents the culmination of the CA process sequence: the student has internalized sufficient knowledge as an expert and evaluative criteria to represent the expert that is, they have made the reasoning done in Phase 1, the scaffolding done for them in Phase 2, and the gradual fading of that scaffolding in Phase 3 part of the student's own thinking.

2.5.4 The Problem of Invisible Cognition and the AI Solution

The implementation of Cognitive Apprenticeship in ASIM is significant because it provides a solution for the problem of invisible expert cognition identified by Collins, Brown and Newman (1989) as the major flaw in formal schooling. In a traditional foreign language classroom, the teacher's language reasoning is not visible to students, not because of a desire to keep it hidden but because the structure of the instruction does not provide a way to make that reasoning visible at the level of the individual. The teacher can model thinking aloud for the whole class but this cannot answer to each student's specific invisible gaps of understanding their individual needs for explanation. In contrast to this, the AI provides visibility to expert language reasoning through an answer to that student's specific question, at the level of detail that question requires and within the domain of difficulty that the question identifies. The way it provides this information is, in several ways, a more complete realization of the Cognitive Apprenticeship modeling method than would typically be able to be achieved through human expert instruction.

In summary, Collins, Brown and Newman (1989) theorized an understanding of how to use the principles of Cognitive Apprenticeship through the instructional sequence logic that ASIM's 3-phase architecture is built upon. Using a cognitive coaching model assists learners to increase their performance through an expert-based model of cognition (expert cognitive processes) by providing models to develop them for learners through training and guidance & increasing their capability through deliberate practice. Through the use of fading, clarification, and reflection; cognitive coaches progressively decrease this support. Therefore, the use of a cognitive methodology aligns with the learning phases of the ASIM System that include all three stages (educational) providing a defined, explicit model of instruction. The use of Artificial Intelligence as the primary modeler and scaffolder will resolve the structural limitation that has prevented Palestinian EFL educators from delivering individualized CA-style instruction at scale. However, regarding how the learner's relationship to technology

itself becomes an outcome of learning in addition to a pedagogical tool -- that question must be addressed through a different theoretical lens than the Cognitive Apprenticeship model of instructional design -- this is the issue that Section 2.6 of this paper will address using the TPACK framework and an extension of the framework that is student-centered, S-TPACK.

2.6 TPACK and S-TPACK: Relocating Technological-Pedagogical Integration from Teacher to Student

2.6.1 The Problem and the Framework

In the last few years, the introduction of technology into the classrooms of Palestinian EFL has greatly increased due in large part to the increase of smartphones, the access to the internet and the advent of more accessible AI tools. National authorities have connected over 70 schools in the West Bank to the internet in response to COVID-19, allowing greater access to digital tools by students (United Nations in Palestine, 2023 as cited in Abu Hashish & Al-Halimi, 2023). However, the mere presence of technology within a classroom does not inherently guarantee that students will receive the maximum amount of learning benefits through technology-enhanced instruction. In a study of 207 teachers and 276 students in 53 Palestinian public schools, Qaddumi, Bartram and Qashmar (2021) found that students stated that Information and Communication Technology (ICT) had only moderate impact on learning and students continue to struggle with their ability to conduct information research; therefore, it appears that students are using technology but do so without the necessary knowledge framework needed to utilize it in an effective manner. For instance, if a student uses their smartphone to translate a word that they do not know, they have used technology without developing language competency; if a student copies AI-created text into their essay; they have used technology but at the same time bypassed the cognitive processes that must take place when learning a language. The problem is not the technology, but rather the lack of knowledge framework that permits the learner to utilize technology in such a way that promotes or substitutes true learning.

Mishra and Koehler's (2006) Technological Pedagogical Content Knowledge (TPACK) framework was created in part to address the issues stated above; however, the issue is addressed from the teacher's side of the equation rather than students (Mishra & Koehler, 2006). TPACK, or Technological Pedagogical Content Knowledge, describes the combination of knowledge needed for effective teachers who use technology in their instruction. This includes knowing the content they are delivering (Content Knowledge (CK)), knowing how to teach that content (Pedagogical Knowledge, or PK), and finally, knowing how technology can be utilized to assist both PK and CK to accomplish their instructional goals (Technological Knowledge (TK)). The major premise of the TPACK framework argued by Mishra and Koehler (2006) is that to effectively integrate technology into instruction, teachers must be simultaneously knowledgeable about all three of these areas, and also how each of these areas

interacts when developing lessons using technology. The unique aspect of the ASIM model created in this study is that the agency of the three areas of integrated knowledge does not belong solely to the teacher. In this study, the researcher expanded on the TPACK framework to develop Student-TPACK (S-TPACK); S-TPACK refers to the simultaneous application of CK, PK, and TK by the student not when designing instruction, but through the teaching of AI (in Phase 3) in an attempt to enact that instruction through the conduct of the act of teaching.

The framework of TPACK, established by Mishra & Koehler (2006), is based upon an earlier concept first outlined by Shulman (1986) called PCK (Pedagogical Content Knowledge). PCK states that expert teachers have subject matter knowledge, general pedagogical knowledge, as well as a third type of knowledge that is a product of the interaction among the first two types — knowledge of how to make the subject matter accessible/understandable to learners with particular learner characteristics in particular contexts. TPACK, while maintaining the essence of PCK, adds the technology dimension. An expert teacher who uses technology effectively not only has extensive subject area and general pedagogical knowledge, but also knows specifically how to use various forms of technology to create particular types of pedagogical affordances for subject area content. This type of integrated knowledge allows teachers to effectively integrate all three components in their teaching, without the individual components losing their power in an educational sense if separated.

Since its inception, TPACK has developed into one of the most influential frameworks in the field of educational technology research and has garnered a vast array of academic attention across many subject areas, levels of education, and countries. This academic attention has been documented in Rosenberg & Koehler's (2015) systematic review of empirical research related to TPACK in 193 studies. When discussing ASIM, the essential relevance of TPACK will be within the implications that arise for the student as opposed to the teacher, and this is something that the original design of TPACK does not explore, yet will be central to the design of ASIM.

2.6.2 S-TPACK: From Teacher Integration to Student Agency

According to the original TPACK framework, teachers serve as the agents of the integration of technology and pedagogy; it is the teacher's role to have CK, PK and TK and to use the three types of knowledge to create lessons and deliver instruction. ASIM intends to turn the definition provided in the TPACK framework upside down; this theoretical basis is founded on the principles of constructivism and Learning by Teaching established in Sections 2.2 and 2.4 of this chapter. If true understanding is the learner creating their own understanding, and if true learning is the learner assuming the role of the teacher, then the most advanced way of using technology to teach language is not for the teacher to have combined CK, PK, and TK to

create lessons, but rather, for the student to have to combine CK, PK, and TK in order to teach the AI.

ASIM's Phase 3 operationalizes the S-TPACK inversion. When a student teaches an AI, configured as a naive learner, about present perfect tense, the student has to deploy Content Knowledge (accuracy of the grammar rule); Pedagogical Knowledge (teaching how to sequence the explanation, what examples to offer, how to respond to the AI's simulated confusions); and Technological Knowledge (how to engineer the prompt with which to configure the AI's behavior to serve as a challenging and instructive teaching target rather than a submissive one). Each of the three types of knowledge is required; none is an optional prerequisite for the proper functioning of Phase 3. The integration of all three types of knowledge through the enactment of teaching is what TPACK describes as the highest level of technology-enhanced learning competency, but in ASIM, the responsibility to achieve this integration lies with the student, not with the teacher.

The theoretical implications of this reversal extend beyond just providing students immediate pedagogical experience. In traditional technology-enhanced EFL teaching, the student is viewed as a passive consumer of the teacher's TPACK integration; the original framework presupposes this relationship but offers no interrogation (Mishra & Koehler, 2006). In such instruction, the teacher makes all decisions regarding which technologies to use, in what ways, and for what pedagogical purpose, and the student simply experiences the corresponding outcome without any technology or pedagogy judgment. Consequently, the student's relationship with technology remains passive, and unthoughtful, rather than purposeful and creative; this distinction underpins ASIM's S-TPACK framework. By comparison in the S-TPACK framework, the student must develop a thought-provoking and integrated relationship with technology, whereby technology is not simply a tool for producing an action but rather a medium for exercising one's content knowledge, pedagogical skills, and technology power as a teacher. The contrast between technology literacy as consumption and technology literacy as agency is the specific type of agency that truly provides the Palestinian EFL student with the ability to survive in a world that will be dominated by AI technology and the ability to use AI technology optimally, as opposed to simply utilizing AI technology.

While S-TPACK is at its fullest development in Phase 3, all three phases scaffold the development of S-TPACK in its progression through the CA sequence. In Phase 1 the student exercises their TK in the most basic manner through formulating prompts. They are learning that the quality of the AI's response is directly related to how well they formulate their prompts. A vague prompt produces a vague response. This establishes the foundation of TK — that an AI is not a search engine, but is an interactive partner that requires precision when giving it directions. In Phase 2 the student is exercising their TK at an intermediate level by directing the AI to accomplish specific language tasks for a specific writing or speaking goal, which combines PK with TK in a manner which was not required during Phase 1. In Phase 3,

the student will demonstrate all three knowledge domains (TK, PK, and CK) simultaneously and integrated to achieve S-TPACK competence, which is the central purpose for which ASIM was designed.

The TPACK framework (Mishra & Koehler, 2006) concludes that an authentic use of technology involves not three disparate areas of knowledge but their simultaneous, contextualized intersection; the S-TPACK extension of ASIM places this integration away from the teacher and replaces it with the student and means that the requirement of Phase 3 to utilize simultaneously, Content Knowledge, Pedagogical Knowledge, and Technological Knowledge via the act of teaching constitutes the highest level of technology-facilitated language learning competency that exists in the EFL literature. In a Palestinian context in which AI tools are becoming increasingly accessible but seldom used in a way that fosters true agency, S-TPACK provides ASIM with theoretical justification for its emphasis on the students owning the technology rather than being owned by the technology. The remaining question then is whether this cumulative progression of cognitive responsibility — to understand (Phase 1), to apply (Phase 2), and to evaluate and create (Phase 3) — corresponds with an established framework of cognitive development. Section 2.7 addresses this concern using Bloom's Revised Taxonomy and Flavell's theory of metacognition.

2.7 Higher-Order Thinking and Metacognition: The Cognitive Progression Built into ASIM

2.7.1 The Problem: A Cognitive Ceiling in Palestinian EFL Instruction

Available evidence and theoretical reasoning indicate that if systematically observing EFL lessons in Palestinian secondary schools, classify each student activity according to the cognitive operation required to perform that activity, it is likely that there would be considerable similarity in cognitive demand across the board. The classroom appears to be occupied for a substantial proportion of time doing activities that require students to operate at the lowest two levels of cognitive demand - lowest - remembering (i.e., decoding the definition of vocabulary words and grammar rules, recalling events in a text) and understanding (i.e., answering comprehension questions, paraphrasing sentences, and determining correct answers). A proportionately smaller amount of time may be allocated to using knowledge (e.g., completing grammar exercises, writing sentences using a specified pattern, or producing a controlled paragraph). According to claims made by researchers in educational psychology and applied linguistics, activities requiring students to analyze, evaluate, or create are exceedingly rare compared with activities requiring students to write, read, or speak (Anderson & Krathwohl, 2001; Nguyen & Nguyen, 2017). Therefore, it is widely accepted by experts that in order for EFL students in Palestine to develop genuine linguistic competence, they must engage in activities that are cognitively demanding and develop transferable language abilities (Al-Mahrooqi & Denman, 2020).

This situation is not solely the fault of teachers; it occurs because of the structure of the assessment system (which has the most emphasis on remembering and understanding), the nature of curriculum (which seems to focus on covering content instead of developing cognitive abilities), and the typical size of students in classrooms (which makes implementing, monitoring and providing feedback for higher-order tasks difficult) (Al-Yahyaie et al., 2022). The cumulative effect might be a systematic cognitive ceiling; children who are only instructed at lower levels of the taxonomy are unlikely to develop higher-order cognitive skills, thus making them more likely to have difficulty using language for genuine communication (which requires the continual use of analysis, evaluation, and creation) than they would if their instruction had included a higher proportion of higher-order levels of cognitive development.

2.7.2 Bloom's Revised Taxonomy as a Design Framework

According to Anderson and Krathwohl's (2001) revision of Bloom's original taxonomy, there are six types of cognitive operations that can be classified into a hierarchy: remember, understand, apply, analyzing, evaluating, and creating. This framework for organizing cognitive operations in instructional design has become very widespread. The revised taxonomy is important to the development of ASIM in two respects; as an assessment tool for identifying cognitive ceilings (as described above) and as a design guideline in establishing cognitive ceiling levels across three phases of ASIM.

The taxonomy shows the reality of the way our mind works. Higher-level operations take more work than the lower level; the lower levels are needed to support the higher; and higher levels create different, more expansive, longer-lasting, types of knowledge. A student who remembers a grammar rule can get the right answer to a highly-structured question. A student who decides if an AI-generated paragraph is of good grammatical quality and explains why is exhibiting understanding, application, analysis, and evaluation at the same time, and has done so in an application in which they must apply their knowledge in an authentic, flexible, and non-routine manner. There is a different relationship to the content between the two types of evidence regarding cognitive performance; one is quantitative and the other indicates a qualitatively different relationship.

ASIM's three-phase architecture was purposely created to help support a progression through Bloom's taxonomy (i.e., with higher levels of cognitive operation required in each of the three phases). The reason for this is intentional design rather than coincidence or emergent properties; ASIM's Phase One has the target of remember and understand through AI-assisted, self-directed inquiry, while Phase Two has the targets of apply and analyze through authentic text production with the help of AI linguistic feedback and teacher conceptual coaching, and Phase Three has the targets of evaluate and create through the role reversal that occurs when students teach, correct, and question the AI. Therefore, the three-phased progression through

Bloom's Taxonomy provides an entire cognitive progression from the bottom to the top when compared with a single lesson using traditional methods because students are limited to the cognitive ceiling imposed by traditional instruction in reaching these two upper-level taxonomies.

2.7.3 Metacognition: The Thread That Runs Through All Three Phases

Flavell's work (1979) was the first to define metacognition (knowledge of knowledge), which adds a new aspect of learning not included within Bloom's taxonomy, but still impacts our understanding of how ASIM can result in the intended learning. Within the definition of metacognition are two elements: the learners have explicit declarative knowledge about their own cognitive processes (e.g., the learners know that they learn grammar by example better than by explanation); and that the learners are performing self-regulation behaviors when learning (e.g., the learners may decide to read again because they did not understand; or may decide to ask further clarification due to an additional area of confusion that arises from what the AI search returned).

In standard Palestinian EFL classes, metacognitive growth seems to be mostly left to chance. Students are not often expected to self-regulate their comprehension, assess their own performance, or adapt their strategies based on self-assessment. The teacher assesses the student; the student receives feedback. Students typically do not receive any instruction in how to develop such metacognitive strategies as (Wenden, 1998; Anderson, 2002): assessing their comprehension while writing; understanding where comprehension falters; and the appropriate use of different learning resources. As noted in literature on EFL learning strategies, students tend to possess fairly low levels of metacognitive awareness when it comes to the level of self-regulation that is needed to demonstrate true language competence (Wenden, 1998; Anderson, 2002).

ASIM's three phases incorporate metacognitive development as an essential component, rather than as an optional addition. The end of Phase 1 culminates in a formal self-diagnosis report, where the learner presents what they grasped, what they still have questions about, and what questions they would like to ask the teacher. This acts as a structured metacognitive task requiring intentional regulation of their own comprehension. In Phase 2, learners use AI-generated feedback to determine if they will use the suggested changes, or challenge those suggestions (a form of metacognitive judgment concerning the quality of an external source compared to their own). And in Phase 3, learners use the most intensive form of metacognitive activity: the third-party monitoring of their knowledge as described in Section 2.4. This entails viewing their knowledge from an outside perspective, evaluating whether they could explain it to another person, and determining the difference between what they believe they know, and what true mastery would necessitate. Repeated use of the ASIM model will develop metacognitive practice in learners and promote the development of the self-regulatory abilities

that characterize expert language users, which current EFL instruction typically fails to accomplish.

2.7.4 The Cognitive Progression across ASIM's Phases

By linking Bloom's revised taxonomy to the three phases of the ASIM model, a clear and logical cognitive sequence is produced, which cannot be achieved through traditional methods of teaching English as a foreign language. Each phase requires distinctly different cognitive functions, and the phases build on each other's cognitive functions. In Phase 1 student who has not developed a genuine understanding cannot complete Phase 2 and will not be able to perform tasks involving Phase 2 at the level of analysis. A student who has completed the work in Phase one cannot complete the work in phases two and three. The three phases represent distinct cognitive stages of development that follow a defined order.

The combination of Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001) and Flavell's (1979) metacognitive framework indicates that the architecture of the ASIM program is not just a sequence of activities, but rather a principled approach to cognitive development. The ASIM program begins with the point at which conventional Palestinian English as a Foreign Language (EFL) instruction ends and develops toward higher order competencies that students currently do not receive prior to leaving school. Through the self-diagnosis, evaluative judgment of performance, and teaching role assessment built into each of the ASIM phases, students will develop true competence (rather than performing as competent). The theoretical foundations of constructivism, self-directed learning, learning by teaching, cognitive apprenticeship, technological integration, and cognitive development have now been established; therefore the chapter will focus on the educational context in which the previously described theoretical frameworks converge. This will be done through Section 2.8, which will focus on the role of artificial intelligence in education, and the implications for EFL learners.

2.8 Artificial Intelligence in Education

The integration of generative AI into education has created both extraordinary opportunities and serious pedagogical challenges. This section discusses the ASIM in terms of the current landscape of AI-enhanced learning, including major capabilities, roles, literacy requirements, ethical concerns, and recent empirical evidence. Integration of AI is especially important to the Palestinian education system, which has large class sizes, limited contact hours, and constrained resources stand to gain most from AI's scalable, on-demand support, yet they also face heightened equity and access challenges that must be carefully managed.

2.8.1 Generative AI: Capabilities and Limitations

Generative AI (GAI) refers to AI systems capable of producing new content—text, audio, images, code—in response to natural-language prompts (Bommasani et al., 2021). Powered by

large language models (LLMs) built on the transformer architecture (Vaswani et al., 2017) and trained on billions of documents, allowing them to produce human-like responses to natural language input through probabilistic pattern recognition rather than explicit programming. For language education, several capabilities are particularly relevant.

Generative AI brings six capabilities that are particularly consequential for language learning in contexts such as Palestine. First, large language models process natural language input without requiring learners to master specialized commands, lowering the cognitive barrier to productive engagement and enabling interaction in plain English or Arabic (Kohnke et al., 2023). Second, AI adjusts its explanations dynamically to match each learner's current proficiency level, functioning as what Vygotsky (1978) termed a more capable other operating within the Zone of Proximal Development—supporting learners to accomplish tasks they could not achieve independently. Third, AI offers unlimited opportunities for practice without the social anxiety that inhibits performance before peers or teachers, a feature that is especially critical in Palestinian classrooms where individual speaking time is severely limited by large class sizes and curriculum constraints (Nation & Newton, 2021); as Hwang and Chang (2021) documented, chatbots provide cost-effective and flexible support tailored to individual learner needs precisely because they are unconstrained by the social dynamics of the classroom. Fourth, AI's 24/7 availability removes the temporal barriers that confine learning to the scheduled class period, enabling self-directed practice outside school hours (Kohnke et al., 2023). Fifth, contemporary large language models support Arabic alongside English, enabling learners to use their first language as a comprehension scaffold, a strategy grounded in Cummins' (1979) linguistic interdependence hypothesis, which holds that competence developed in L1 transfers to and supports the acquisition of L2. Sixth, AI provides immediate corrective feedback on grammar and vocabulary, allowing rapid revision cycles; Kohnke et al. (2023) identified this capacity for instant, on-demand feedback as one of ChatGPT's central affordances for language learning, particularly in contexts where teacher feedback is structurally limited.

Generative AI also carries well-documented limitations that educators must understand. Large language models are prone to hallucination—generating fluent but factually inaccurate content—a problem systematically documented across multiple natural language generation tasks (Ji et al., 2023). Bender et al. (2021) argued that LLMs perform sophisticated pattern matching rather than genuine language comprehension, processing statistical regularities in training data without grounded understanding of meaning. LLMs also reflect the biases embedded in their training corpora; Navigli et al. (2023) documented that this includes cultural under-representation, whereby languages and cultural perspectives with limited digital presence, including those of the Arab world, are systematically marginalized in model outputs. Finally, AI lacks authentic emotional intelligence and cannot replicate the relational dimension of teaching; research confirms that while integrating AI into EFL instruction positively affects students' cognitive, emotional, and social engagement, participants expressed concerns about over-reliance on technology, suggesting that AI integration should be balanced with traditional

methods (Guo & Wang, 2024). ASIM's three-phase structure is designed precisely to harness AI's strengths while compensating for these limitations through structured human oversight.

2.8.2 AI as Tutor, Tool, and Tutee: ASIM's Triadic Framework

ASIM's central innovation is integrating all three sequentially within a single instructional model.

AI as Tutor (Phase 1 – AI-Guided Discovery): AI provides explanations about content, answers questions, and scaffolds understanding before classroom instruction. Students interact with AI to preview lesson material, creating a Self-Diagnosis Report (showing what they understood and what is still unclear). In Palestinian classrooms, where a single teacher may serve 35–45 students, AI tutoring democratizes access to personalized explanation that was previously unavailable to most learners.

AI as Tool (Phase 2 – Human-AI Co-Creation): AI assists writing and speaking practice through grammar feedback, idea generation, and conversational rehearsal. Research confirms AI is effective at surface-level feedback with meta-analytic evidence indicating moderate effects on grammatical accuracy (Zhang et al., 2023; Lyu et al., 2025) however, while there is some evidence of AI's ability to assess argumentation and critical thinking, the primary responsibility for this will remain with the teacher. The human-AI division of labor in Phase 2 explicitly reserves higher-order coaching for the teacher, ensuring AI augments rather than replaces professional judgment.

AI as Tutee (Phase 3 – Student-as-Teacher Mastery): According to learning-by-teaching research, which includes the protégé effect, students must have a true understanding of these concepts in order for them to accomplish this goal (Section 2.4), this phase requires genuine understanding that cannot be faked or delegated to AI. Students prompt AI to act as a confused student, then diagnose and correct its deliberate errors—simultaneously developing mastery, metacognition, and critical AI evaluation skills.

Phase	AI Role	Student Activity	Cognitive Level (Bloom's)
Phase 1	Tutor	Preview content; produce Self-Diagnosis Report	Remembering, Understanding
Phase 2	Tool	Write and speak with AI feedback; teacher coaches	Applying, Analyzing
Phase 3	Tutee	Teach AI; correct AI errors; self-reflect	Evaluating, Creating

The movement from guided to independent phases of learning follows the natural order of learning: instruction, supported practice, and then teaching others. The development of independent learning capabilities in students who have experienced teacher-directed instruction is particularly important in this context and will empower students to build confidence in their ability to learn and access all subjects.

2.8.3 AI Literacy and Prompt Engineering

As (GAI) becomes embedded in professional and academic life, AI literacy—defined by Long and Magerko (2020, p. 2) as "a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace"—has become a foundational competency comparable to reading and numeracy. Southworth et al. (2023) argue that AI literacy must be integrated across all disciplines rather than confined to computing courses, and ASIM enacts this vision within English language education.

Prompt engineering, crafting precise natural-language instructions to guide AI behavior—is taught explicitly across the three phases, progressing from simple Phase 1 prompts such as "Explain present perfect tense" to sophisticated Phase 3 prompts that require students to specify AI's learning gaps, confusion profile, and desired questioning style. Reynolds and McDonell (2021) demonstrated that even minor changes in prompt structure can substantially improve model output quality, a finding that informs ASIM's decision to teach prompt engineering explicitly rather than assuming students will acquire it intuitively.

The ability to speak English fluently and interact with AI will be particularly advantageous for students from Palestine. Students graduating with a strong command of professional English and the capability to interact with AI will develop a definite advantage in the growing use of AI in the workplace due to their ability to guide, evaluate, and use AI to communicate with the labor market.

2.8.4 Ethical Considerations and Pedagogical Concerns

Since the release of ChatGPT to the public in November 2022, Educational institutions around the world have grappled with how best to develop an adequate response to the emergence and proliferation of (GAI) writing tools. The ASIM initiative provides structural redesign-based solutions to address six primary areas of concern versus prohibiting the use of GAI.

Academic Integrity: The easy access to writing tools that employ AI has created a large number of opportunities for AI-assisted cheating on traditional text-based assessments (Cotton et al., 2023). Therefore, the ASIM approach is one of redesigning scopes of assessment versus detecting cheating. During Phase 3 of the ASIM process, an oral teaching demonstration

where students respond to extemporaneous questions from AI, diagnose errors and misunderstandings, and re-state their explanations with clarity—the demonstration cannot be scripted or outsourced to a third party; therefore, if a student has cheated using AI in previous phases of the ASIM process, he/she will fail Phase 3 and be unable to demonstrate genuine knowledge of subject material. This move towards performance-based assessment represents a very significant shift in pedagogy in the Palestinian context (where the exam culture is well-established) toward a more authentic and transferable competency development process.

Over-Reliance and Dependency: Warschauer et al. (2023) documented that over-reliance on AI writing tools risks compromising students' capacity to develop independent writing skills, a concern they describe as the "with or without" contradiction: students may perform adequately with AI support yet remain unable to write effectively without it. ASIM counters this through progressive scaffold removal: Phase 1 offers high AI support, Phase 2 moderate support with human oversight, and Phase 3 requires fully independent performance. AI functions as a temporary scaffold removed as competence develops, consistent with Vygotsky's (1978) scaffolding principle.

Equity and Access: In the Palestinian educational context, equity concerns carry particular weight. Persistent resource disparities between urban and rural schools, disruptions to education caused by conflict and instability, and limited institutional technology infrastructure mean that AI integration risks deepening rather than reducing educational inequality. ASIM addresses this through free-tier compatibility (ChatGPT basic tiers suffice), school-based computer-lab access during Phase 1, mobile-device compatibility, and teacher facilitation for students with limited digital literacy. Critically, because ASIM builds on pedagogical principles rather than specific proprietary platforms, it remains implementable even when technology access is intermittent.

Teacher Role: Rather than replacing teachers, ASIM repositions them for higher-value work. Freed from correcting 200 grammar essays individually, teachers provide conceptual coaching, cultural mediation, emotional support, and higher-order assessment. Palestinian teachers, who often work across multiple schools or under severe workload pressures, benefit particularly from this redistribution: AI handles mechanical feedback at scale, preserving teacher capacity for the relational and intellectual work that defines professional teaching (Guo & Wang, 2024).

Bias and Critical Evaluation: LLMs exhibit cultural bias, providing more detailed and accurate responses for Western contexts than for others (Navigli et al., 2023). Palestinian students should be explicitly taught that AI may underrepresent, misrepresent, or omit Palestinian cultural and historical contexts. Phase 3's AI Confusion Strategy, where students must identify and correct AI's deliberate errors, develops critical evaluation skills that extend naturally to detecting cultural bias, hallucinations, and contextual inaccuracies. Teachers can supplement Phase 3 with targeted activities comparing AI-generated descriptions of Palestinian society,

history, or culture with authoritative local sources, building critical AI literacy grounded in students' own context.

Data Privacy: Commercial AI platforms that students use generate data that could potentially fall under corporate privacy policies that lack transparency (Holmes & Porayska-Pomsta et al, 2023); ASIM recommends choosing a platform that will provide clear and readily-accessible student data protections, using pseudonyms where possible, and providing students with explicit instruction on how to read their data. Palestinian schools working under Palestinian Authority (PA) education standards should seek out guidance on culturally appropriate data privacy standards and look into whether or not there are locally-hosted or open-source LLM solutions that can fulfill this goal.

2.8.5 Empirical Evidence: GAI in Language Learning

The research on generative AI in language education has expanded rapidly since 2022. Key findings drawn from peer-reviewed, empirically verified studies directly inform ASIM's design decisions.

Writing Development

In a meta-analysis that evaluated the impact of chatbot-assisted language studies, a total of 61 Effect Sizes based on 18 different studies were included and an overall moderate effect size of ($g=0.527$) indicated the support provided by chatbots for language learning was positive. The effect sizes were strongest where the study duration was between 1-3 months ($g=0.681$) but diminished considerably for studies less than 1 month ($g=0.208$). Improvements were mostly found in the areas of grammatical accuracy and vocabulary Range with regard to chatbot support with less improvement with regard to argumentation quality and depth of critical thinking. These results support ASIM's Phase 2 Division of Labor in which AI addresses the surface level linguistic accuracy while teachers address the development of a student's conceptual and argumentative depth (Zhang et al., 2023).

Lyu et al. (2025), in addition to these findings, reported an overall effect of ($g = 0.543$) for chatbots as aids for second language acquisition in their review of 59 experimental studies published in the International Journal of Applied Linguistics. Learning context (foreign vs second language setting), mode of organization (individuals vs groups), and target language were all identified as statistically significant moderators of the effects, thus demonstrating the importance of cultural and linguistic context to design decisions based on the ASIM project description for Palestinian participants.

Three fundamental contradictions exist for L2 writers using generative-AI tools, according to research published by Warschauer et al. (2023)

1. The Mimicry Contradiction states that prose produced via AI will sound like a native speaker but not reflect the writer's own developing voice.
2. The Matthew Effect Contradiction means that students who are already at a higher proficiency level will be able to more readily obtain advantages from using AI than will those writers at lower proficiency levels, resulting in increasing separation between these two groups of writers.
3. The With-or-Without Contradiction suggests that those students who are exposed early to writing using AI tools are likely to become dependent upon their use of the tool(s) and, consequently will be unable to produce their writing independently once they have adapted to using such tools over an extended period of time.

The risk of this occurring is directly addressed by removing the AI scaffolding, as required by ASIM Phases 1 through 3, prior to a student's participation in summative assessment.

Speaking and Interaction

Through a systematic review of the literature, Kohnke et al. (2023) have demonstrated that non-judgmental interaction is one of the most valued characteristics of chatbot-based language-learning tools. This finding is especially important for Palestinian learners since the fear of speaking in front of others creates an added barrier to participating orally. ASIM's Phase 2 will facilitate a low-anxiety, non-evaluative speaking practice environment for learners to interact with an AI prior to being evaluated by a teacher or peer.

Usage Patterns and Academic Integrity

Public and institutional conversations about ChatGPT in relation to education are predominated by the subject of academic integrity and redesign of assessments, while student voice was largely absent according to a content analysis of 100 news articles from Australia, New Zealand, the U.S., and the U.K. Sullivan et al. (2023) concluded that unguided access to AI is pedagogically flawed and that a structured framework is needed to help students maximize their learning through use of AI and that this conclusion supports the architecture of ASIM in its three-phases.

In a huge survey across North America, Wiley (2024) reported that (850) faculty members and (2,067) students believe that (96%) of faculty think their students use AI to cheat and also noted that (53%) of students believe cheating has increased since the advent of generative AI. The survey data also indicated that students used AI primarily for writing and brainstorming; however, in many cases, students did not have a clear understanding of institutional limits or guidelines related to AI usage. These data further support the need for clear role definitions of the ASIM across all three phases of use in order to discourage misuse and provide structure to productive engagement with AI technologies.

2.8.6 Distinctive Contributions of ASIM

Situating ASIM within the AI education literature reveals six distinctive contributions:

1. ASIM combines all three roles of AI (tutor, tool, and tutee) into one comprehensive developmental progression; thus, maximizing the pedagogical affordances of AI while reducing the potential limitations through supported human-AI complementarity.
2. ASIM is also doable in a resource-constrained environment (e.g., Palestinian schools) using commercially available generative AI tools rather than specific generative AI software, which cannot be used without institutional software licenses and requires a certain level of technical infrastructure.
3. ASIM develops AI literacy (particularly prompt engineering and skill evaluating critically) organically through practical practice as a precursor to students' deployment of AI in an automated professional capacity.
4. ASIM also contains a series of structural safeguards (e.g., Phase 3 verification, the systematic removal of scaffolds, and ongoing teacher support), which address specific concerns regarding the integrity, dependence on conclusion of construction, and quality of evidence reported in empirical-based literature (Warschauer et al., 2023; Sullivan et al., 2023; Wiley, 2024).
5. ASIM also applies a combination of proven forms of pedagogical theories (e.g., constructivism, self-directed learning, learning through teaching, Bloom's Taxonomy) into an innovative technological context, ensuring that ASIM is based on theoretical grounds, rather than technology.
6. ASIM also provides a complete implementation framework (consisting of a series of specific procedures, structured prompts, and assessment criteria), allowing the adoption of AI to be implemented without requiring in-depth experience in AI.

Additionally, in the Palestinian context, ASIM is designed to address the following previously documented gaps associated with the teaching of English language: limited opportunities for speaking practice; large class sizes; inadequate teacher capacity and resources; underdeveloped student independence and critical thinking resulting from a predominantly transmission-based pedagogical approach to education. With AI at the core of a systematic, teacher-supported framework, the ASIM program presents a comprehensive, equitable process by which to provide quality English language education for all Palestinian students.

2.9 English Language Skills Development: Writing and Speaking in the Palestinian EFL Context

The preceding sections have established the pedagogical and technological foundations of ASIM. The focus now shifts to the two specific language skills the model was designed to

develop: academic writing and conversational speaking. Before examining each skill, however, it is essential to establish a distinction that shapes every dimension of ASIM's design. In the Palestinian context, English functions as a Foreign Language (EFL) rather than a Second Language (ESL) a difference that is not merely terminological but structural and pedagogical. ESL learners encounter the target language continuously outside the classroom, through daily social interaction, media, and environmental print. EFL learners in Palestine do not: English exists almost exclusively within formally scheduled lessons, and once a student leaves school, meaningful exposure to the language effectively ceases. This absence of an ambient English environment means that the classroom, and whatever supplementary resources the learner can access independently, must bear the entire burden of language development, a structural reality that directly motivates ASIM's use of AI as a surrogate language environment available beyond school hours.

This distinction is not terminological but pedagogical and structural. In ESL contexts, such as an immigrant learning English in the United States or the United Kingdom, the target language permeates the learner's daily environment: it appears on street signs, in shops, on television, and in workplace interactions; input is continuous, exposure is unavoidable, and the social motivation to communicate is immediate (Krashen, 1982). In EFL contexts, by contrast, none of these conditions apply: the foreign language exists almost exclusively within formally scheduled classroom lessons, and once learners leave school, meaningful exposure effectively ceases (Hyland, 2019). In the Palestinian context specifically, English instruction is limited to three 45-minute sessions per week, and Al-Ureimi (2021) confirmed that Palestinian students and teachers alike identify the absence of any English-speaking environment outside school as one of the primary barriers to language development. There is no ambient linguistic environment to reinforce what is taught, no incidental acquisition through daily interaction, and no communicative necessity that extends beyond assessed school tasks.

The practical consequences of this distinction are severe and directly inform ASIM's architecture. First, speaking practice in EFL classrooms is structurally insufficient: when instructional time is distributed across four language skills and a full class of students, the individual speaking time each learner receives falls dramatically short of the more than 100 hours that Nation and Newton (2009) identify as necessary for fluency development, and unlike ESL learners, Palestinian EFL students have no out-of-school interaction to compensate for this deficit. Second, writing development depends entirely on formal instruction because students have no access to authentic English text environments outside school. Third, and most significantly, the absence of a living English environment means that the classroom and any supplementary tools the student can access independently must collectively bear the entire burden of language development. It is precisely this vacuum that ASIM's AI-mediated phases are designed to fill: the AI functions not merely as a pedagogical tool but as a surrogate language environment, available at any hour, infinitely patient, and capable of providing the authentic, responsive English interaction that the Palestinian EFL context structurally denies.

As Hyland (2019) observed, the capacity to write and speak effectively in English has become an essential prerequisite for academic success and professional advancement in an increasingly globalized world. In the Palestinian EFL context, both skills face structural barriers that traditional classroom instruction cannot adequately resolve. Large-size classes prevent teachers from providing the individualized feedback that writing and speaking development requires (Hammad, 2016b; Ahmed et al., 2020). Instructional time falls far short of the practice thresholds necessary for fluency development (Nation & Newton, 2009). The complete absence of out-of-school English input, which is the defining condition of all EFL environments, means there is no ambient channel through which classroom learning can be reinforced (Krashen, 1982). High affective barriers, particularly speaking anxiety driven by fear of negative evaluation and the avoidance cycle it produces, further suppress the limited practice opportunities that do exist (Horwitz et al., 1986; Qaddumi et al., 2023). Empirical evidence reviewed below demonstrates that ASIM's three-phase architecture directly addresses each of these barriers through the complementary deployment of AI-mediated and human-guided instruction.

2.9.1 Academic Writing Development in EFL Contexts

2.9.1.1 Cognitive and Linguistic Demands of L2 Writing

Writing in a foreign language imposes simultaneous cognitive and linguistic demands that are qualitatively more challenging than those encountered in first-language composition. Cumming (2016) identified what he termed a compounding burden on EFL writers: learners must manage limited vocabulary and grammatical knowledge, sustain high cognitive load due to a lack of automaticity in linguistic encoding, and resist the rhetorical influence of their first language, all concurrently. These three pressures interact in ways that make writing disproportionately demanding compared with other language skills. In the Palestinian EFL context, this burden is intensified by the absence of any reinforcing English environment outside the classroom: unlike ESL learners who encounter English incidentally in daily life, Palestinian EFL writers rely exclusively on formal instruction for exposure, leaving no natural channel through which vocabulary, grammar, or rhetorical conventions can be consolidated between lessons.

A critical dimension of this difficulty for Arab EFL learners is rhetorical transfer, which refers to the systematic importation of Arabic compositional conventions into English academic prose. Kaplan's (1966) foundational work established that Arabic and English academic writing follow structurally distinct patterns: English paragraphs develop linearly and thesis-first, while Arabic prose is characterized by circular organization, parallel constructions, and a tendency to approach the central point from multiple angles before arriving at it. Although Kaplan's original framework has been critiqued and refined over subsequent decades, its core

observations regarding Arabic rhetorical organization have been consistently confirmed by later scholarship. Connor (2011), whose concept of intercultural rhetoric represents the most widely accepted contemporary development of Kaplan's work, documented that Arabic academic writing employs lexical repetition and parallelism for cohesion, strategies that English academic readers interpret as digression or redundancy, and privileges appeals to authority as evidence, in contrast to the empirically grounded citation norms of English academic writing. These contrasts are not errors but distinct rhetorical traditions; however, they conflict systematically with English academic expectations and require explicit, sustained instruction.

2.9.1.2 The Writing Process and Its Breakdown Points

Flower and Hayes' (1981) cognitive process model describes writing as moving recursively through planning, translating (drafting), and reviewing stages. Arab learners face characteristic difficulties at each stage. During planning, students typically brainstorm in Arabic and then translate, converting a creative generative process into a cognitive translation exercise that introduces both linguistic errors and rhetorical mismatches before a word is written in English. During drafting, the mental effort of concurrent vocabulary retrieval, grammar application, and L1-to-L2 conversion exhausts working memory resources before ideas are fully developed — what Cumming (2016) calls premature closure, in which the writer stops before the idea is complete because cognitive resources are depleted. During revision, students who lack intuitions about natural-sounding English restrict themselves to surface corrections, spelling, punctuation, rather than reorganizing content or deepening arguments.

Research on Palestinian learners specifically confirms these patterns. Al-Ureimi (2021) conducted a descriptive survey study examining the difficulties Palestinian students face in learning English language skills, finding that writing posed the most severe and persistent challenges across all proficiency levels, with grammar accuracy, vocabulary range, and rhetorical organization identified as the primary deficit areas. These findings underscore that the obstacles described in the broader L2 writing literature are not abstract but directly and acutely operative in the target context of ASIM.

2.9.1.3 The Feedback Bottleneck

The challenges identified above traditionally require sustained, individualized teacher feedback. Research on written corrective feedback (WCF) has consistently demonstrated that the most effective form of feedback, metalinguistic explanation, which provides students with explicit accounts of why an error is incorrect, produces significantly greater accuracy gains than direct correction alone (Ferris, 2011; Kang & Han, 2015). Yet metalinguistic feedback is also the most time-intensive feedback type to produce.

In Palestinian EFL classrooms, the large number of students per class creates an intractable structural problem. A teacher managing multiple classes faces an overwhelming volume of

writing papers per assignment, and providing even minimal written responses to each student would require a time investment that far exceeds what is realistically available within a teacher's working schedule. In practice, teachers respond by reducing the frequency of writing assignments to well below what the process writing literature identifies as optimal for skill development (Flower & Hayes, 1981; Hammad, 2016). The consequence is a feedback bottleneck in which the most educationally necessary intervention is also the least structurally feasible.

Automated Writing Evaluation (AWE) systems have been proposed as a partial solution, but their limitations are well documented. Research consistently shows that AWE tools perform adequately on surface-level features such as grammar and spelling, while their feedback on rhetorical organization, content coherence, and argumentative structure remains unreliable and generic (Strobl et al., 2019; Wilson & Roscoe, 2020). These are precisely the dimensions most critical for Palestinian EFL learners who struggle with rhetorical transfer from Arabic. Generative AI offers a more sophisticated alternative through conversational, adaptive feedback that can be configured by the student to target specific aspects of their writing. Guo and Wang (2023) compared ChatGPT-generated feedback with teacher feedback on EFL students' argumentative writing and found that combining AI feedback for surface-level accuracy with teacher feedback for higher-order content development produced superior writing outcomes compared to either source alone. This finding directly informed ASIM's dual-scaffolding design in Phase 2, where AI handles linguistic feedback while the teacher focuses on conceptual and argumentative depth. Table (2.3) maps each key writing challenge to the traditional approach, its structural limitation in the Palestinian context, the ASIM innovation that addresses it, and its theoretical basis.

Table 2.3.A: Writing Challenges and ASIM Solutions.

Writing Challenge	Traditional Approach	Limitation	ASIM Innovation	Theoretical Basis
L1 rhetorical transfer (circular Arabic vs. linear English)	Teacher explains once in class	Students have no reference when writing at home	Phase 1: Students query AI about English rhetorical conventions on demand	Contrastive Rhetoric (Kaplan, 1966)
Cognitive overload during drafting (L1-to-English translation)	Advise students to think in English	Not actionable without support structures	Phase 2: AI provides real-time grammar support, reducing translation burden	Cognitive Load Theory (Sweller, 1988)

Table 2.3.B: Writing Challenges and ASIM Solutions.

Inability to revise effectively	Peer review	Peers also lack proficiency; blind leading the blind	Phase 2: AI gives immediate linguistic revision suggestions; teacher adds conceptual depth	Process Writing (Flower & Hayes, 1981)
Unsustainable teacher feedback burden.	Reduce assignments to 2–3 per semester	Far below the 8–10 optimal for skill development	Division of labor: AI handles surface feedback; teacher focuses on argumentation and ideas	Ferris (2011); Kang & Han (2015)
Writing anxiety	Encouragement and reassurance	Does not address structural root cause	Phase 2: AI provides non-judgmental, private, repeatable feedback	Affective Filter (Krashen, 1982)
Absence of authentic communicative purpose	Simulated writing tasks	Students know the task is artificial; low investment	Phase 3: Students write to teach the AI — a genuinely communicative, high-stakes task	Authentic Assessment (Wiggins, 1990)
Difficulty faced by Palestinian EFL learners (vocabulary, grammar, expression)	Traditional teacher-led instruction	Insufficient individualization in large classes (Al-Ureimi, 2021)	ASIM integrates AI for personalized, adaptive language support	Al-Ureimi (2021); Özkan et al. (2024)

Note. Phases refer to the three instructional phases of ASIM described in Chapter Three.

2.9.2 Speaking English Skills Development

2.9.2.1 The Time and Anxiety Problem

Unlike writing, which permits planning, revision, and reflection, speaking demands that the learner process meaning, select grammar, retrieve vocabulary, and manage interaction all simultaneously and in real time? This fundamental difference makes speaking the most cognitively demanding of the productive skills, and also the most anxiety-inducing in formal classroom settings (Horwitz et al., 1986). In the Palestinian EFL context, this challenge is

compounded by a structural reality that has no equivalent in ESL environments: because English has no presence outside the classroom, every speaking opportunity a Palestinian learner encounters is confined to formal, teacher-evaluated instruction. There is no incidental conversation, no overheard dialogue, and no low-stakes social interaction in the target language that might build confidence and fluency gradually. Nation and Newton (2009) established that meaningful speaking fluency requires sustained, repeated practice far exceeding what any formal classroom schedule can provide, and in an EFL context like Palestine, the classroom is the only practice environment that exists. It is this structural vacuum, rather than any deficiency in teaching quality, that ASIM's AI-mediated Phase 2 is designed to address by providing Palestinian EFL learners with an accessible, low-stakes, and unlimited speaking environment beyond school hours.

This critical shortage of practice is compounded by pervasive speaking anxiety. Horwitz, Horwitz, and Cope (1986) identified three interrelated sources of foreign language classroom anxiety: communication apprehension, fear of negative evaluation, and test anxiety. The interplay of these sources is especially damaging in speaking contexts because, when each speaking opportunity is rare and public, errors become memorable failures and the teacher functions as the sole judge of correctness. The anxiety, avoidance cycle that results, where anxiety reduces willingness to speak, reduced speaking maintains low proficiency, and low proficiency intensifies anxiety which is particularly resistant to classroom-based intervention.

Al-Ureimi's (2021) survey of Palestinian learners confirmed that speaking was ranked as the most anxiety-inducing of the four language skills, and that insufficient practice time was identified by both students and teachers as the primary barrier to speaking development. These findings provide direct empirical grounding for ASIM's Phase 2 design, which positions AI-mediated conversation practice as a structural solution to the time and anxiety barriers that Palestinian teachers cannot resolve within the limits of formal instruction.

2.9.2.2 Communicative Competence and Its Components

The communicative competence framework proposed by Canale and Swain (1980) has been central to EFL pedagogy over the past 40 years. The framework identifies grammatical, sociolinguistic, discourse and strategic competence as the core components of effective communication, each of which develops through qualitatively different types of practice. What is critical for understanding the structural limitations of Palestinian EFL classrooms is that the majority of these components are largely impossible to develop through teacher-fronted instruction alone. Sociolinguistic competence, which governs the appropriateness of register, politeness conventions, and contextual language variation, cannot develop when students are exposed exclusively to the formal and decontextualized English of textbook dialogues, since sociolinguistic norms are acquired through repeated exposure to authentic communicative interaction rather than through formal instruction (Taguchi, 2015). Discourse competence, involving the capacity to produce coherent and extended speech, atrophies in classrooms

where teacher-controlled interaction rarely permits students to speak beyond a few sentences at a time (Loewen & Sato, 2018). Strategic competence, defined as the ability to repair communication breakdowns through circumlocution, clarification requests, and paraphrase, remains dormant when teachers pre-empt difficulties by rephrasing questions rather than requiring students to manage the breakdown themselves (Loewen & Sato, 2018). Formulaic competence, encompassing the idioms, collocations, and fixed expressions that make speech sound natural and fluent, suffers when vocabulary is presented in decontextualized word lists rather than embedded in authentic communicative use, since formulaic sequences are acquired through repeated exposure in meaningful context rather than through explicit memorization (Wray, 2002). Critically, three of these four components depend on sustained, authentic, interactive practice that classroom schedules and class sizes structurally prohibit. ASIM's Phase 2 AI conversation activities are designed to address this gap by providing Palestinian EFL learners with the kind of extended, low-stakes, and varied interaction that teacher-fronted instruction cannot sustain.

ASIM's Phase 2 conversational activities have been developed to allow for the simultaneous development of all four components of communicative competence: AI simulates varied interlocutors (sociolinguistic competence), requires extended topic discussion (discourse competence), creates authentic misunderstandings that demand repair (strategic competence), and introduces lexical chunks in meaningful context (formulaic competence).

2.9.2.3 Technology-Mediated Speaking Practice: Evidence Base

Research on AI-mediated speaking practice provides robust empirical support for ASIM's design. Kohnke, Moorhouse, and Zou (2023) identified AI chatbots as a pedagogically significant tool for language learning, noting their capacity to provide immediate, patient, and contextually responsive interaction — qualities that align directly with the private, non-judgmental practice environment those high-anxiety Palestinian learners require. Critically, anxious or reticent students reported feeling more relaxed practicing with an AI partner than in face-to-face classroom settings, because the social stakes of error are absent.

This evidence gains additional weight when considered alongside the EFL teacher perspective documented by Özkan, Erdemir, and Coşkun (2024), whose systematic review of 29 studies found that EFL teachers broadly recognize AI tools as helpful for increasing student engagement, promoting personalized learning, and reducing teacher workload on routine tasks. However, the same review identified substantive concerns: the potential for AI to facilitate academic dishonesty, the risk of student over-reliance that may undermine independent critical thinking, and the inadequacy of training available to both teachers and students for using AI effectively. These concerns are not peripheral to ASIM's design but central to it: ASIM's Phase 2 teacher role specifically addresses over-reliance by requiring teachers to observe AI conversations and intervene on higher-order competencies, and Phase 1

and Phase 3 both require metacognitive reflection that explicitly counteracts passive dependence on AI output.

2.9.3 Integration of Writing and Speaking in ASIM

In Palestinian schools, conventional teaching methods have historically divided writing from speaking, largely due to time restrictions rather than effective teaching practice. This has led to ASIM implementing writing and speaking in an integrated manner across all three phases of the program. This decision is based on three inter-related reasons.

First, Cumming (2016) documented those learners who practice skills in isolation risk developing what might be termed modal fossilization: incorrect patterns become fixed within one modality, and corrections applied in another modality fail to transfer. A student who repeatedly writes "He go" may correctly say "He goes" in speech, yet the written error persists because the two channels have never been activated simultaneously with the correct form. ASIM addresses this by requiring students in Phase 3 to explain their writing choices orally to the AI, activating written and oral channels at the same moment. Richards (2008), in his comprehensive account of speaking pedagogy, similarly emphasizes that productive integration of oral and written tasks produces qualitatively superior learning outcomes compared with either skill practiced in isolation, particularly for EFL learners who, unlike ESL learners, cannot rely on their daily environment to bridge the gap between the two modalities.

Second, writing and speaking share several cognitive processes, vocabulary retrieval, grammar selection, idea organization, audience awareness, and self-monitoring, and practice in either modality therefore strengthen both (Bereiter & Scardamalia, 1987). Oral brainstorming, which imposes lower production demands than writing, scaffolds subsequent written organization; written preparation, which permits editing and reflection, develops grammatical accuracy that gradually transfers to speech (Swain, 2006). By integrating the two skills, ASIM exploits this cognitive overlap rather than duplicating effort across separate lessons (Ellis, 2003).

Second, in authentic real-world English use, writing and speaking are rarely isolated. Business professionals write slide decks and present them orally; academics write papers and defend them in discussion; job applicants write cover letters and elaborate on them in interviews. Instruction that separates these skills creates artificial boundaries that learners must bridge independently when they enter professional or academic contexts (Widdowson, 1978). ASIM's Phase 3 mirrors authentic integration by requiring students to write a teaching plan for a grammar concept and then deliver it orally to the AI, encountering the same concurrent cognitive and communicative demands they will face outside school (Ellis, 2003).

Furthermore, constructivist approaches to language learning — particularly as articulated by Szabó and Szoke (2024) in the context of AI-assisted education, emphasize that when learners

interact actively with generative AI as a creative and explanatory partner, rather than as a passive information retrieval tool, they build richer and more durable knowledge structures. ASIM's integration of writing and speaking in Phase 3's teaching activities exemplifies this active constructivist engagement: students do not simply produce language but use it to construct and communicate meaning in response to an interlocutor that pushes back.

2.9.4 Summary: How ASIM Addresses Writing and Speaking Development

The review in this section demonstrates that ASIM's design directly and systematically targets the challenges that traditional Palestinian classroom instruction cannot adequately address. Table (2.4) provides a consolidated comparison of each challenge, the limitation of conventional teaching, and the ASIM advantage.

Table 2.4.A: ASIM's Comprehensive Response to Language Development Challenges.

Challenge	Traditional Limitation	ASIM Advantage
L1 rhetorical transfer	Teachers provide one-time explanation of English rhetoric	Students can use AI to ask questions about English rhetoric at home anytime
Insufficient practice time	40-min classes, six times per week	AI practice is unlimited and available outside class
High anxiety (speaking and writing)	Reassurance without structural change	AI practice is private, non-judgmental, and repeatable
Burden of Feedback is Unsustainable for Teachers	There are too many students (30-40) and not enough teachers to provide quality feedback	AI can scale technology to provide immediate feedback to all students
Limited interlocutors	Students usually only have the same teacher/classmates for practice	AI provides varied speakers, registers, and contexts
Lack of Authentic Communication	Role-plays and simulations feel artificial	With AI, students are engaging in genuinely "communicative" and "unpredictable" communication
Lack of Formulaic Competence	Students often memorize phrases out of context	With AI, students are using phrases in their authentic situations

Table 2.4.A: ASIM's Comprehensive Response to Language Development Challenges.

Underdeveloped Strategic Competence	Controlled interactions prevent repair practice	With AI, students are participating in authentic communication with a high number of breakdowns, which requires students to initiate repairs
Writing and speaking skills taught separately	Time constraints force skill separation	All phases of ASIM are intentionally taught with integrated writing and speaking skills
Concerns about over-using AI (e.g., cheating)	teachers need additional training & guidelines around using AI effectively in EFL classrooms	ASIM embeds explicit AI literacy training and maintains dual teacher-AI scaffolding (Özkan et al., 2024)

A critical aspect of ASIM's design is what it does not attempt. ASIM does not position AI as a replacement for human teaching. Phase 2's human and AI co-creation structure and Phase 3's teacher observation and assessment role preserve the teacher's irreplaceable functions, namely cultural mediation, motivational support, human connection, and the evaluation of higher-order competencies such as argumentation, pragmatic appropriateness, and critical thinking. The concerns raised by Özkan et al. (2024), specifically that AI might foster over-reliance, undermine academic integrity, or substitute for skills learners should develop independently, are structurally addressed in ASIM precisely by maintaining teacher authority over the dimensions that require human judgments. Ultimately, in the Palestinian EFL context where the absence of an English-speaking environment outside school is an immovable structural reality, AI does not merely supplement instruction; rather, it stands in for the language environment that ESL learners take for granted. ASIM harnesses this capacity deliberately and responsibly, treating AI as a surrogate language world that Palestinian EFL learners can enter on demand, while ensuring that the irreplaceable human dimensions of language education remain firmly in the hands of the teacher.

2.10 Previous Studies

The present section reviews a considerable amount of highly relevant recent empirical and theoretical work bearing directly on the present research questions, theoretical framework, and experimental design. The studies are not presented in isolation but rather are grouped thematically, and written synthetically: rather than advocating for one study or another, what is meant is that studies are grouped according to the argument they make as a contribution to an overall coherent, cumulative account of what the existing literature knows, what it

questions, and where it fails to answer. The reviewed literature bears on six interrelated themes, the theoretical frameworks underpinning AI in education, generative AI and EFL writing skills, generative AI and EFL speaking skills, self-directed learning and generative AI, critical thinking and motivation in AI-enhanced EFL learning, and AI and technology integration in the Palestinian educational context, that together form the empirical and theoretical ground on which the AI-Student Inversion Model (ASIM) is built.

2.10.1 Theoretical Frameworks for Artificial Intelligence in Education

Any serious proposal for a generative AI-infused instructional model requires an upfront decision on how our field presumes AIs and human learners interact, and how that interaction allocates cognitive authority. Two recent systematic reviews confirming the need for a theoretical re-focusing also helped us to define (with significant breadth of view and high-level analytical precision) the developmental span of this field. Li, Tan, Wang, and Lowell (2025) analyzed a total of 144 empirical peer-reviewed studies published between 2023 and 2024 using PRISMA methodology and the relevant databases of Web of Science, Scopus, and ERIC. Their analysis provided evidence to support an exponential increase in published research over this time frame as well as a substantial methodological maturation within the field, moving from exploratory and descriptive to a more systematic and theoretically driven approach.

Despite this maturation, the authors identified three ongoing critical gaps in research: almost total dominance in postsecondary settings (86.7% of studies were completed in postsecondary education); overwhelming focus on writing rather than speaking (42.4% of the studies examined writing only); and missing research in K-12 education, especially at the secondary school level, which is the focus of this study.

At the same time that we were reviewing macro-level research, Lee et al. (2025) conducted a review of 49 research- and classroom-based studies published within the period of their review (January 2023 to December 2024). They found evidence of significant variation in how generative AI was used in different methods of learning and/or teaching (i.e., as a provider of feedback, as a tutor for learning, as a stimulator of cognitive growth, as a facilitator of interaction, and as a partner in conversation). They also identified recurrent issue barriers to integrating AI into classrooms regardless of method of integration, including; limitations of technology related to quality of content and quality of querying, educational concerns of over reliance, pedagogical limitations, and academic integrity issues.

The most critical finding related to the current study was the discovery of a structural gap between the operational versatility of generative AI and the lack of a theoretically grounded instructional framework to support the use of that versatility to promote argentic, critical, and

self-directed learning rather than merely using that versatility for convenience; the authors specifically called for additional research to examine K-12 education, the use of longitudinal designs, and the development of a pedagogical structure for integrating AI into schools of learning (Lee et al., 2025).

The foundational contribution to this question came from Ouyang and Jiao (2021), who offered a theoretical framework dividing AI applications in education into three paradigms. These are: AI-directed learning "we are told what AI is learning... so the direction of learning is controlled by AI"; AI-supported human learning, in which someone other than the learner is doing the teaching: "AI helps human teachers teach, but does not replace actual human teaching/learning classwork"; and human-directed AI learning, which "places human learners in the firmament as they learn how to direct an AI to learn". Ouyang and Jiao argued that a fundamental shift is needed to the third paradigm of learning where it is the student, not the technology, who possesses cognitive authority over the learning process. This significant shift will be accomplished through the three phased design of ASIM, through which cognitive authority is progressively transferred from AI to the student until Phase 3, in which the student takes on the role of teacher by directing the behavior of the AI, evaluating AI's response to his/her requests and providing explicit linguistic justification for any corrections made to the AI's intentional errors.

In summary, these three foundational studies: Ouyang and Jiao's—tripartite approach; Li et al.'s—systematic review; and Lee et al.'s—systematic review define an empirical and conceptual framework within which ASIM is situated; that is the purpose of the ASIM: to operationalize this under-researched third paradigm at the high school level while explicitly addressing agency, criticality, and self-direction of different AI integrations that thus far have not been addressed by the field.

2.10.2 Generative Artificial Intelligence and EFL Writing Skills

The infusion of ChatGPT into EFL writing pedagogy has unleashed a fast-growing stream of empirical data which is nevertheless characterized by a fruitful tension: all report measureable improvements in writing outcomes even while consistently showing that improvements are clearly not inevitable - they are dependent on careful pedagogical structuring, learner readiness and on scaffolding of instruction that 'saves students' from 'sliding into consumption of AI'.

However, the picture becomes considerably more complicated when the conditions for effective AI-assisted writing development are scrutinized more closely. Two findings recur with notable consistency across the reviewed studies. First, the effectiveness of ChatGPT as a writing support tool is highly mediated by the quality of instructional scaffolding surrounding

its use. Jmaiel, Abukhait, Mohamed, and Shaaban (2025) reported in an experimental study with 117 university students who exhibited significant improvement across all measured dimensions of writing with female students attaining significantly greater results than male students, though they highlight the need for larger populations, longitudinal studies, and comparative studies with other learning tools before making definitive conclusions about the benefits of ChatGPT for writing. Mahapatra (2024) investigated the role of ChatGPT as a formative feedback tool with undergraduate ESL students through a mixed-methods intervention study using three different writing tests and a series of focus group discussions. Their findings show a statistically significant positive impact on students' academic writing skills, with students' perceptions overwhelmingly corroborating the quantitative picture. Their most practically relevant conclusion is that with appropriate student training, ChatGPT can function as an effective feedback tool in large-size writing classes -a finding with direct relevance for the overcrowded Palestinian classroom context that motivates our own study. Second, and more provocatively, the presumption that AI-assisted instruction will always outperform traditional approaches receives direct empirical challenge from Sapan and Uzun (2024), whose mixed-methods study of 66 tenth-grade Turkish EFL learners -conducted over five weeks using SPSS 29 for the quantitative analysis and MaxQDA24 for thematic analysis-finds that traditional instruction shows greater gains in writing and vocabulary development than instruction using ChatGPT. As a highlight to their findings, Sapan and Uzun revealed qualitative data from semi-structured interviews conducted with seven participants, in which participants viewed ChatGPT as predominantly positive, but expressed concerns about academic integrity, ethics, and 'laziness' in learning -concerns which Sapan and Uzun translated into 'pedagogically meaningful implications for teachers, students, and future researchers'.

The clearest evidence of ChatGPT's ability in writing development comes from experimental/quasi-experimental studies. Poláková and Ivenz (2024), whose quasi-experimental study working with 110 Czech undergraduate EFL students detected significant increases in conciseness, grammar, inclusion of key information, and frequency of passive constructions subsequently to ChatGPT-augmented feedback sessions, while their focus group data revealed students expressed a positive impression of ChatGPT's effects on their writing skills, and regarded it as a promising tool for teaching innovation to meet the challenges of modern learners. At scale, these single study findings are supported by the systematic review conducted by Lo, Yu, Xu, Ng, and Jong (2024), who tracked 70 empirical studies on the use of ChatGPT in education published within the 18 months after the Chatbot's release, which found that writing study dominates the research landscape (in number of studies); speaking, listening, and reading remain under-studied. In this review, they documented ChatGPT's affordances around personalized learning, increased learning opportunities, and teacher support, but also identified drawbacks including incorrect information, privacy issues, and cheating. Critically, the authors note that few studies "thoroughly investigated ChatGPT's

quantitative effects on student performance using objective sources of data" - and called for more rigorous experimental designs. Song and Song (2023), working with 50 Chinese university EFL students who were randomly divided into experimental and control groups, reported statistically significant enhancements in writing organization, coherence, grammar, and vocabulary among those who had received ChatGPT-assisted feedback, along with higher levels of motivation and a greater inclination towards autonomous study. However, qualitative findings provided useful nuance: participants concurrently expressed worries over the situational validity of AI-generated texts, and from the possibility of becoming overly dependent on such outputs - echoing future pedagogic questions around how to structure AI to support learner agency rather than supersede it.

The outcome of all of this research can be summarized with two important concepts related to the central focus of this study; ChatGPT is capable of helping EFL learner's develop their writing skills and this will only happen when exposed to formal educational practices that will provide learners with opportunities to interact with ChatGPT as active and critically involved participants as opposed to passively receiving AI-created items or responses. That is the design principle of ASIM incorporated into each of its three phases, so that learners can go from using AI to explore initially while learning the structure of co-created pieces together, and finally, in the last phase, be responsible for creating the AI piece themselves (by teaching it, assessing it, and correcting it).

2.10.3 Generative Artificial Intelligence and EFL Speaking Skills

While somewhat less extensive than that for AI-supported writing research, the evidence base for AI-assisted EFL speaking development paints a similar converging picture: AI tools, most often chatbot-based conversation partners, can be expected to reliably improve measurable aspects of oral proficiency and, even more importantly, lower the affective barriers that conventional classroom speaking practice consistently fails to overcome - anxiety, fear of judgment, and communication inhibition.

The grounding significance of AI speaking instruction's affective gains in low-resource contexts is further reinforced by Alenezi and Alenezi (2025), whose quasi-experimental study with 30 Saudi secondary school students from the Northern Borders Region (a geographically peripheral, low-resource educational context with structural parallels to Palestinian EFL settings) found that three weeks supervised interaction with GPT-powered chatbot produced statistically significant speaking gains (mean gain = 5.24 vs. 1.08 for the control group, $p < .001$), and a student attitude mean of 4.35 out of 5. Teacher observations recorded high ratings for engagement (4.63), confidence (4.52), and spontaneous speech (4.33), while usage log analysis showed a moderate-strong positive correlation between total chatbot interaction time and speaking score improvement ($r = .61$)—establishing that consistent, settled engagement

with structured AI activity rather than mere access opportunity is the operative condition for measurable speaking gains.

Beyond these specific experimental findings, Al-Obaydi, Pikhart, and Hossain (2025) also provided complementary exploratory evidence from a phenomenographic design with interviews following a six-week experiment conducted with a cross-national sample of 30 Iraqi and Czech EFL college students. The findings indicate that an interactive guide and practice partner using ChatGPT substantially helps in all language skills, with 75% of participants finding it indispensable in developing speaking, 100% for writing, 82% for listening and reading, and 75% for communicating. The authors note that the findings suggest that the language practice model of ChatGPT is most effective when used regularly in extended language practice over time—a finding that reinforces the scaffolding argument rising across the writing literature and that directly motivates ASIM's collective sustained multiweek design.

The most compelling experimental evidence comes from Fathi, Rahimi, and Derakhshan (2024), whose controlled mixed-methods study with 65 EFL learners—(33) in experimental group using the Andy English Chatbot and 32 in control group engaged in the peer interaction—showed that AI mediated interactive speaking activities produced significantly greater gains in speaking skills, speaking fluency, speaking coherence, grammatical accuracy, and willingness to communicate than peer based activities. Qualitative data from semi-structured interviews suggest that learners had positive attitudes and perceptions toward AI mediated speaking instruction, indicating that the interactive and personalized nature of chatbot engagement provides an affective gain over conventional formats of speaking practice delivery is further corroborated at elementary school level by Tai and Chen (2024), who worked with 85 Taiwanese EFL learners randomly assigned to three groups: individual interaction with a CoolE Bot, paired interaction with a CoolE Bot, and a conventional classroom control group in a three-week summer program of daily 45-min AI chatbot interactions. Results showed both Bot Groups significantly outperformed the control group on post-test speaking skills, with no significantly greater gains for individual Bot use compared to paired use. Semi-structured interviews highlighted CoolE Bot as capable of coherent human-like conversation, and interestingly even chat discussions topics discussed dependent learners' interests, and more interestingly able to foster a supportive low-anxiety environment that learners described as enjoyable, motivating, and engaging, certainly one in which learner anxiety about EFL speaking was reduced and at the same time encouraged personalized engagement and self-directed learning. The convergence of those findings across age groups, across educational level, across national contexts suggest that the affective dimension of AI assisted speaking, the presence of a safe, non-judgmental, enjoyable practice environment, indexing one of the most educationally significant and practically neglected affordances of AI integration in EFL, and one of more tightly packed significance in resource constrained

contexts in which authentic English speaking opportunities outside EFL class time often scarcely exist.

The specifics of the Palestinian context are the focus of Qassrawi, ElMashharawi, Itmeizeh, and Tamimi (2024). This mixed-method investigation combines pre- and post-speaking tests with focus group discussions to investigate the use of Google Assistant for speaking development in Palestinian higher education. Results revealed that students made substantial improvement in their overall speaking performance as well as the sub-skills of speaking fluency, interaction, pronunciation, vocabulary and spoken grammar. It also found that AI-based applications can be pleasantly engaging, motivating, and supportive of enjoyable self-directed language learning experiences; but nevertheless that challenges related to accessibility, connectivity, and technical issues persist, concluding that the conditions for AI-assisted speaking development exist among Palestinian EFL learners but that the theoretically grounded instructional scaffolding and institutional support to realize their fullest potential is lacking. ASIM's phased, multiweek designs for Phases 2 and 3 make precisely this implication, embedding AI-assisted conversational practice within a thoughtfully structured pedagogical sequence, providing the sort of scaffolding the Palestinian AI literature suggests is missing.

2.10.4 Self-Directed Learning and Generative Artificial Intelligence

The relationship between generative AI and self-directed learning is both the most theoretically valuable and the most politically risk-laden aspect of AI integration in language education. Valuable because generative AI tools - available on demand, are highly responsive to individual user's needs, and can provide high-quality individualized feedback - would seem to be custom-made to facilitate the development of learner autonomy, self-monitoring, and purposefully-directed activity, as required by SDL theory. Risk-laden because a significant body of empirical evidence indicates that this potential can be realized only under specific, well-designed conditions, and when these conditions are not present, generative AI tools are more likely to create passive, unmonitored, and unconsidered upon use of resources than to facilitate learner autonomy.

By contrast, Alfares (2025) ascertained what makes outcomes that ASIM is being designed for become possible when SDL is purposefully made: purposefully scaffolding a self-directed learning plan based on self-determination theory, with (31) EFL students at Umm Al-Qura University in a one-group pre- and post-test design, produced significant gains in motivation, self-efficacy, reduced anxiety, and increased learner confidence; these gains emerged not from learner autonomy, but were the outcome of deliberately scaffold autonomy behaviors in terms of choice of "completing a course using the Instagram app," for instance, within a securely theorized instructional frame. The Palestinian side of this angle is put forward by Hamdan, Al-

Atrash, and Dweikat (2025), whose qualitative thematic study of (30) undergraduate EFL students at two Palestinian universities uncovers a perception of ChatGPT encompassing an accessible and responsive tool enhancing confidence control and motivation but raising concerns about over-reliance and misinformation and reducing cognition and engagement - and the absence of human-mediated feedback. According to the authors' conclusion, "AI is changing the way we work in social learning environments; however, pedagogical support, ethics and critical thinking skills are required to use AI successfully." This view aligns well with the instructional design of ASIM and the basis of the three-phase instructional design used at ASIM to create pedagogical scaffolding. ASIM addresses the fundamental issue of keeping learners actively engaged in metacognition and avoiding being passive consumers of AI and therefore directs them in their use of AI.

Liang and Wu (2024), using a post-humanist lens, implemented a structured ChatGPT-based critical thinking program with 58 university EFL learners in Macau, collecting pre- and post-tests of critical thinking skills and student reflection journals. They statistically found no significant changes in measured critical thinking skills after the program — but the reflection journals told another story: students felt ChatGPT supporting better cognition and described becoming more contextually aware and spatially informed critical thinkers, while stressing the importance of maintaining a balanced trade-off between the two domains of technology and human thought. The divergence between statistically flat quantitative and upwardly sloping qualitative is theoretically important in itself: it indicates that the effect of ChatGPT on critical thinking may indeed be real, but takes place on a level of cognitive sensitivity that conventional pre- and post-test instruments are not sensitive to being picked up by them — and suggests that its actualization relies on structured pedagogical design rather than incidental AI use.

Li, Bonk, Wang, and Kou (2024) researched the theoretical reconceptualization of SDL during the AI Age using thematic analysis of interviews conducted with 19 language learning YouTube Content Creators. The authors identified new constructs of ChatGPT-integrated SDL operating at two levels: (a) local, which refers to the interactive relationships between the personal characteristics of the learner and their ongoing language learning, and (b) global, which describes the ever-evolving social-political-cultural foundations of AI that are shaping a continuous changing environment in which SDL occurs. The results from their study indicated that a reconceptualization of SDL framework exists in the generative AI era through an understanding of how ChatGPT could promote SDL for language learners. They also provided implications for the design of AI-integrated language learning technologies using theoretical insights. Li, Wang, and Bonk (2024) supported this theoretical insight with their empirical research that included a mixed-methods (survey and interview) study on 276 respondents and 11 one-on-one interviews with language learners in the U.S. Their study found that learners were primarily motivated to use ChatGPT due to its flexibility and personalization, which

provided them with access to language learning resources that aligned with their levels of knowledge, areas of interest and language-learning goals. Further, they also found that strong self-monitoring skills were a condition that allowed learners to use ChatGPT effectively and achieve their language-learning goals. Therefore, it can be concluded theoretically that, within the context of AI-mediated SDL, the learner's metacognitive architecture – rather than the technological capabilities of ChatGPT – determines SDL outcomes the most.

2.10.5 Critical Thinking, Motivation, and Engagement in AI-Enhanced EFL Learning

The research conducted on the relationship of AI to higher-order cognitive processes such as critical thinking, self-regulation, and metacognition indicates a fundamental paradox: AI can both facilitate and inhibit the very cognitive skills that language education must develop. Whether AI becomes a catalyst for deeper thinking or an alternative to deeper thinking depends mostly upon both the intentionality of pedagogical design applied to AI and how the teachers who use AI are trained in their use.

Liang and Wu (2024) and Darwin, Rusdin, Mukminatien, Suryati, Laksmi, and Marzuki (2024) conducted two complementary investigations examining the same paradox from different quantitative and qualitative perspectives. As discussed earlier, Liang and Wu concluded that ChatGPT had a small but real effect on critical thinking but the presence or absence of a structured pedagogical approach would determine whether the impact was realized or not. This was confirmed through qualitative analysis of student reflection journals. Darwin et al. found that the student participants had a more complex understanding of the concept of critical thinking, which included questioning assumptions, considering context, and evaluating evidence, but also identified several limitations associated with AI, such as that AI does not personalize information to the learner, may create echo chambers, and has difficulty with nuanced interpretation. To address this, they suggest an approach to the use of AI that utilizes its strengths while also taking its limitations into consideration; pedagogical design must integrate critical thinking and the implementation of AI so that learners do not navigate AI independently, without guidance from a structured pedagogical framework. Both studies agree that any design of tasks that uses AI as a tool for students to complete should also require that students evaluate, question, and add to the AI produced text rather than passively accepting it.

The larger motivational and achievement landscape is charted by Wei (2023), whose mixed-methods study of 60 university EFL learners compared an experimental group provided with AI-mediated instruction against a control group receiving traditional instruction. Quantitative analysis found significantly higher outcomes in delivery of grammar, vocabulary, reading and writing among the experimental group, alongside higher levels of L2 motivation and greater use of self-regulated learning strategies. Qualitative data from semi-structured interviews with

14 participants in the experimental group additionally indicated that the AI platform's capacity to enhance engagement and provide personalized learning experiences was the main mechanism believed to underlie these motivation and regulation improvement gains. The importance of feedback configuration in such outcomes can be seen in Zou, Xie and Wang (2023) and Vattøy and Gamlem (2023). The first study compared 90 university students, divided into three groups of 30 receiving either technology-enhanced peer, teacher, or self-feedback respectively over 14 weeks, and found that peer and teacher feedback technologies outperformed traditional self-feedback techniques in enhancing collaborative writing quality, while peer- and self-feedback - were superior to teacher feedback. Peer- and self-feedback, moreover, outperformed teacher feedback in improving students' critical thinking tendency, behavioral and emotional engagement. The study highlights that the most effective "configuration" of language learning support is not a single source but the myriad voices contained in multiple feedback types, the precisely will multi-source architecture that ASIM employs through dual AI-teacher scaffolding in Phase 2 and student-directed self-evaluation in Phase 3. Vattøy and Gamlem (2023), citing survey data from 2,832 lower secondary school students aged thirteen to fifteen in EFL and mathematics teaching identify the most important variable for students to experience peer feedback to be awareness-raising of learning goals, with other factors including self-monitoring, self-efficacy, and enjoyment positively associated with peer feedback practice. The former study identifies self-monitoring as key enabling condition for productive feedback engagement, justifying ASIM's embedding of structured metacognitive self-monitoring activities at the conclusion of every phase — ensuring the development of evaluative disposition required for students to convert feedback from a received commodity into processed learning resource.

2.10.6 AI and Technology Integration in the Palestinian Context

The Palestinian context presents its own challenges for AI integration, challenges mired as they are in a broader context of political instability, uneven infrastructure and institutional cultures that have historically privileged examination compliance over pedagogical exploration. Fully understanding these conditions is as important as context-setting for the current study: the design of ASIM and analysis of its results cannot be dissociated from the reality to which it was attuned.

The most comprehensive contemporary account of these realities is to be found in the British Council-LearnJam (2025) evaluation of primary and secondary English teaching and learning in Palestine, revealing (among many other insights): that 32% of children are studying in classes of over 30 children, that the Tawjihi examination includes no assessment of speaking or listening (creating a significant wash back effect which sidelines oral development) and that teachers and supervisors frequently cited the absence of explicit institutional policies for addressing learning loss due to closures, movement restrictions and armed conflict. Given this

context, the literature on AI use in Palestinian higher education provides a nuanced narrative of both AI appetites and the impediments thereof. Omar, Shaqour, and Khlaif (2024) found, using a descriptive quantitative approach and two purpose-built questionnaires administered to 130 faculty members at An-Najah National University, that attitudes to AI integration in Higher Ed were generally strong (indicative of the broad recognition of AI's ability to support learning and teaching) but that the recommendations of their study highlighted ongoing fears around risks and barriers to adoption, and called for intensive training programs within institutions to enhance teaching competence when their study had found only moderate competency in AI use.

The faculty experience of negotiating AI integration in Palestine is explored phenomenologically by Hamamra, Khlaif, and Mayaleh (2025) and Hamamra et al (2025), whose complementary phenomenological studies — again using semi-structured interviews with 21 PhD-holding instructors in three universities in Palestine — showed a simultaneous sense of excitement regarding AI's opportunity for enhancing teaching quality and deep concern for the impact of AI in the performance of existing pedagogical roles, students' cognitive engagement, and the relationship with the act of teaching. The two studies called for urgent professional development and strong administrative support for responsible AI integration — a criterion that echoes the presence of a vacuum of these supports structures in Palestinian higher education. Hamdan et al (2025) provided probably the most directly applicable evidence relevant to our Palestinian SDL focus. They showed that Palestinian EFL students view ChatGPT as having both an authentic role and a positive influence on their SLD development and perception of personal autonomy and motivation, while also feeling anxious about dependency on AI-generated content, misinformation, cognitive disengagement, and losing feedback from their human instructor — these are all considered paradoxes, and that is precisely what ASIM is designed to address by providing appropriately designed human-enabled instructional sequences for and about the use of AI.

Dimensions of the student experience of technology integration in the Palestinian context reveal a continuing, consequential gap between availability and meaningful use. Alhur, Khlaif, Hamamra, and Hussein (2025), whose qualitative inquiry - semi-structured interviews with 22 subjects and focus groups of 24 medical educators from Palestinian universities - examined the cognitive dangers of unstructured AI dependence in this context, identified 5 significant drivers of AI dependence, including workload at the institution (reported by over 80% of students), and low academic self-belief (60%). The six consequences of AI over-reliance reported by the same study - skills atrophy concerning mere critical thinking and decision making (89% of participants), 'erosion of pedagogical profiles' (76%), and dampening of creativity (43%) - are a stark empirical warning of the potential dangers of adopting generative AI in educational contexts without meaningful attention to structuring how AI are used in a given local institution. The student-level experience of generative AI is captured by Hamamra

et al (2024), whose qualitative case study using student-sourced essay text assisted with ChatGPT and follow up interviews revealed that Palestinian university students' acceptance of ChatGPT as facilitating entry to academic content in principle starkly contrasted with the rote memorization of prior pedagogical approaches, but that ChatGPT use also produced declining participation in class, upticks in absence, and loss of motivation for exams - signs that ChatGPT usage, without pedagogical structure, may inhibit rather than enhance appropriation of engagement in formal learning.

Also examining teacher education, Qaddumi, Smith, Masd, Bakeer and Abu-Ulbeh (2023) highlighted ICT-beliefs of 780 Palestinian English teachers (primary & secondary) in-service across 260 schools in their study; given the ubiquity of AI tools as they compose their paper, one might almost suggest the teachers might be inclined to "AI-beliefs", but in fact their large study finds teachers confidently recognizing clear ICT potential to support English language teaching, while simultaneously providing the finding that they need a stronger emphasis on the training dimension of ICT integration, as the most urgent condition for this latter potential to be realized — with overtones of the training/scaffolding element built into the ASIM preparatory phase here. In tracing back the origins of these risks to structural causes, Khlaif and Salha (2022) employed a sequential mixed-methods design, where qualitative data comprised of 39 subjects along with 300 graduate students from two of northern Palestine's universities, explicates how factors such as students' attitude and beliefs, quality of service, cost, and instructor's effect significantly shaped mobile technology integration, and stresses how mere access does not support meaningful educational use, therefore inseparably linking structural and relational conditions that must be grappled with in any integrative technology-mediated instructional model. These structural conditions are also captured well in Qaddumi, Bartram, and Qashmar (2021), who surveyed a total of 207 teachers and 276 students from 53 Palestinian schools participating in an ICT project, finding students rated ICT as producing only moderate influence on their learning, persisting in having challenges with lesson duration limits and access to the most current devices and information research skills, a situation quite at odds with teacher's stronger impressions of ICT's promoting greater influence on their teaching. The present project's studied positioning of the student as cognitive agent rather than passive beneficiary emerges in graduate response to this empirical warning foregrounding the fact that students themselves often find less substantial impact on their learning in relation to educational technology than teachers perceive them in relation to their teaching.

Together, these studies in the Palestinian context, highlight the specific court that ASIM plays in: one in which the appetite for learning through AI is present but coexists with structural, institutional, and pedagogical conditions that, without careful thought by the educator, conspire to produce dependent yet disengaged students.

2.10.7 Summary and Research Gap

In summary, the literature indicates that generative AI tools in general, and ChatGPT in particular, have a significant potential to improve English as a Foreign Language (EFL) learning outcomes in writing and speaking skills; support learner independence/self-regulated learning; enhance learner motivation/engagement; and facilitate higher-order thinking. Systematic reviews of the literature by Li et al. (2025) and Lee et al. (2025) indicate that since the introduction of ChatGPT, the field has become more mature and has increased the focus on conducting sound empirical research that reflects theoretical understanding. While the literature has documented the advances that have been made since the introduction of ChatGPT, it has also identified a number of serious limitations and pedagogical risks to using AI tools when there is not adequate structured guidance; a clear ethical framework; and/or intentional instructional design. Studies across various regions of the world (China, Macau, Indonesia, Czech Republic, Norway, Saudi Arabia, South Korea, Turkey, United States, and Palestine) have consistently shown that unguided use of AI usually promotes learner passivity, cognitive offloading, and surface engagement in learning rather than deeper, self-regulated, metacognitive learning processes associated with true language learning.

An initial methodological overview of the studies reveals a relative diversity of designs circa 2025, reflecting the increasing empirical maturity of the field. Most studies drew on quantitative approaches gauging gains in language skills using pre- and post-test quasi-experimental or experimental designs - these include Song and Song (2023), Jmaiel et al. (2025), Fathi et al. (2024), Tai and Chen (2024), Alenezi and Alenezi (2025), Wei (2023), and Alfare (2025). A substantial number of studies used some kind of mixed-design approach, with quantitative measures being supplemented by qualitative data from interviews, focus group discussion, or reflective journals to present both statistical evidence as well as which provide interpretive depth - among these are Mahapatra (2024), Sapan and Uzun (2024) Qassrawi et al (2024), Al-Obaydi et al (2025), Zou et al. (2023). Finally, though a smaller but theoretically substantive case of studies employed a purely qualitative approach - including phenomenological, case study, and thematic designs - to investigate students lived experiences and understandings of AI-based pedagogies and contexts through “thick description”: Hamdan et al (2025), Hamamra et al (2024), Hamamra et al (2025), Omar et al (2024), Darwin et al (2024), Alhur et al (2025). Finally, two foundational studies took the form of systematic reviews, supplying macro-level synthesis of the empirical literature: Li et al (2025), Lee et al (2025). This diversity of methods strengthen the evidence base on which findings of this study can build on while simultaneously exposing a likely enduring imbalance - experimental and quasi-experimental designs capable of being able to draw out stronger inferences around AI integrated instruction and language learning were underrepresented containing rather “lightweight” lines of evidence, with the present study contributing a mixed-design quasi-experimental finding from the Palestinian EFL context.

Three specific gaps emerge from this review and whose presence the present study attempts to address directly. First, research conducted at the secondary school (as opposed to college/university) level - specifically with high school learners in public school contexts - remains profoundly scarce, with nearly all existing studies conducted in higher ed (Li et al., 2025; Lee et al., 2025). The only secondary study directly comparable to the present study — Sapan and Uzun's (2024) Turkish study — found that unstructured ChatGPT integration performed even worse than traditional instruction, indicating the need for theoretically grounded, pedagogically useful AI frameworks at this level. Second, the Palestinian context, given its unique structural, political, and infrastructural aspects, has received no attention in experimental studies examining AI-driven development of writing and speaking skills at the secondary school level - a gap rendered even more germane by the findings of British Council and LearnJam (2025) of continuing and systemic underperformance in English language abilities across Palestinian public schools, and the body of qualitative Palestinian-context studies reviewed in section 2.9.6 which collectively present an educational environment in which AI penetration is clearly broad yet pedagogically unstructured. Finally, and most importantly, the literature reveals a near zero amount of studies which address what occurs when the role of the learner is flipped entirely on its head - when the learner, rather than being a passive recipient of AI-mediated instruction, is the human agent who is designing, directing, evaluating, and correcting the behavior of the AI. Ouyang and Jiao's (2021) foundational paper identified this third paradigm — human-directed AI learning — as the most hopeful yet the least empirically examined configuration of AI in education. It is precisely this gap - conceptual, empirical, and contextual - that the present finds itself attempting to fill through the design and examination of ASIM: a theoretically-grounded, researcher proposed instructional model that aims to instantiate the third paradigm at the secondary school level in the Palestinian EFL context, with writing and speaking skills as its targeted outcomes and student cognitive agency as its design principle.

Chapter Three

Methodology

This chapter introduces the methodological framework of this study, covering the research design, the AI-Student Inversion Model developed by the researcher, the study's target population and sampling method, data collection procedures, measurement instruments, study variables, data analysis procedures, and measures of validity, reliability, and ethical considerations. The first part of this chapter describes the explanatory sequential mixed-methods research design adopted to examine ASIM's effectiveness in developing the academic writing and English-speaking skills of eleventh-grade students attending public schools in Palestine. The chapter then reviews ASIM's conceptual basis, theoretical framework, three-phase structure, and implementation procedures across the participating schools during the 2025–2026 academic year. Throughout, methodological choices are justified with reference to the theoretical framework presented in Chapter Two and to the practical realities of conducting educational research within a politically constrained, infrastructure-limited context that calls for innovative, context-sensitive instructional models.

3.1 Research Methodology

This research employs an explanatory sequential mixed-methods design that integrates two complementary approaches: a quasi-experimental quantitative approach and an exploratory qualitative approach. The goal of this design is to provide a comprehensive understanding of ASIM's effectiveness in developing the academic writing and English-speaking skills of eleventh-grade students, and of whether these effects differ between male and female students. The rationale for this design rests on several methodological considerations, including the use of precise quantitative measurements to assess the impact of ASIM, and qualitative insights to explore how learning occurs and the contextual influences shaping that process. This design also offers a comprehensive view of the skills under investigation by capturing both the general improvement in language skills and the underlying processes and influences associated with that improvement. In addition, the design allows for methodological triangulation, which enhances the credibility and trustworthiness of the findings by drawing on multiple sources of data. Furthermore, the qualitative data are used to add meaning and depth

to the quantitative findings, supporting a more complete understanding of the phenomena under investigation. Finally, given the innovative nature of using generative artificial intelligence within a theoretical instructional model, this study requires a combination of quantitative methods to quantify effects and qualitative methods to explore students' personal experiences and perspectives.

3.2 Research Design

The research employed an explanatory sequential mixed-methods design comprising two phases.

Phase 1 (Quantitative Component)

A quasi-experimental pretest-posttest design was used with two intact groups: one composed of male students and the other of female students, drawn from two separate public schools in the Ramallah and Al-Bireh Governorate.

A control group was not included in this design for two reasons related to the study context. First, students could not be randomly assigned to groups, as the two participating schools are single-sex institutions; each school's student body was therefore determined by existing administrative structure rather than by random assignment. Second, intermittent road closures between neighboring villages in the Ramallah Governorate during the study period restricted feasible access to schools outside the researcher's immediate geographic reach, making it impractical to identify and recruit a comparable external site for a control group. As Cohen et al. (2018) observe, true experimental designs involving random selection and random assignment to experimental and control groups are often not feasible in research conducted within existing educational structures. Given these constraints, a two-group quasi-experimental pretest-posttest design, using intact rather than randomly assigned groups, was identified as the most appropriate methodological option for this study.

The design is represented as follows:

Group 1 (Female): $O_1 \rightarrow X \rightarrow O_2$

Group 2 (Male): $O_1 \rightarrow X \rightarrow O_2$

Where:

O_1 = Pre-test: assessment of academic writing and English speaking skills prior to the intervention

X = Experimental Intervention: implementation of ASIM, supported by two volunteer cooperating teachers who teach the same grade level at the selected schools

O₂ = Post-test: assessment of academic writing and English speaking skills following the intervention

Phase 2 (Qualitative Component)

Qualitative data were collected at a later time using in-depth teacher and student interviews, classroom observations, and student reflections after completion of the quantitative portion of the study. These qualitative data will provide further explanation, context, and enrichment of the quantitative data.

3.2.1 Justifications for Selecting this Design

Developmental Focus: The goal of the study is to assess the researcher-developed ASIM by assessing the performance changes for each of the groups prior to and after the intervention, and also to compare the results of the two groups.

Practicality and Realism: This type of design allows application of the model with intact student cohorts in "natural" classrooms, thereby eliminating exclusion of any group from the intervention.

Assessment of Growth: It allows for monitoring the actual progress of academic writing, speaking, and independent learning skills for each of the groups, as well as cross-group comparisons of growth.

Ethical Considerations: The design is free of any potential ethical issues regarding withholding a potentially effective instructional model from any group of students.

Integration with Qualitative Data: This design provides baseline quantitative data that will be able to be interpreted and augmented with future qualitative analysis consistent with the explanatory sequential design.

3.3 The AI-Student Inversion Model (ASIM)

3.3.1 Model Overview and Conceptual Origins

AI Student Inversion Model (ASIM) is the researcher-designed instructional model is grounded in educational theory, centering on generative AI tools, ChatGPT, specifically, around a self-directed learning cycle that encompasses three sequential phases to be used in an EFL context. "Inversion" is a controlling idea of the model that becomes evident in a systematic way in which the cognitive load or cognitive control of all educational actors, the student, the AI, and the teacher, is transferred in the model. In traditional AI-assisted instruction, the teacher chooses how to use AI tools, and the student is a passive consumer of AI-generated content. ASIM systematically inverts the roles, moving to Phase 3, in which the student becomes teacher and the AI a learner.

ASIM was developed based on the researcher's observation of a significant discrepancy between the availability of software (AI tools) to EFL students and the current instructional methods, which classify EFL students as just consumers of software (AI output) rather than active participants who create, assess and critically interact with the software (AI behaviors). This observation is consistent with the theoretical literature reviewed in Chapter Two and has led to the creation of a new model of technology literacy in education. Technology literacy will no longer be defined as simply the ability to use software, but rather as the ability to control, question and teach software.

3.3.2 Theoretical Foundations

ASIM has six various related theoretical frameworks that are fully detailed in Chapter Two of this dissertation:

- Constructivism (Piaget, 1954; Vygotsky, 1978): the epistemological framework for the model.
- Self-Directed Learning (Knowles, 1975): the philosophical basis for learner-agency as seen in Phase 1's design.
- Mastery Learning (Bloom, 1968): the basis of the phase progression as a series of mastery checkpoints.
- Learning by Teaching (Duran, 2017; Whitman, 1988): the basis for student-as-teacher design in Phase 3.
- Cognitive Apprenticeship (Collins, Brown and Newman, 1989): provides four specific methods (modeling, coaching, instruction, and fading) for implementation during each of the three phases.
- S-TPACK (an adaptation of TPACK, proposed by the current researcher), in which a student integrates content, pedagogical, and technological knowledge and skills.

A full description of ASIM, its theoretical framework, phases structure, examples of activities, and definitions of key terms may be found in the Appendix (see Appendix 1).

3.3.3 Model Structure: The Three Phases

ASIM's instructional cycle consists of three interrelated phases, each characterized by distinct cognitive demands, participant roles, and session organization. As a model, ASIM defines the theoretical structure and underlying principles governing the interaction between students, teachers, and AI; as a cycle, it describes the sequential and recursive progression through which students move across the three phases within and across sessions. The table below

outlines the primary functions of each phase, and Figure 3.1 illustrates how the three phases relate to the overall ASIM structure.

Table 3.1: ASIM Instructional Cycle: Phase Structure and Cognitive Alignment

Phase	Name	Sessions	Mode	Work Type	Bloom's Level
Phase 1	AI-Guided Discovery	1 × 40 min per cycle	Asynchronous	Pre-class, Individual	Remember / Understand
Phase 2	Human-AI Co-Creation	2 × 40 min per cycle	Synchronous	In-class, Individual and Group	Apply / Analyze
Phase 3	Student-as-Teacher Mastery	1 × 40 min per cycle	Synchronous	In-class, Individual	Evaluate / Create

AI-Student Inversion Model

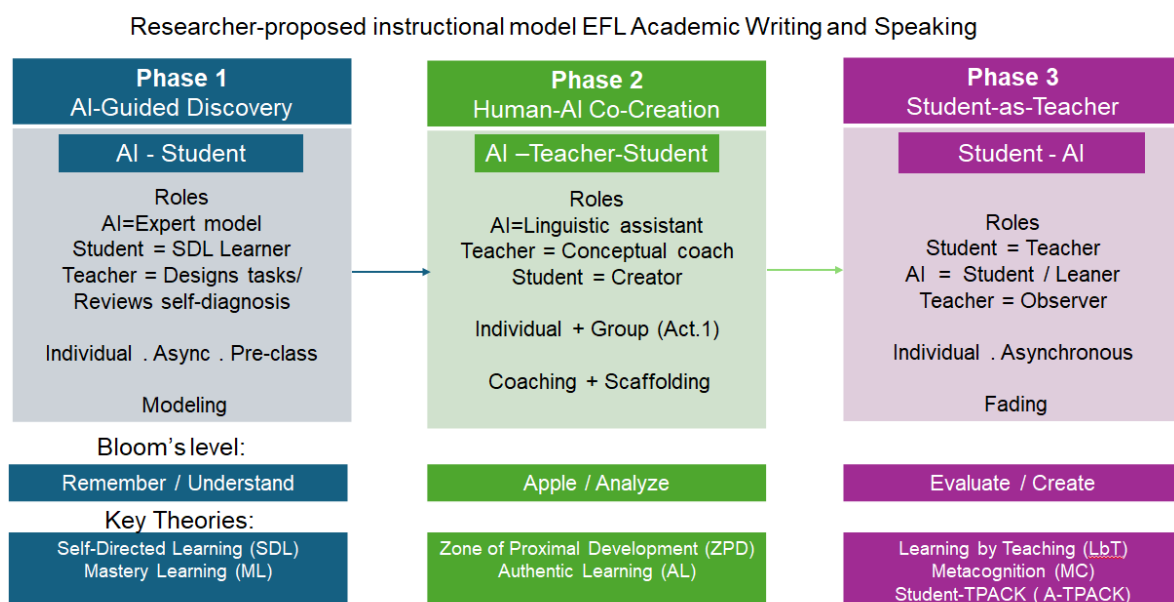


Figure 3.1. The AI-Student Inversion Model (ASIM): Three-phase instructional cycle showing role distribution, Bloom's cognitive levels, and theoretical foundations.

Note. Figure created by the researcher using Microsoft PowerPoint.

In Phase 1 (AI-Guided Discovery), students work independently with the AI in an asynchronous online environment prior to attending class. The AI functions as an expert model, providing individualized support in reading comprehension, grammar, and vocabulary directly relevant to the writing and speaking tasks students will undertake in Phase 2. Phase 1 concludes with students submitting a self-assessment diagnostic report to their teacher before class, a metacognitive task that operationalizes the self-directed learning principle of learner-initiated needs identification and that informs the teacher's planning for Phase 2.

In Phase 2 (Human-AI Co-Creation), students undertake authentic writing and speaking tasks supported by two distinct forms of scaffolding. The AI provides immediate, language-focused feedback on grammatical accuracy, vocabulary use, and fluency, functioning as an on-demand linguistic resource available throughout the task. The teacher, by contrast, provides conceptual coaching focused on content development, organization, argumentation, and rhetorical structure, drawing on judgment and contextual understanding that the AI does not provide. This division of labor is grounded in Vygotsky's (1978) Zone of Proximal Development, which supports distributing scaffolding across two complementary sources of expertise, human and technological, rather than positioning either as a substitute for the other.

In Phase 3 (Student-as-Teacher Mastery), responsibility for instructional delivery shifts to the student, who teaches a linguistic or grammatical concept to the AI, positioned as a novice learner that asks questions, expresses uncertainty, and requests clarification. This phase incorporates the AI Confusion Strategy, implemented in the following steps: (1) the student instructs the AI to produce a response containing a deliberate linguistic error related to the concept being taught; (2) the student identifies the error within the AI-generated response; (3) the student explains why the response is incorrect, using appropriate metalinguistic terminology; and (4) the student corrects the error and justifies the correction to the AI in the role of teacher. This sequence engages the evaluate level of Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001) and is intended to develop metalinguistic awareness in ways that extend beyond what conventional teacher-led error-correction activities typically require of students.

Prior to implementation, ASIM was reviewed by a panel of five specialists in English language education, curriculum and instruction, and educational technology, each holding a doctoral degree and a minimum of ten years of relevant professional experience. Panel members evaluated the model using a structured review form addressing its theoretical coherence, alignment between phases and stated learning outcomes, pedagogical soundness, and practical feasibility within the Palestinian public school context. Reviewer feedback led to revisions in [specify: e.g., the wording of the Phase 1 diagnostic report, the session timing, the AI Confusion Strategy steps], after which the model was judged suitable for implementation (see Appendix 2).

3.3.4 ASIM in the Current Study

In the current study, ASIM was applied to Units 7, 8, and 9 of the eleventh-grade EFL curriculum. These three units were selected primarily on the basis of their sequential position within the second academic term, corresponding to the period available for the intervention. This choice was methodologically defensible because all units in the eleventh-grade curriculum follow a consistent structural template, each containing an equal number of writing and speaking activities; consequently, the selection of any three consecutive units would not have altered the nature or intensity of the intervention. Within each of these three units, ASIM was implemented specifically with the two writing activities and two speaking activities designated in the curriculum, rather than across the unit as a whole, because these are the lessons most directly aligned with the dependent variables of academic writing and English-speaking skills defined in Chapter One. Across the three units, this yielded a total of twelve ASIM cycles over the intervention period. Each cycle consisted of four 40-minute sessions, one session for Phase 1, two sessions for Phase 2, and one session for Phase 3, totaling 160 minutes per cycle.

The intervention was conducted concurrently in two public schools located in the same town within the Ramallah and Al-Bireh Governorate. The first school was a public girls' school with 28 students, and the second was a public boys' school with 22 students. Each school had one volunteer cooperating teacher who taught Grade 11 English.

The two cooperating teachers completed a two-week training program consisting of structured workshops covering ASIM's theoretical basis, its three phases, and their role as facilitators throughout the process. Ongoing support continued throughout the implementation period, with the researcher available to address teachers' questions through in-person meetings, Zoom sessions, and WhatsApp communication as needed. Implementation fidelity was further monitored through weekly on-site visits by the researcher to both schools throughout the intervention period.

Cooperating teachers collected students' Phase 1 self-assessment reports prior to each session; these reports were subsequently reviewed by the researcher to inform the focus of Phase 2 instruction and were retained as part of the study's qualitative data set. Cooperating teachers also facilitated Phase 2 class discussion and observation and reviewed Phase 3 teaching transcripts, in cooperation with the researcher, throughout the intervention.

The implementation of ASIM took place under the exceptional conditions affecting Palestinian public education during the 2025/2026 academic year, including a school week reduced to three days because of economic pressures, and repeated restrictions on movement between communities. These conditions required an adaptive, contingency-based approach to delivery, applied consistently across both participating schools.

For approximately fifty days of remote learning mandated by the authorities, sessions were delivered via Zoom rather than Microsoft Teams, the platform recommended by the Palestinian Ministry of Education and Higher Education for remote instruction. This decision was based on several considerations. First, Zoom was more accessible to students with limited prior experience with video-conferencing tools, most of whom were beginner or intermediate users of such technology. Second, Zoom's core features, including screen sharing, session recording, file sharing, and text chat, aligned closely with the interactive, activity-based structure of ASIM sessions, allowing the researcher and cooperating teachers to observe student-AI interactions in real time and to review recordings for observational data collection. This choice is consistent with Kohnke and Moorhouse (2023), who found that the benefits of synchronous Zoom instruction outweighed its drawbacks for remote learners, with screen sharing identified as a particularly effective pedagogical feature. Similarly, in a Palestinian university EFL context, Badah et al. (2024) found that Zoom's recording, screen-sharing, file-sharing, and chat functions enhanced the online language-learning experience for both students and instructors, with no meaningful difference between male and female participants in their perceived effectiveness of the platform.

Throughout the remote-learning period, the researcher maintained regular contact with cooperating teachers and participants to ensure continuity of the ASIM cycle. Additional sessions were conducted face-to-face in a local community hall, with verbal parental consent, to prevent disengagement among students with limited access to remote instruction. These adjustments were necessitated by circumstances beyond the researcher's control; however, ASIM's core structure was not altered. Across all twelve cycles, the three-phase sequence, role distribution, and Bloom's cognitive progression were maintained consistently throughout the study.

3.4 Study Population and Sample:

The study population comprised all eleventh-grade students (N = 3,901) enrolled in public schools within the Ramallah and Al-Bireh Governorate during the 2025/2026 academic year.

3.5 Study Sample

Sample Type	Purposive sampling
Sample Size	50 students, distributed across two groups
	Group 1: 28 female eleventh-grade students from a girls' public school Group 2: 22 male eleventh-grade students from a boys' public school
Implementation Support	The intervention was supported by two volunteer cooperating teachers, one in each school, both of whom taught English at the eleventh-grade level. Their participation helped ensure consistent implementation of ASIM's three phases across both groups.

3.5.1 Sample Selection Criteria

Selection criteria were developed prior to recruitment so that all of the students who participated in the study met the minimal conditions required for meaningful and sustained use of the ASIM intervention.

- Regular attendance with no frequent absences
- Access to required technologies and infrastructure (e.g., tablet, smart phone, and internet)
- Intermediate to upper-intermediate English Language.
- Willingness to participate in research intervention voluntarily.
- No diagnosed disabilities or difficulties with learning that may affect student's ability to acquire language.

3.5.2 Sample Size Justification

The sample of 50 students, across both groups, was selected on the basis of the following scientific and practical considerations, each of which support the adequacy of this sample size for the mixed-method design adopted by the present study.

- A sample size of fifty participants is broadly consistent with general guidelines for quantitative educational research; Gay, Mills, and Airasian (2012) recommend a minimum of approximately 30 participants per group for experimental and correlational designs, a threshold the present study's two groups ($n = 28$ and $n = 22$) approximate. While no formal a priori power analysis was conducted, this sample size reflects conventions common in applied educational research employing quasi-experimental designs (Shadish, Cook, & Campbell, 2002) and was further shaped by the limited number of eligible intact classrooms available within the study's geographic and political context, as discussed in Section 1.8.
- The sample size will provide sufficient statistical power to complete reliable statistical analyses using paired t-tests to measure pre-to-post within groups as well as independent samples t-tests to compare male/female improvement.
- Because of the sample size, individualized and exact monitoring of each student's progress within each group is possible.
- The sample size reflected the practical limitations of the researcher, as well as the voluntary nature of participation by the cooperating teachers.
- The sample size is large enough to generate rich qualitative data for the second phase of the mixed-methods design.

- Conditions related to the political and socioeconomic environment of Palestine (reduced school days, closure of schools repeatedly) have had an impact on developing a larger sample size enough to be distributed geographically.
- The intervention was implemented in three instructional units of the eleventh grade curriculum, with a total of two writing activities and two speaking activities for each unit, resulting in twelve cycles of ASIM implementation during the intervention period.

3.5.3 Sample Distribution

The two public schools participating in this research are located in Ramallah and Al-Bireh area. They include one all-girls school and one all-boys school. Each school provides one complete 11th-grade class for this study. The outcome of this sampling strategy is that both groups of students have similar demographic backgrounds, allowing for a valid comparison of the effectiveness of ASIM across different demographic groups while maintaining ecological validity by using naturally occurring, intact classes.

3.5.4 Statistical Analysis Procedure

The two-group design included the following statistical analyses:

- Before conducting inferential analyses, descriptive statistics were computed in order to provide a general overview of participants' performance on the two skills, writing and speaking. Means (M) and standard deviations (SD) were computed for the pre- and post-test for each skill across two groups (female and male), in order to provide a general account of those variables used as outcome measures in the further inferential analyses. All analyses were undertaken using SPSS.
- The paired samples t-test utilized to evaluate improvements within each group (female and male) by comparing pre-test and post-test scores.
- An independent samples t-test utilized to evaluate the magnitude of improvement between groups and to determine if there were statistically significant differences in response to ASIM based on gender.

Statistical analyses performed using SPSS with a significance level of $p < .05$.

3.6 Data Collection Procedures and Implementation Context

Data collection and implementation of the AI-Student Inversion Model (ASIM) were conducted within a quasi-experimental design embedded in an explanatory sequential mixed-methods approach. This approach integrates both quantitative and qualitative strands to

provide a comprehensive account of the intervention's outcomes. The quantitative phase was implemented first, followed by the qualitative phase, in accordance with the explanatory sequential design adopted for this study.

Both phases of data collection were carried out under the exceptionally challenging circumstances affecting Palestinian public education during the 2025/2026 academic year, as described in Section 1.2. These conditions required flexibility and adaptation in how the intervention was delivered, while maintaining the theoretical integrity and pedagogical cohesion of ASIM throughout.

3.6.1 Pre-Intervention Phase

The researcher secured an official letter of facilitation from her university prior to starting the intervention (see Appendix 3) and submitted this letter to the principals of both schools, who verbally agreed to allow the study to be conducted at their respective institutions, consistent with standard administrative procedures for public schools in Palestine. Following this agreement, parental consent was obtained through two complementary channels. For most participants, consent was obtained during an in-person meeting held at each school, in the presence of the school principal, during which the researcher explained the purpose, procedures, and voluntary nature of the study to attending parents. For parents unable to attend this meeting, the researcher subsequently contacted them by telephone to provide the same information and obtain their verbal consent. In both cases, consent was verbal rather than documented through a signed written form, consistent with common practice for classroom-based research in the Palestinian public school context; this reliance on verbal consent, though administratively supported by the schools' principals, is acknowledged as a limitation of the study's ethical procedures and is addressed further in Section 1.8. Each cooperating teacher, one female and one male, received thorough training on ASIM's theoretical basis, three-phase cycle, and their role in supporting implementation. Prior to the intervention, pre-tests of students' academic writing and English-speaking abilities were administered to both groups to establish baseline performance levels.

3.6.2 Intervention Phase

The ASIM intervention was administered over twelve weeks across three instructional units of the eleventh-grade EFL curriculum, with one ASIM cycle implemented per week. Within each unit, ASIM was applied specifically to the two writing activities and two speaking activities it contained, that, to the productive language skills (academic writing and speaking) constitutes the dependent variables of the study, rather than to the unit as a whole.

In-Person Delivery:

When regular schools were open (limited to three days a week due to limited resources), the ASIM program was delivered in-person at the participating partners' schools, with the trained teachers conducting classroom sessions at their sites and being partnered with the researcher in the delivery of the ASIM program to ensure fidelity to all three phases across both groups of students.

Remote Delivery via Zoom:

Educational authorities declared a period of about 50 days in the study to be Remote Learning. The researcher used Zoom to facilitate individual and continuous mediate with each of participants throughout this time. The researcher hosted regularly scheduled live online meetings to implement the intervention, check participants' progress, provide feedback on writing and speaking tasks, and collect data through observations. Participating teachers were actively involved in all Zoom meetings; this ensured both pedagogical consistency and dual support for students from both the researcher and the teacher.

Supplementary Face-to-Face Sessions:

To help counteract the limitations of complete remote instruction, particularly for those students who had trouble accessing the internet or devices, the researcher and cooperating teachers planned several additional hands-on sessions in the local community hall. Parental consent for these extra sessions was obtained through direct personal communication between the researcher and parents. This arrangement was made in order to ensure that no participant was negatively affected by the move to complete remote learning and continue with the ASIM intervention throughout the time of study.

3.6.3 Post-Intervention Phase

At the conclusion of the intervention, two groups of students were given two post-tests that were constructed and scored the same way as the pre-tests so that the two groups could be compared on the basis of their growth after the intervention in both their academic writing skills and their English speaking skills. The quantitative data that was collected from the pre- and post-tests were then analyzed using paired samples t-tests to measure improvement within each group and an independent samples t-test to compare the results between males and females.

3.6.4 Qualitative Data Collection Phase

Consistent with the explanatory sequential design, the qualitative phase was not intended to generate independent or exploratory findings, but rather to serve as evidentiary support for the quantitative results, namely, to trace and illustrate how individual students' academic writing

and speaking skills developed over the course of their engagement with ASIM, and to document the self-directed learning behaviors underlying that development. Four sources of qualitative evidence were collected for this purpose:

- Individual in-depth interviews with a purposively selected group of participants from both cohorts, focused on each participant's experience of engaging with ASIM across its three phases and their perception of their own developing writing and speaking abilities.
- Classroom observations, conducted in person and via Zoom, documenting the nature and quality of students' interaction with AI tools and cooperating teachers, and capturing observable changes in student performance and engagement across sessions.
- The researcher's reflective journal, maintained throughout the intervention period with a minimum of one entry per week, recording observations on model implementation, changes in student behavior and performance, and emerging patterns relevant to the research questions.
- Cooperating teachers' journals completed after each intervention session, providing an independent record of implementation quality, student engagement, and self-directed learning behaviors observed across both groups.

These four sources were analyzed thematically to identify recurring patterns in students' developmental trajectories, with particular attention to evidence that corroborates, qualifies, or contextualizes the quantitative gains observed in academic writing and speaking skills.

3.7 Research Instruments

Consistent with the quasi-experimental design embedded in an explanatory sequential mixed-methods approach used in this study, the instruments used in the study are grouped into quantitative measurement instruments, control variable instruments, and qualitative data collection instruments. Together, these instruments enable a thorough examination of the effectiveness of ASIM, while controlling variables that could potentially affect the outcome, and help researchers derive rich meaning from students' experiences with the model.

3.7.1 Quantitative Measurement Instruments

These constitute the main data collection devices of the quantitative phase and were used to test the two dependent variables of the study both before and after the intervention to comprise:

A. Academic Writing Skills Test

An academic writing test (see Appendix 4) was constructed by the researcher and validated by experts in English language education and curriculum development (see Appendix 5). The test

aimed to assess students' ability to produce a clear, grammatically accurate, and well-organized text in English. The scoring rubric assesses five components:

- Topic sentence construction.
- Supporting sentences.
- Concluding sentence.
- Grammar and sentence structure.
- Vocabulary.

The test was administered twice, as a pre-test before the intervention and a post-test after. To avoid possible test-retest effects arising from exposure to identical items, the pre-test and post-test were constructed as parallel forms: identical in type, format, length, required effort, and scoring rubric, but using different prompts. This design supports the interpretation that gains observed in post-test scores reflect genuine learning rather than familiarity with specific test items.

Raters scored each response using a holistic-analytic approach: they first read the response in its entirety to form a general impression of the writing, then assigned independent scores for each of the five components listed above, based on their complete reading of the piece rather than isolated sentences. This approach reflects the accepted principle in ESL writing assessment that a written composition is an integrated communicative act best evaluated through a combined holistic and analytic reading process.

B. English Speaking Skills Test

An oral proficiency exam (see Appendix 4) was developed and validated using the same process as the written exam and assessed students' ability to engage in structured academic speaking tasks embedded within an academic context aligned with the ASIM framework. The exam consists of two tasks: Task One (Responding to a Prompt), in which students speak for three minutes on an assigned topic following one minute of preparation, and Task Two (Interactive Conversation), in which students respond to five spoken questions drawn from an assigned question set.

As with the writing test, this instrument was administered both as a pre-test, prior to the intervention, and as a post-test, upon its completion. To prevent students from recalling task content between administrations and thereby confounding the results, the pre-test and post-test were constructed as parallel forms, identical in structure, format, task type, cognitive demand, and scoring criteria, but differing in topic and content.

Each task was scored independently. Task One was scored out of 25 points across four criteria: fluency and pronunciation, content coverage, vocabulary range, and grammatical accuracy. Task Two was scored out of 25 points, with each of the five questions assigned an independent score out of 5 points using a holistic performance descriptor. Combined, the two

tasks yielded a total speaking score out of 50 points. This independent scoring by task was intended to capture the distinct communicative demands of each task and to provide more specific diagnostic information about students' oral proficiency profiles. The complete scoring rubric is provided in Appendix 4.

3.7.2 Control Variable Instrument

This instrument is not a primary outcome measure; rather, it provides a measure of pre-existing differences between participants that may influence the level of engagement with ASIM and the outcomes produced by ASIM.

C. Pre-Intervention Technology Familiarity Questionnaire

A short structured questionnaire (see Appendix 6) was given to all 50 participants as part of the process to measure previous exposure/expertise with digital technologies and generative AI tools before the intervention. This questionnaire comprised two sections two sections:

Section 1: General Use of Technology

- Device ownership: smartphone, computer, tablet
- Average internet usage per day
- Digital apps and platforms most frequently used

Section 2: Generative AI Familiarity

- Prior knowledge of ChatGPT. (Yes/No)
- Prior use of either tool. (Yes/No)
- What was the purpose of your original use? (Academic, Personal, Other)
- How experienced are you with using AI tools (5 Point Likert 1: Not at all to 5: Very experienced)?

On the basis of responses participants were classified into three levels:

Beginner: No prior knowledge or use of AI tools

Intermediate: Occasional or irregular prior use

Advanced: Frequent and intentional prior use

This classification was used in statistical analyses as a control variable to account for pre-existing differences in technological familiarity that might affect the interpretation of pre and post test results.

The researcher conducted three online training sessions on zoom with all participants, regardless of prior level of familiarity, prior to beginning ASIM.

3.7.3 Qualitative Data Collection Instruments

These instruments were administered after the completion of the quantitative phase and were intended to articulate, interpret and contextualize the statistical reports, to be consistent with the explanatory sequential design.

D. Semi-Structured Interviews

Semi-structured interviews (see Appendix 7) were conducted with a purposively selected subset of six participants (three male, three female) drawn from both cohorts, at the completion of the full intervention period. This timing enabled participants to describe their complete ASIM experience rather than impressions limited to a partial, mid-intervention stage. Participants were selected using two combined criteria: their post-test performance levels, to ensure representation across high, moderate, and low levels of improvement within each gender group, and their willingness to volunteer for the interview.

Parameter	Details
Participants	Six students purposively sampled from both cohorts (three male, three female)
Selection criteria	Variation in post-test performance levels (high, moderate, low) combined with student volunteering
Number of interviews	One per selected student.
Timing	End of full intervention.
Duration	45–60 minutes
Recording method	Audio recorded (with the student's consent), annotated in field notes.
Analysis	Thematic analysis

The interview guide covered the following aspects or domains:

- Students' experience of ASIM three-phase model.
- Academic writing and speaking skills development.
- Difficulties experienced during the intervention.
- Development of self-directed learning skills and strategies.
- Attitudes towards AI-assistant language learning.

E. Structured Observation Protocol

During the intervention period, non-participant observations (see Appendix 8) took place in-person and remotely (via Zoom) in order to observe and document students' engagement with ASIM during the implementation stage.

Parameter	Details
Type	Non-participant structured observation
Frequency	4-6 sessions across the intervention period
Delivery modes	In-person sessions, Zoom sessions, and community hall sessions
Analysis	Descriptive qualitative analysis and integrated thematic findings

Each observation session used structured observation protocols that included the following focus areas:

- Self-directed learning behaviors such as taking initiative, planning, and monitoring.
- Quality and nature of interaction with generative AI tools.
- Strategies used to overcome language difficulties.
- Engagement levels, motivation levels, and persistence levels.
- Context and situation of learning.
- Interpretive comments and timestamps of significant events

F. Reflective Journals — Researcher and Cooperating Teachers

During the intervention period, the two volunteer cooperating teachers worked alongside the researcher using a structured reflective journal (see Appendix 9), based on a framework developed by the researcher. The researcher completed one entry per week, while each cooperating teacher completed one entry after every intervention session, together yielding an average of two to three entries per week across both journals. This unified format provided a standardized basis for data collection across all three contributors.

Unlike the interview data, which were examined through thematic analysis, the reflective journals were analyzed as a chronological evidentiary record, tracing individual students' and groups' progress across the intervention period. Entries were reviewed sequentially to identify documented instances of change in students' writing, speaking, and self-directed learning behaviors, which were then used as illustrative evidence to corroborate and contextualize the quantitative findings, rather than being aggregated into cross-cutting thematic categories.

Prior to the intervention, the researcher trained both cooperating teachers on how to complete journal entries and maintained regular supervisory contact with them throughout implementation to ensure their entries were consistent, timely, and substantively documented.

Parameter	Details
Contributors	Researcher + two cooperating teachers
Frequency / Researcher	At least one entry per week
Frequency / Teachers	After every intervention session (2–3 per week)
Format	Unified digital journal

Supervision	Ongoing by researcher
Analysis	Chronological evidentiary review, integrated with interview and observation data

Journal entries were reviewed sequentially by contributor and by week to construct a documented, time-ordered account of each participant group's progress across the intervention period. Rather than being coded into cross-cutting thematic categories, entries were treated as a longitudinal evidentiary record, from which specific, dated observations were drawn to illustrate and substantiate students' developmental trajectories in writing, speaking, and self-directed learning. These documented instances were subsequently read alongside the thematically analyzed interview and observation data to provide a triangulated, evidence-based interpretation of the quantitative findings.

Each journal entry included information in the following areas:

- Observations on model application and practical implementation difficulties.
- Observed changes in student performance and behavior across lessons.
- Decisions and modifications made during implementation, with justification for these changes.
- Specific, dated instances illustrating students' progress or difficulty in relation to the research questions.
- Points flagged for continued tracking in subsequent entries.

3.8 Study Variables

3.8.1 Independent Variable

The independent variable of the study, the AI-Student Inversion Model (ASIM), is a theoretically-informed, researcher derived self-learning instructional model designed to infuse generative artificial intelligence tools into a strategic three-phased learning sequence. ASIM is operationalized through implementation across both participant groups over an 8-12 week period in the 2025/2026 scholastic year.

3.8.2 Dependent Variables

The dependent variables both measured through pre-tests and post-test proceeding and subsequent to exposure to ASIM are as specified in Chapter One:

A. Academic Writing Skills

Operationalized in this study across five measurable dimensions:

- Content and idea development.
- Organization and coherence.
- Grammatical accuracy.
- Vocabulary range and appropriateness.
- Writing mechanics.

B. English Speaking and Conversation Skills

Operationalized in this study across five measurable dimensions:

- Fluency & spontaneity of expression.
- Pronunciation & intelligibility.
- Grammatical accuracy in spoken production.
- Vocabulary range & contextual appropriateness.
- Communicative competence & interaction strategies.

3.8.3 Moderating and Control Variables

The following variables are controlled or monitored in order to reduce their potential influence on the dependent variables and to enhance the internal validity of the study

A. Students' Initial Language Proficiency

It is measured by means of writing and speaking pre-tests given before the intervention. Pre-test scores were treated in the analysis as a covariate which partitions out pre-test differences in English language proficiency between and within groups.

B. Prior Experience with Digital Technologies and Generative AI

It was evaluated through the Pre-Intervention Technology Familiarity Questionnaire distributed to all 50 participants prior to the beginning of the intervention. Participants were divided into three tiers (beginner, intermediate, advanced) of digital experience familiarity, and this tier will be entered into the statistical analysis as a control variable. To counteract the effects of a potential difference in level of experience (whether one group is more familiar with ChatGPT than another), the researcher held one or two training sessions held via Zoom before the intervention proper to ensure all students have some experience with ChatGPT before ASIM starts.

C. Overall Academic Performance

Students' general academic performance level was tracked as a background control variable in order to account for the possibility that being academically high- or low-achieving overall

(independent of command of English) may influence students' aptitude to engage with ASIM's cognitively demanding tasks.

3.9 Validity and Reliability

The current study's credibility, trustworthiness, and scientific rigour rely on having valid and reliable research materials. Mix-methods research uses different approaches to addressing validity and reliability for quantitative and qualitative measurements since each approach's epistemological assumptions are different.

3.9.1 Validity

A. Content Validity — All Instruments

All instruments developed for this research (Academic Writing Skills Test, English Speaking and Conversation Skills Test, Pre-Intervention Technology Familiarity Questionnaire, Semi-Structured Interview Guide, Structured Observation Protocol, and both reflective journal frameworks) received expert content validation prior to field use. Each instrument submitted to the expert review was given to a panel of specialists who hold expertise in English language education, applied linguistics, and educational technology. They reviewed the instruments for the following criteria:

- Clarity and understandability of language and instructions.
- Item relevance to the construct being measured.
- Comprehensive coverage of the skills or theme or construct being assessed.
- Appropriateness for the participant's age and educational context.

Reviewer feedback was incorporated into the final versions of all instruments before any data was collected.

B. Face Validity — Quantitative Tests

Besides specialist evaluation of the content, the writing and speaking assessments were evaluated for face validity to confirm their relevance and clarity in the context of Palestinian public high schools and eleventh graders. Minor language or format updates were made following this analysis to optimize accessibility and minimize possible causes of anxiety or confusion with future tests.

C. Parallel Forms Validity: Writing and speaking test

To reduce the potential for recall bias due to repeated exposure to the tests, the Academic Writing Skills Test and the English Speaking and Conversation Skills Test were developed in

parallel forms for the pre-test and post-test so it was necessary to review both these parallel forms with the expert panel to ensure that they are comparable in terms of cognitive demand, difficulty, structure and scoring criteria, but sufficiently different in content and prompts to avoid recall effects.

D. Pilot Study Validity — Quantitative Tests and Questionnaire

The pilot study was performed before the main intervention on a sample of 15 students who were in eleventh grade from a school that was not included in the main study but was still in the same governorate. The purposes of the pilot study included:

- To verify that the test instructions and items were clear and understandable.
- To determine an appropriate time for each test.
- To determine the difficulty and discrimination indices of writing and speaking tests.
- To calculate the internal consistency of the Technology Familiarity Questionnaire using the Cronbach's Alpha coefficient.
- To identify and resolve any logistical or procedural concerns before the full implementation of the tests.

The instruments were revised based on the pilot study results prior to administering them to the main study sample.

E. Credibility — Qualitative Instruments

The qualitative strand of this research prioritized credibility, that is, the extent to which the findings accurately reflect participants' experiences and perspectives. To establish credibility, the following strategies were used:

1. Methodological triangulation across multiple qualitative data collection methods. Four complementary sources of data were used, each contributing a distinct vantage point on students' experience of ASIM:
 - a. Semi-structured interviews, providing participants' own retrospective accounts of their experience, perceptions, and challenges across the full intervention.
 - b. Structured observations, capturing real-time, externally observed evidence of student-AI and student-teacher interaction that participants themselves might not report or recall.
 - c. The researcher's reflective journal, providing a longitudinal, session-by-session record of implementation and observed change from the researcher's perspective.

d. Cooperating teachers' reflective journals, providing an independent, classroom-embedded perspective on student engagement and performance, distinct from the researcher's own observations. During analysis, interview and observation data were examined through thematic analysis to identify recurring patterns in participants' experiences, while the reflective journals were reviewed chronologically as an evidentiary record, as described in Section 3.7.3.F, above. Findings from these sources were then read together: journal entries were used to corroborate or contextualize themes emerging from the interviews and observations, and points of convergence or divergence across sources were noted to strengthen the credibility of the overall qualitative interpretation.

2. A subgroup of participants was invited to review their interview transcripts and preliminary thematic interpretations, to confirm that the findings accurately represented their intended meaning and experience (member checking).

3. The researcher maintained sustained engagement with participants over the twelve-week intervention period, which supported a more contextually grounded understanding of participants' experiences and reduced the likelihood of researcher misinterpretation (prolonged engagement).

3.9.2 Reliability

To ensure the consistency and dependability of scoring the writing and speaking performance tests, both intra-rater reliability and inter-rater reliability were established.

Intra-rater reliability was examined to ensure the consistency of the researcher's scoring over time. The researcher rescored all students' responses on both the writing and speaking pre-tests and post-tests after an interval of two weeks without referring to the initial scores. The degree of consistency between the two rounds of scoring was calculated using the Pearson correlation coefficient. The obtained coefficients were 0.912 for the writing test and 0.889 for the speaking test, indicating a high level of scoring consistency and confirming the stability of the researcher's application of the scoring rubrics over time.

To establish inter-rater reliability, two experienced English language teachers independently scored a representative sample of students' writing and speaking responses using the same analytic scoring rubrics adopted in the study. The degree of agreement between the researcher's scores and those of the external raters was calculated using the Pearson correlation coefficient. The obtained coefficients were 0.901 for the writing test and 0.856 for the speaking test, indicating good to high agreement between raters. These findings confirm that the scoring criteria were applied consistently by different raters and that the obtained scores were sufficiently reliable for the purposes of the study.

Overall, the obtained reliability coefficients demonstrate that the writing and speaking assessment instruments yielded consistent and dependable scores, supporting the reliability of the data collected for the present study.

Internal Consistency — Technology Familiarity Questionnaire

The Pre-Intervention Technology Familiarity Questionnaire's internal consistency was judged using the Cronbach's Alpha (α), which evaluates how well the individual items in the questionnaire consistently measure the same underlying construct. A minimum of 0.70 was established as an acceptable level of Cronbach's alpha for the educational research community as a general guideline for determining an instrument's internal consistency.

Dependability — Qualitative Instruments

Reliability in qualitative research can mean the extent that the research process is well-documented, consistent, and can be replicated in similar situations. The following methods were used to assure dependability:

Audit trail: All processes involved in the collection of data, analytical decision making, coding of data, and making methodological adjustments to the research were documented systematically. This provided an audit trail that would allow independent reviewers to evaluate the consistency of the research process.

Structured instrument: The use of structured observation protocols and structured reflective journals rather than purely open-ended documentation allowed the researcher to maintain a degree of consistency in the collection of data across the various sessions, raters, and settings.

Reflexivity: In addition to being a source of data, the researcher's reflective journal was a means of monitoring and documenting the researcher's own assumptions, decisions, and developing interpretations through the course of the study, thereby creating additional means through which to ensure transparency and dependability in the qualitative data analysis.

3.10 Ethical Considerations

The current study was carried out within the framework of ethical considerations pertaining to educational research that involves human participants. As this research was conducted among adolescent students in a secondary school setting and given the extraordinary political and socio-economic conditions existing in Palestine during the 2025/2026 academic year at the time of research, extra care was taken to protect all participants both as individuals and as members of a larger group.

3.10.1 Institutional and Official Permissions

Before collecting data, the researcher obtained an official letter of facilitation from AL-Quds University (The researcher's University) (see Appendix 3). This letter was presented to both participating school administrations to formally introduce the study and its aims. Both school administrations verbally approved the researchers' ability to conduct the research at their respective schools. No contact with participants or their families occurred until all required procedures were completed. Both school administrations verbally approved the research to take place following receipt of the official facilitation letter from the researcher's university and were therefore deemed adequate sufficient to continue with the research. The researcher followed the normal policies and procedures for conducting research in the institutional administrative environment for Palestinian public schools.

3.10.2 Informed Consent

Parental Consent:

Since all participants are under the age of eighteen,, informed consent was obtained from parents or legal guardians of the 50 students prior to the intervention. The researcher contacted each of the families directly and personally by telephone to explain in clear and straightforward terms the nature, purpose, and duration of the study and what their family could expect from the study. Parents were informed that:

- There is no requirement for their child to participate in this study.
- They have the ability to withdraw their child from this study at any point with no consequences.
- That their child will be asked for (audio recorded) interviews, and that all data collected will be used only for academic and scientific purposes research.
- That they are taking steps to protect their child's identity and the confidentiality of the data collected from the study.

Student Assent:

All students participating in the study were requested to give assent (in addition to parental consent) to take part in the study. All students were informed of their right to withdraw from the study at any time without penalty academically or in any other way. They will be viewed as having provided their ongoing assent to participate and be treated as an ongoing condition of their participation.

3.10.3 Confidentiality and Anonymity

The data obtained in the study (i.e. pre-test and post-test data and test results, interview transcripts and audio format, observation notes, and the reflective journal) were treated as strictly confidential, with the following procedures in place to safeguard the privacy of individuals involved: all participants were assigned a number that was used instead of their name on all data, audio recordings were stored only on a password protected device, which was only accessible by the researcher, raw data was never viewed by the school head or teachers, and all data will be kept securely for the time period required by institutional policy and shredded or deleted or other means of anonymising data. Likewise, the cooperating teachers that aided in the study were also coded in all research documents, and this was done to protect them in the published details of findings.

3.10.4 Voluntary Participation and Right to Withdrawal

Participation was completely voluntary for both students and cooperating teachers. Students were informed that whether they participated or chose not to participate would have no impact on their grades, standing, or relationship with their teachers or school. The two cooperating teachers participated voluntarily and had no professional requirement to take part. Both were free to withdraw their cooperation from the intervention at any stage without consequence.

3.10.5 Protection from Harm

The researcher took steps to ensure that no participant experienced psychological harm, or unjustifiably negative impacts on their educational or social wellbeing, from participating in the current study. Specifically:

- The intervention was designed so that all students would benefit educationally, no group being disadvantaged from access to the ASIM model.
- Tasks involving AI tool tenderness while ensuring students were not exposed to harmful, misleading, or inappropriate AI-generated outputs
- The additional face-to-face sessions took place in the community hall with parental consent and were supervised, safe and fit-for-purpose
- The researcher remained attentive to wellbeing throughout the study making adjustments where necessary.

3.10.6 Data Use and Research Integrity

All data from this study will be used only for the purposes of this doctoral dissertation. Findings will be reported honestly, avoiding spin and selective reporting. There is no conflict of interest for the researcher in relation to the participants, schools, or AI tools studied in this

research. The use of generative AI tools - ChatGPT, as objects of study in this research is fully disclosed and not intended as endorsements of any commercial tool.

3.10.7 Ethical Considerations in the Palestinian Context

The circumstances under which this study was conducted in the Palestinian context, including three-day school weeks, repeated closures, and approximately fifty days of remote learning, raised distinct ethical considerations, which the researcher addressed as follows:

- Supplemental sessions held at the community hall were organized only after parental permission had been obtained, consistent with the consent procedures described in Section 3.6 (in-person meetings held at each school in the presence of the school principal, supplemented by direct telephone contact for parents unable to attend), so that no student felt pressured to participate beyond the standard school context.
- Zoom was used for some remote sessions with awareness that not all students had access to private, consistent internet connectivity or study space; attendance was monitored in coordination with parents/guardians and cooperating teachers, and individual arrangements, including participation in face-to-face community-hall sessions, were made as needed.
- Cooperating teachers were informed of the voluntary nature of their participation, which is acknowledged accordingly in the relevant sections of this dissertation.

Chapter Four

The Results of the study

4.1 Introduction

This chapter presents the findings of the present study in accordance with the experimental method with a quasi-experimental design and a mixed approach (quantitative, qualitative), whereby quantitative results are presented first, followed by qualitative findings that serve to interpret and contextualize the statistical outcomes. The chapter is organized around the study's eight research questions and four null hypotheses, each addressed in turn through a structured sequence of descriptive statistics, inferential analyses, and qualitative thematic findings.

The first section below commences with descriptive input relating to the Pre-Intervention Technology Familiarity Questionnaire, providing a contextualized grounding of each participant's prior experience with digital technologies and generative AI tools, prior to the start of the ASIM intervention. Participants' scores on the academic writing and English speaking skills pretests and posttests are then presented for female ($n = 28$) and male ($n = 22$) participants separately. Within each group, scores on the two tests were compared using paired samples t-tests, while the two groups themselves were compared using independent samples t-tests. The quantitative phase concludes with the results of the structured observation checklist, graphically documenting students' self-directed learning performances, their manner of interaction with the AI, and language skill indicators at six sessions.

In the qualitative phase, thematic data derived from semi-structured interviews conducted with a purposively selected subgroup of participants after the intervention were integrated with data derived from the researcher's reflective journal and from the cooperating teachers' reflective journals. Quantitative data was analyzed by the researcher using SPSS with the level of significance set at $p < 0.05$, while the qualitative data was analyzed using thematic analysis that included open and axial coding in an iterative manner.

4.2 Pre-Intervention Technology Familiarity Questionnaire

The ASIM participants completed a short-structured questionnaire before the ASIM intervention. All 50 participants' prior digital technology and generative artificial intelligence experience were assessed as a way to provide an indicator for the baseline measurement of pre-existing differences in technology familiarity. This instrument was not designed to generate new primary research but rather serves as a variable for controlling any baseline differences that may have existed in prior technology use among participants. Fifty percent of the participants were classified as "beginner," thirty percent were classified as "intermediate," and twenty percent were classified as "advanced" with generative AI tools according to the results of the questionnaire. This classification matches the general Palestinian secondary school setting included in the theoretical framework, where having access to technology does not necessarily equate to pedagogical uses of that technology. This classification system was included as a covariate in the statistical analysis to ensure that any observed increase in post-test scores for students who participated was the result of ASIM and not the result of a prior advantage related to the student's familiarity with technology.

4.3 Results related to the first research question

Question 1: What is the effectiveness of the proposed self-learning model utilizing generative artificial intelligence (ASIM) in developing secondary school students' academic writing skills in English?

Hypothesis 1 (H_{01}): There are no statistically significant differences at ($\alpha \leq 0.05$) between eleventh graders' mean scores on the pre- and post-tests of English academic writing skills attributable to the implementation of the ASIM model.

To examine the effectiveness of the ASIM model in developing students' academic writing skills, a Paired Samples t-test was conducted to compare students' pre-test and post-test scores as shown in Table (4.1) below..

Table 4.1: Paired Samples t-test for Students' Pre- and Post-Test Scores in Academic Writing Skills

Test	N	Mean	SD	Std. Error Mean	T	Df	Sig. (2-tailed)
Pre-test Writing Scores	50	27.64	3.67	.52	-14.420	49	.000
Post-test Writing Scores	50	37.96	4.03	.57			

Results presented in Table (4.1) showed that the mean score of students on the pre-test was ($M = 27.64$, $SD = 3.67$), whereas the mean score on the post-test increased to ($M = 37.96$, $SD =$

4.03). The calculated t-value was (-14.420) with (df = 49) and a significance level of ($p = .000$), which is lower than the adopted significance level ($\alpha \leq 0.05$).

These findings indicate statistically significant differences between students' pre-test and post-test scores in favor of the post-test. Therefore, the null hypothesis was rejected.

The improvement in post-test mean scores suggests that the implementation of the ASIM model had a significant positive effect on developing eleventh graders' English academic writing skills.

4.4 Results related to the second research question

Question 2: Are there statistically significant differences between eleventh graders' mean scores on the post-test of English academic writing skills attributable to gender?

Hypothesis 2 (H_{02}): There are no statistically significant differences at ($\alpha \leq 0.05$) between eleventh graders' mean scores on the post-test of English academic writing skills attributable to gender.

An Independent Samples t-test was employed to investigate gender-based differences in post-test writing scores and the results are presented in Table (4.2)

Table 4.2: Independent Samples t-Test for Gender Differences in Post-Test Academic Writing Scores.

Gender	N	Mean	SD	T	Df	Sig. (2-tailed)
Female	28	40.07	3.16	5.156	48	.000
Male	22	35.27	3.40			

Results presented in Table (4.2) showed that female students achieved a higher mean score ($M = 40.07$, $SD = 3.16$) compared with male students ($M = 35.27$, $SD = 3.40$). The calculated t-value was ($t = 5.156$) with ($df = 48$), and the significance level was ($p = .000$), which is lower than the adopted significance level ($\alpha \leq 0.05$).

These findings indicate the presence of statistically significant differences between male and female students' post-test scores in English academic writing skills in favor of female students. Therefore, the null hypothesis was rejected.

The results suggest that female students benefited more from the implementation of the ASIM model in developing academic writing skills.

4.5 Results related to the third research question

Question 3: What is the effectiveness of the proposed self-learning model utilizing generative artificial intelligence (ASIM) in developing secondary school students' English-speaking skills?

Hypothesis 3 (H_{03})

There are no statistically significant differences at ($\alpha \leq 0.05$) between eleventh graders' mean scores on the pre- and post-tests of English-speaking skills attributable to the implementation of the ASIM model.

To examine the effectiveness of the ASIM model in developing students' English-speaking skills, a Paired Samples t-test was conducted to compare students' pre-test and post-test speaking scores as shown in Table (4.3) below.

Table 4.3: Paired Samples t-test for Students' Pre- and Post-Test Scores in English-speaking Skills

Test	N	Mean	SD	Std. Error Mean	T	df	Sig. (2-tailed)
Pre-test (Responding to a Prompt)	50	13.46	1.96	.28	-13.182	49	.000
Post-test (Responding to a Prompt)	50	18.88	2.29	.32			
Pre-test (Performance Descriptors)	50	13.30	2.06	.29	-16.467	49	.000
Post-test (Performance Descriptors)	50	19.20	2.20	.31			
Speaking total pre-test	50	26.76	3.55	.50	-16.237	49	.000
Speaking total post-test	50	38.08	4.13	.58			

The findings displayed in Table (4.3) revealed noticeable improvement in students' speaking performance after implementing the ASIM model. For the "Responding to a Prompt" task, students' mean score increased from ($M = 13.46$, $SD = 1.96$) in the pre-test to ($M = 18.88$, $SD = 2.29$) in the post-test. The paired-samples t-test produced a value of ($t = -13.182$) with ($df = 49$) and a significance level of ($p = .000$). Since the obtained significance value is lower than the established alpha level (0.05), the improvement was statistically significant.

Likewise, students demonstrated higher achievement in the "Performance Descriptors" task. The mean score rose from ($M = 13.30$, $SD = 2.06$) before the intervention to ($M = 19.20$, $SD = 2.20$) after the intervention. The statistical analysis showed a significant difference between the two administrations, as indicated by ($t = -16.467$, $df = 49$, $p = .000$), where the significance value was below ($\alpha \leq 0.05$).

Concerning the total speaking score, the results also reflected substantial progress. The pre-test mean score of ($M = 26.76$, $SD = 3.55$) increased to ($M = 38.08$, $SD = 4.13$) in the post-test. The obtained t-value was ($t = -16.237$) with ($df = 49$) and a significance value of ($p = .000$), which is less than the significance criterion of (0.05). Therefore, the difference was statistically significant in favor of the post-test.

In light of these results, the null hypothesis was rejected. The statistical evidence indicates that the ASIM model was effective in enhancing eleventh graders' English-speaking skills across all evaluated speaking dimensions.

4.6 Results related to the fourth research question

Question 4: Are there statistically significant differences between eleventh graders' mean scores on the post-test of English-speaking skills attributable to gender?

Hypothesis 4 (H_{04})

There are no statistically significant differences at ($\alpha \leq 0.05$) between eleventh graders' mean scores on the post-test of English-speaking skills attributable to gender.

An Independent Samples t-test was utilized to compare male and female students' post-test speaking scores and the results are shown in Table (4.4) below.

Table 4.4: Independent Samples t-Test Results for Gender Differences in Students' Post-Test English Speaking Skills

Test	Gender	N	Mean	SD	T	Df	Sig. (2-tailed)
Responding to a Prompt	Female	28	19.92	1.92	4.234	48	.000
	Male	22	17.54	2.04			
Performance Descriptors	Female	28	20.17	2.05	4.064	48	.000
	Male	22	17.95	1.73			
Speaking Total	Female	28	40.10	3.53	4.676	48	.000
	Male	22	35.50	3.36			

As presented in Table (4.4), female students obtained higher mean scores than male students in all post-test speaking measures. In the "Responding to a Prompt" task, female students achieved a mean score of ($M = 19.92$, $SD = 1.92$), whereas male students obtained a mean score of ($M = 17.54$, $SD = 2.04$). The calculated t-value was ($t = 4.234$) with ($df = 48$). The significance value was ($p = .000$), which is lower than the adopted significance level ($\alpha \leq 0.05$). Therefore, the difference was statistically significant in favor of female students.

Similarly, in the "Performance Descriptors" task, female students recorded a higher mean score ($M = 20.17$, $SD = 2.05$) compared with male students ($M = 17.95$, $SD = 1.73$). The independent-samples t-test yielded a t-value of ($t = 4.064$) with ($df = 48$) and a significance level of ($p = .000$), which is less than ($\alpha \leq 0.05$). This indicates a statistically significant difference between the two groups in favor of female students.

Regarding the total speaking score, female students achieved a higher overall mean score ($M = 40.10$, $SD = 3.53$) than male students ($M = 35.50$, $SD = 3.36$). The obtained t-value was ($t = 4.676$) with ($df = 48$), and the significance level was ($p = .000$), which is lower than the

adopted alpha level (0.05). Accordingly, the difference was statistically significant in favor of female students.

Based on these findings, the null hypothesis was rejected. The results indicate that statistically significant gender-based differences existed in students' post-test English speaking skills following the implementation of the ASIM model, with female students outperforming male students across all speaking measures.

4.7 Results related to the fifth research question

Question 5: What are eleventh-grade students' experiences and perceptions of engaging with the three phases of ASIM, including their development of self-directed learning behaviors (self-monitoring and autonomous AI-directed learning)?

4.7.1 Quantitative Findings: Self-Directed Learning across Sessions

Level of Students' Self-Directed Learning Behaviors across ASIM Sessions

Descriptive statistics were computed to determine the overall level and developmental pattern of students' self-directed learning behaviors across the six ASIM intervention sessions.

Table 4.5: Descriptive Statistics for Self-Directed Learning Behaviors across Intervention Sessions

Session	Mean	SD	N
Session 1	31.67	5.96	6
Session 2	31.83	5.04	6
Session 3	34.67	6.44	6
Session 4	39.50	7.50	6
Session 5	40.83	6.52	6
Session 6	46.50	4.81	6

The descriptive results show a clear upward trend in students' self-directed learning behaviors across the ASIM intervention sessions. The mean scores increased steadily from Session 1 ($M = 31.67$) to Session 6 ($M = 46.50$), indicating continuous improvement in students'

engagement with AI interaction, self-monitoring, and language development skills throughout the intervention period.

Overall Effect of Sessions

To determine whether the observed differences across the six sessions were statistically significant, a one-way repeated measures ANOVA was conducted.

Assumption Check (Sphericity)

Mauchly's Test of Sphericity indicated that the assumption was violated ($W = .000$, $\chi^2(14) = 27.95$, $p = .039$). Therefore, Greenhouse–Geisser corrected values were used for interpretation.

Table 4.6: Tests of Within-Subjects Effects for Session Improvement

Source	Type	SS	df	MS	F	Sig.	Partial Eta Squared
Session	Greenhouse-Geisser	1021.67	1.78	573.70	23.58	.000	.825
Linear contrast	—	963.09	1	963.09	138.49	.000	—
Error	Greenhouse-Geisser	216.67	8.90	24.33	—	—	—

The repeated measures ANOVA revealed a statistically significant effect of session on students' self-directed learning behaviors, $F(1.78, 8.90) = 23.58$, $p = .000$. This indicates that students' performance differed significantly across the six ASIM intervention sessions.

In addition, the linear contrast indicated a significant upward trend across sessions, $F(1, 5) = 138.49$, $p = .000$, confirming that the improvement followed a consistent and progressive pattern.

The effect size indicated a very large impact of the intervention ($\eta^2 = 0.825$), suggesting that the ASIM sessions accounted for a substantial proportion of the variance in students' self-directed learning behaviors.

Overall, the findings indicate a statistically significant and sustained improvement in students' self-directed learning behaviors throughout the ASIM intervention. The results suggest that the

ASIM model was associated with continuous development in students' AI interaction, self-monitoring, and language skill development across the intervention period.

Comparison of Session 1 and Session 6 across Dimensions

To further examine the development of students' self-directed learning behaviors under the ASIM intervention, paired-samples t-tests were conducted comparing Session 1 and Session 6 across the overall score and its three key dimensions: AI interaction, self-directed learning behaviors, and language skill development.

Table 4.7: Paired Samples t-Test Results for Session 1 vs Session 6 across Dimensions

Dimension	Session	N	Mean	SD	Std. Error Mean	T	df	Sig. (2-tailed)
Overall Score	Session 1	6	31.67	5.96	2.43	-6.37	5	.001
	Session 6	6	46.50	4.81	1.96			
AI Interaction	Session 1	6	9.50	2.07	0.85	-5.32	5	.003
	Session 6	6	14.00	1.67	0.68			
Self-Directed Learning	Session 1	6	12.33	2.80	1.15	-4.79	5	.005
	Session 6	6	17.83	2.56	1.05			
Language Skills	Session 1	6	9.83	1.60	0.65	-8.91	5	.000
	Session 6	6	14.67	0.82	0.33			

The results of the paired-samples t-tests presented in Table (4.7) indicate a statistically significant improvement from Session 1 to Session 6 across all measured dimensions of self-directed learning behaviors.

At the overall level, students' mean score increased markedly from 31.67 in Session 1 to 46.50 in Session 6, with a statistically significant difference observed, $t(5) = -6.37$, $p = .001$. This finding demonstrates a strong positive effect of the ASIM intervention over time.

Similarly, all sub-dimensions showed statistically significant gains. AI interaction skills improved significantly from 9.50 to 14.00 ($p = .003$), reflecting increased engagement with generative AI tools throughout the intervention. Self-directed learning behaviors increased

from 12.33 to 17.83 ($p = .005$), indicating enhanced learner autonomy, goal-setting, and self-monitoring abilities. In addition, language skill development showed the strongest improvement, rising from 9.83 to 14.67 ($p = .000$), suggesting substantial gains in students' linguistic performance and communicative output.

Since all p-values are below the significance threshold of $\alpha \leq 0.05$, the null hypothesis of no difference between Session 1 and Session 6 is rejected across all dimensions.

4.7.2 Qualitative Findings: Structured Observation Notes

Apart from the scores recorded in the numeric scoring of the checklist, the 'score' in qualitative terms was recorded by the observer's field notes in Section D of each completed checklist form. These notes are analyzed thematically in this section to show how students engaged with each of ASIM's phases during the six observed sessions and thus provide context for and elaboration of the numerical results we reported above.

Increased Learner Initiative in AI Interaction

Observation notes recorded a shift in the nature of students' engagement with AI tools across the six sessions. In Sessions 1 and 2, field notes across all six observed students described interactions as predominantly reactive: students accepted the AI's first response without further questioning and showed limited evidence of directing the interaction themselves. By Sessions 5 and 6, notes recorded longer, more elaborated prompts and, during Phase 3 specifically, sustained turns in which students maintained the AI in a learner role across multiple exchanges.

This improvement, while consistent in direction across all six students, was not uniform in degree. Field notes indicate that some students (e.g., those identified in Theme 2 below as advanced-level, such as F21 and M19) showed marked and rapid gains in the sophistication of their AI interaction, while others showed more modest, incremental change across the same sessions. No student's engagement pattern moved in the opposite direction; the variation observed was one of pace and extent of improvement rather than of direction.

This documented shift, and its uneven pace across students, provides qualitative context for the quantitative gains reported in the AI Interaction subscale of the structured observation checklist (Table 4.7) and for the individual variation reflected in its standard deviations.

AI Confusion Strategy as Trigger for Metalinguistic Awareness

Observation notes were particularly detailed during Phase 3 (Sessions 5 and 6), when students applied the AI Confusion Strategy. Both observers recorded a shift in students' attentiveness to AI-generated errors, noting increased care and longer written comments compared to the more rapid, less deliberate responses recorded in earlier sessions. Advanced-level students (F21,

M19) were observed constructing increasingly sophisticated error-detection prompts; one such student instructed the AI to "delimit its brain" and produce a "metalinguistic hash of itself" before generating a deliberate error, and subsequently provided extended metalinguistic corrections in response. Students identified as beginner-level in the pre-intervention technology familiarity questionnaire (F05, M03) approached the task more tentatively at first, but by Session 6 were engaging with AI-generated errors substantially more often than in earlier sessions. This pattern suggests that, while the AI Confusion Strategy places greater cognitive demands on students, it was associated with observable learning gains across proficiency levels, not only among advanced-level students.

Gender Differences in SDL Behavioral Patterns

While the upward trend in self-directed learning behavior was observed across all six sessions for both groups, analysis of the observation notes indicated differences in how female and male students exercised these behaviors. Female students were more frequently observed setting explicit goals before a session began, verbalizing self-assessments during tasks, and persisting with difficult aspects of the AI Confusion Strategy in Phase 3. Male students were more frequently observed excelling in AI interaction specifically, particularly in prompt refinement and critical engagement with AI-generated output, but were noted as requiring more prompting from teachers to engage in self-monitoring and reflection on their own strengths and weaknesses. These observed patterns are consistent with female students' numeric advantage on both the checklist scores and the pre-/post-test results reported in Sections 4.3 through 4.6, and suggest that the gender difference observed in this study may be more pronounced in self-monitoring and reflective behaviors than in AI interaction skills specifically.

4.7.3 Qualitative Findings: Semi-Structured Interviews

The qualitative phase of this research question also drew on semi-structured interviews with a purposively sampled group of six students (three female, three male) from the two participating schools, representing different proficiency levels (advanced, average, and beginner/intermediate, as classified in Section 4.2). Interviews were conducted upon completion of the full ASIM intervention, audio recorded with participant consent, and transcribed for thematic analysis. The interview guide addressed both students' experiences and perceptions of engaging with ASIM's three phases and the development of their self-directed learning behaviors, consistent with the dual focus of Research Question 5. Thematic analysis produced five themes, presented below.

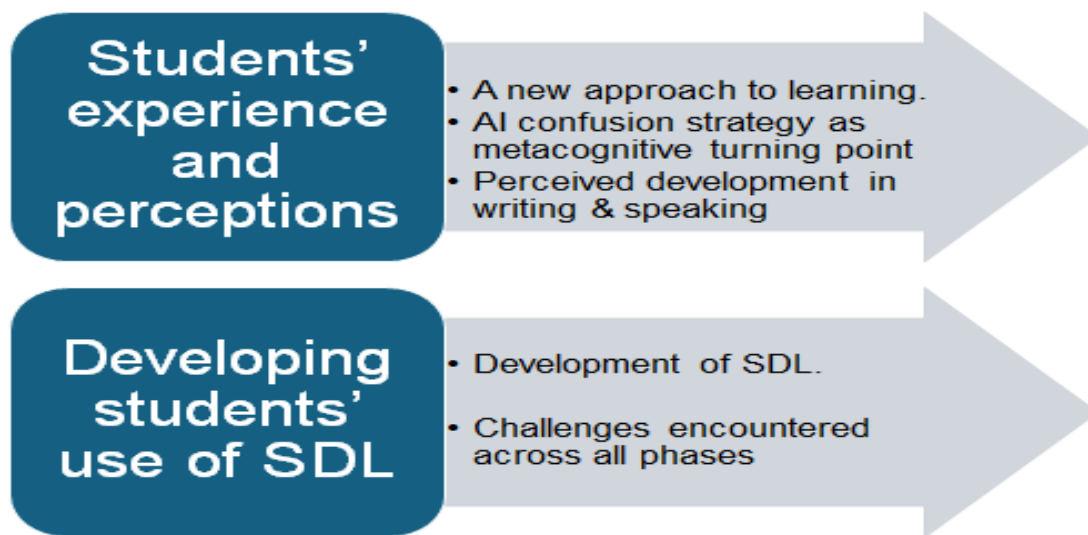


Figure 4.1_ thematic map: Semi-structured interview findings

Theme 1: A New Approach to Learning

The six interviewed students consistently described ASIM as requiring a different mode of engagement than their regular, teacher-centered classroom experience. A female advanced-level student described this shift in agency: "In regular class, the teacher teaches me and I just copy and learn. In ASIM, I decided what to do, I told the AI what I wanted to do, and I made it work for me. I felt like I was actually doing the learning instead of just preparing for a test." A male intermediate-level student described a corresponding shift in cognitive demand: "At first, I thought I would find it easier to do the work because there was no teacher there to watch me. However, I found it to be more challenging in a good way... I had to ask the AI the right question to get the answer."

Phase 3 elicited particularly strong responses. Two students described it as the most memorable and cognitively demanding component of the intervention. A female advanced-level student explained: "When I had to explain the grammar rule to the AI as if the AI did not understand it, I realized I did not understand the grammar rule as well as I thought I did. Teaching the AI made me understand the rule better than I have ever done with any other activity." Taken together, these accounts suggest that, for this subsample, ASIM was perceived as a departure from passive learning, one that students described as initially challenging but ultimately more engaging than their prior instructional experience.

Theme 2: The Turning Point of the AI Confusion Strategy as a Metacognitive Strategy

The AI Confusion Strategy - the third phase of the ASIM experience where learners stimulated errors in the AI and then identified, clarified and corrected them - was identified by all six participants as an effective and memorable element of the ASIM experience. The cognitive processing of the confusion strategy by the participants was reported to be qualitatively different from their cognitive processing of the conventional error presentation and correction tasks.

An intermediate level female learner provided insight into her perspective: "I had to explain the rule that made the mistake an error when I found an AI answer intentionally. I was trying to think of things for the very first time based on my own knowledge instead of what my teachers had told me as mistakes."

An average level male learner described his struggles with the confusion activity in session 5 of the ASIM experience: "In the beginning, I was not sure what type of error I should be looking for. Then by the fifth session of the ASIM experience, I could give the AI an example of a type of mistake, like 'Show me an error in the present perfect tense.' Then I would be able to find the error and explain it to the robot. It became like a game, and I learned from it."

More advanced students provided additional reflections of the metacognitive significance of the confusion strategy. A female learner stated: "I know that when I helped correct mistakes made by the AI I was able to stop making that type of mistake in my own writing. When I do! I correct a mistake for the teacher it feels like I do not learn anything." Those responses indicate that for many of the participants, the confusion strategy served as a metacognitive turning point and a structured opportunity to develop from recognition of surface-level errors to the metalinguistic analysis of errors as viewed through the Phase 3 improvement data gathered from the ASIM Classification System.

Theme 3: Perceived Development in Writing and Speaking Skills

All six interviewed students reported that their academic writing and speaking skills had improved meaningfully through their engagement with ASIM, with writing improvements described in more specific and detailed terms than speaking improvements. These accounts reflect the perceptions of this six-student subsample and should not be read as representative of all fifty study participants; they are considered here alongside, and in relation to, the statistically significant quantitative gains reported in Sections 4.3 and 4.5.

Regarding writing, students most frequently cited vocabulary range, grammatical accuracy, and organizational structure as areas of noticeable improvement. A female advanced-level student explained: "Now I write, check it by using ChatGPT, understand what is wrong and why, and then go back to writing. My paragraphs are much more organized." A male intermediate-level student attributed his improvement to the iterative feedback cycle: "The AI

did not just say 'wrong,' it explained. Then I would go back and try again... I used to never do that; when we used to be in class, the teacher would give you a mark and just move on."

Regarding speaking, students described improvement primarily in affective terms, reporting reduced anxiety and increased willingness to communicate, often attributing this to practicing without fear of peer judgment: "I could make mistakes in using the AI and nobody laughed at me" (female, average level); "After using ChatGPT for speaking I was less scared to speak in English, I no longer have fear" (male, beginner level). These accounts are broadly consistent with the statistically significant quantitative gains reported in Section 4.5, and add an affective dimension not directly captured by the test scores themselves.

Theme 4: From Teacher-Directed Dependence to Self-Regulated Practice

Interview data traced a discernible progression in participants' self-directed learning behaviors, moving from initial reliance on external direction toward independent goal-setting and self-monitoring, rather than a uniform or immediate shift toward autonomy. Early in the intervention, students described continued dependence on external direction: "In the beginning, it was hard. I kept waiting for someone to tell me what to do next" (female, average level). By later sessions, the same students described self-initiated planning and evaluation: "I started planning what I wanted to do before every session... I never did that before. The teacher always told me what to do" (female, intermediate level); "I learned to check my own work, not wait for the teacher to tell me, I was my own teacher and my own student" (male, advanced level). One student attributed this shift specifically to ASIM's structure rather than to AI access alone: "If someone just gave me ChatGPT without the model, the phases and the activities, I do not think I would have used it differently. The structure showed me how to learn by myself" (male, intermediate level).

Read alongside the quantitative findings in Section 4.7.1, this gradual pattern corresponds with the progressive, session-by-session increase in self-directed learning scores reported in Table 4.5: both data sources indicate that autonomy developed cumulatively across the intervention rather than emerging after a single phase or session. In addition, students' accounts of setting explicit goals and monitoring their own progress correspond with the pattern documented among female participants specifically in Section 4.7.2, offering a possible qualitative explanation for the statistically significant gender difference in post-test scores reported in Section 4.4.

Theme 5: Challenges Encountered During the Intervention

Participants described three categories of challenges. Technical challenges, related to internet access, device limitations, and the constraints of the free version of ChatGPT, were mentioned by all six students: "Sometimes ChatGPT stopped responding in the middle of a session... especially during Phase 3 when I was in the middle of teaching it something" (male, beginner

level); "The free version had a limit on how much you could write. Sometimes I had to start a new conversation and explain everything again" (female, intermediate level).

Linguistic challenges were most prevalent among beginner- and average-level students in the early sessions: "At first I did not know how to ask the AI what I wanted. My English was not good enough to explain clearly" (male, average level). One student noted that this difficulty became a source of incidental language practice: "Writing the prompts was like a writing exercise itself. By the end, I could write better prompts, which also meant my English was better" (female, beginner level).

Cognitive challenges, particularly the demands of Phase 3 and the AI Confusion Strategy, were acknowledged but generally framed as productive: "Teaching the AI was the hardest thing I did in this model. But it was also the most useful. I think the difficulty was the point" (female, advanced level). When asked for suggestions, students most frequently requested access to a fuller version of ChatGPT, additional preparatory training in prompt formulation prior to Phase 1, and further opportunities for speaking practice during Phase 2.

4.8 Results related to the sixth research question

Question 6: How do the researcher and cooperating teachers perceive the implementation of ASIM, in terms of students' engagement patterns and the contextual challenges encountered during the intervention?

4.8.1 Reflective Journals —Researcher and Cooperating Teachers

The integrated reflective journal framework was completed throughout the twelve-week intervention period by the researcher, who recorded one entry per week (12 entries total), and by the two volunteer cooperating teachers, one at each school, who completed one entry after each intervention session (28 entries total across both teachers). Consistent with the chronological evidentiary approach described in Section 3.7.3.F, the three journals were reviewed sequentially by week rather than coded thematically. This review traced a consistent developmental arc across the intervention period, documented below in chronological order, and is presented alongside the thematically analyzed interview and observation data (Sections 4.7.2–4.7.3) to provide a triangulated account of ASIM's implementation.

To mitigate potential researcher bias in the interpretation of journal data, several procedural safeguards were applied. First, the researcher's and cooperating teachers' journals were reviewed independently before being compared, reducing the likelihood that the researcher's expectations shaped the teachers' documented observations. Second, the researcher maintained a separate reflexive log, distinct from the main observational entries, in which assumptions and expectations were noted explicitly and revisited throughout the intervention (see the Week

7 entry cited below, in which the researcher explicitly questioned her own tendency to search for confirming evidence). Third, a colleague with expertise in qualitative research methods independently reviewed a sample of journal excerpts and the researcher's interpretation of them, to assess the consistency of these interpretations (peer debriefing).

The patterns described below emerged from a sequential reading of journal entries and are presented as a documented, time-ordered account rather than as themes derived through coding; illustrative excerpts are dated by week to preserve this chronological structure.

Weeks 1–5: Difficulty and Gradual Resolution of Role Transitions

Across the first weeks of implementation, all three journals documented students' difficulty adjusting from a teacher-centered mode of participation to the self-directed, AI-mediated mode ASIM requires. The researcher noted in the Week 1 entry:

"Both groups seem uncertain about what they should do in Phase 1. They keep looking up, waiting for me or the teacher to tell them what they are supposed to do next. The idea of directing your own session seems genuinely an alien one. I need to keep telling myself that it is not failure, it is conditioning."

The cooperating teacher at the girls' school recorded a similar pattern in the early weeks, noting that some students submitted their Phase 1 self-diagnosis reports with vague, general statements, suggesting the metacognitive self-assessment task itself required a type of thinking students were unaccustomed to. By Week 4, both teachers' journals recorded a shift. The teacher at the boys' school wrote: "Something has changed this week. The students are actually commencing Phase 1 without being told to. Two students had begun their AI interaction before I was passing out the task sheet, not happened before." The researcher's Week 5 entry recorded a corresponding observation: "The targeted discussions emerging from the students' own diagnosed gaps seem to be doing something important. Students are beginning to see a relationship between what they diagnosed they do not know and what we talk about together."

This transition was not uniform across all students. Notes for one beginner-level male student (M03) indicated a slower trajectory: as late as Session 4, this student continued to wait for teacher prompting before initiating Phase 1 tasks independently, illustrating that the pace of this shift varied across individuals rather than following a single uniform pattern.

Recurring Pattern: Gender Differences in Engagement

Both teachers' journals contained repeated references, across multiple weeks, to differing patterns of engagement between male and female students. At the girls' school, the cooperating teacher recorded early and sustained engagement with the self-diagnosis and goal-setting components of Phase 1, with female students frequently submitting detailed reports identifying specific grammar or vocabulary targets. The researcher's own observations, recorded across several weeks, aligned with this pattern: "The female students are able to

identify specific grammar structures they had difficulty with, while the male students tended to use general statements such as 'I need to improve my writing.'" At the boys' school, the cooperating teacher instead recorded a pattern of rapid development in AI interaction and prompt refinement, including instances of students constructing AI personas independently of teacher suggestion. By Week 6, the researcher's journal noted this divergence explicitly: "A trend is beginning to emerge in that females are excelling within the reflective/self-monitoring portion of ASIM, while males exhibit greater spontaneity and creativity when interacting with the AI." This pattern, documented consistently from the early to the later weeks of the intervention, is consistent with the quantitative gender difference reported in Table 4.4 and with the interview accounts discussed in Section 4.7.3

Weeks 7–9: The AI Confusion Strategy as a Documented Turning Point

Entries from Weeks 7 through 9, across all three journals, documented a marked change in student engagement coinciding with the introduction of the AI Confusion Strategy in Phase 3. In the Week 7 entry, the researcher described observing a student (F12) engage with an AI-generated text for approximately 14 minutes, correcting errors in a notebook before returning to explain each correction to the AI, engagement the researcher described as unlike anything observed during Phases 1 or 2. The cooperating teacher at the girls' school documented increased participation from a previously disengaged student during this period, while the cooperating teacher at the boys' school recorded an instance of unprompted peer-to-peer discussion, lasting approximately eight minutes, in which two students independently resolved a disagreement about an AI-generated sentence.

As part of her reflexive log during this period, the researcher recorded an explicit acknowledgment of the risk of confirmation bias: "I recognize that I am searching for evidence supporting the effectiveness of ASIM. I must remain vigilant that while I am using engagement as a means of measuring learning, the way in which students are working during Phase 3 indicates a level of engagement beyond rote memorization." Not all documented reactions to this strategy were positive; one beginner-level male student expressed continued discomfort with it even in the final week of the intervention, describing it in his interview as feeling "strange to correct something that is not really wrong."

Recurring Contextual Challenges Documented Throughout the Intervention

Three contextual challenges recurred consistently across the three journals throughout the twelve-week period. First, the three-day school week was repeatedly documented as disrupting continuity, with the researcher noting in a Week 3 entry that the reduced schedule caused students to lose track of their prior AI conversations between sessions. Second, remote sessions conducted via Zoom introduced recurring technical disruptions, documented by both teachers, though both also recorded that Zoom sessions produced some of the most self-directed behavior observed during the intervention, with one teacher noting that a normally

quiet student attempted new prompt types more freely outside the physical classroom. Third, the community-hall sessions were documented by both teachers as producing the highest levels of engagement observed during the intervention, attributed to the informal setting and the absence of institutional pressure; the cooperating teacher at the boys' school described one such session as one in which students "supported one another with prompts, discussed their responses to the AI, and remained engaged the entire time without any prompting."

Methodological Note on Qualitative Trustworthiness

Data saturation was not formally assessed given the study's small qualitative interview sample ($n = 6$); however, no substantially new patterns emerged in the final two interviews relative to the preceding four, suggesting the core patterns identified were reasonably stable within this subsample. This is acknowledged as a limitation in Section 1.8. The patterns reported across Sections 4.7 through 4.8 emphasize experiences that recurred across participants and across journal entries; however, as noted above, not all experiences were uniform, and instances of slower transition, discomfort, or disagreement were retained and reported rather than excluded. The relative consistency of responses may partly reflect the purposive sampling strategy described in Section 3.7.3.D and the small size of the interviewed subsample, and should not be read as evidence that no divergent experiences occurred among the broader group of fifty participants. Qualitative findings were integrated with the quantitative results through explanation-building, consistent with the logic of the explanatory sequential design (Section 3.6): qualitative and journal-based evidence was examined specifically for its capacity to account for the statistically significant patterns reported in Sections 4.3 through 4.6, as illustrated in Section 4.7.3.

Chapter Five

Discussion of the Results

5.1 Introduction

This chapter presents the results of the current study in relation to its eight research questions while placing the results in context with the empirical literature reviewed in chapter two and the theoretical framework developed therein. The format for the discussion of the results follows the order of the eight research questions and integrates the quantitative and qualitative data where possible. Finally, in conclusion, an overview of the findings will be presented together with recommendations to teachers, developers of educational resources, and future researchers.

5.2 Discussions

Discussion of the Results of the first question: What is the effectiveness of ASIM in developing eleventh-grade students' English academic writing skills?

Results of the first research question showed that students' academic writing scores were significantly higher following the ASIM intervention, with statistically significant differences observed between pre-test and post-test scores across the combined sample of female and male participants, $t(49) = -14.42$, $p < .001$, after twelve weeks of ASIM implementation. This finding directly addresses the writing deficits documented in the problem statement (Section 1.2), where Palestinian secondary EFL learners were shown to struggle with limited linguistic competence, weak organization, and heavy reliance on translation from Arabic (Hammad, 2016). Significant improvement was observed across all five components of the writing rubric: topic sentence construction, supporting sentences, concluding sentence, grammar and sentence structure, and vocabulary. That gains were observed across all five components, rather than being concentrated in one or two, suggests that ASIM's three-phase structure engaged multiple dimensions of academic writing simultaneously, though the relative contribution of each phase to specific components was not isolated in this study and cannot be claimed with certainty from the present data.

The most comparable empirical evidence supporting the present study's findings comes from Jmaiel et al. (2025) and Song and Song (2023). Jmaiel et al. (2025) conducted an experimental study examining the impact of a ChatGPT-based pedagogical intervention on university students' writing (N = 117), reporting increases across all writing dimensions measured post-intervention. Similarly, Song and Song (2023) conducted a quasi-experimental study examining ChatGPT-assisted feedback for EFL university students' writing, finding statistically significant positive effects on writing organization, coherence, grammatical accuracy, and vocabulary range. The similarities between the present findings and those of Song and Song (2023) are particularly instructive, as both studies employed a comparable procedure: students submitted written work to an AI program, received specific feedback, and produced multiple drafts based on that feedback. This three-step process closely parallels the operation of Phase 2 within ASIM.

Mahapatra (2024) provides additional support for this finding through an examination of ChatGPT as a formative assessment tool providing feedback to ESL undergraduate students. That study's results indicate that structured student preparation prior to receiving AI-assisted feedback was an enabling condition for measurable writing gains. This finding carries specific implications for the design of ASIM's Phase 1, in which students are prepared with prompt engineering and AI-engagement skills before participating in Phase 2 writing tasks. This sequencing is particularly relevant to Palestinian EFL classrooms, where class sizes often exceeding 30 students make frequent individualized teacher feedback impractical (British Council & LearnJam, 2025). Poláková and Ivenz (2024) similarly found that ChatGPT-enhanced sessions had a statistically significant impact on both grammatical accuracy and writing concision, alongside positive shifts in students' perceptions of the tool as a viable learning resource, an affective dimension consistent with the present study's qualitative findings on student engagement, discussed in relation to Research Question 5.

Lo et al.'s (2024) systematic review of seventy studies on ChatGPT in education, conducted over the eighteen months following its release, situates the present findings within a rapidly coalescing body of scholarship: writing development represents the most widely studied outcome of ChatGPT-assisted instruction, and the evidence for its effectiveness, particularly when supported by structured pedagogical design, is among the most robust in the emerging literature on AI in education. However, few of the reviewed studies evaluated ChatGPT's quantitative impact using objective, pre-post measures, a methodological limitation that the present study's pre-test and post-test design directly addresses.

However, these positive findings should be considered alongside more cautionary evidence presented by Sapan and Uzun (2024), who conducted a mixed-methods study with 66 tenth-grade Turkish secondary school students, comparable secondary-level students to the eleventh-grade Palestinian participants in the present study. Their study found that traditional instruction yielded greater writing and vocabulary development than unstructured integration of ChatGPT. This finding offers an empirical counterpoint to any assumption that access to AI

alone produces writing gains, and provides a useful interpretive frame for the present findings: the improvements observed in this study are attributable to the structured pedagogical architecture within which ChatGPT and Grammarly were implemented, rather than to the tools themselves. This interpretation is further supported by the qualitative finding discussed in relation to Research Question 5, in which students attributed their development to ASIM's structured phases rather than to AI access alone.

Interpreting these empirical findings within both ASIM's theoretical foundations and the specific structural characteristics of Palestinian secondary EFL education offers additional insight. ASIM's constructivist design, grounded in Vygotsky's (1978) Zone of Proximal Development and operationalized through Collins et al.'s (1989) Cognitive Apprenticeship sequence, positions the AI as a scaffolding device whose presence recedes as students gain competence. This progression is reflected in the sustained upward trajectory of writing scores observed across the intervention. The AI Confusion Strategy introduced in Phase 3 engages the evaluate and create levels of Anderson and Krathwohl's (2001) revised taxonomy, prompting the kind of metacognitive monitoring that Flavell (1979) identifies as central to sustainable language acquisition.

This structured approach is particularly significant given the systemic underperformance in writing outcomes documented among Palestinian EFL learners (British Council & LearnJam, 2025) and the risk, noted by Hamamra et al. (2024), that unstructured AI use in this context may foster dependency rather than genuine skill development. Read together, these findings suggest a contribution specific to this study: that writing gains associated with generative AI in a Palestinian secondary-school context depend on the presence of structured pedagogical design, such as that provided by ASIM's three phases, rather than on access to the AI tool alone.

Discussion of the Results of the Second Question: Are there statistically significant differences between eleventh graders' mean scores on the post-test of English Academic Writing Skills attributable to gender?

The second research question revealed statistically significant differences in post-test academic writing scores between female and male students, with female students obtaining higher mean scores than male students. This finding should be interpreted with an important methodological caveat: because female and male participants were drawn from two separate, single-sex schools rather than randomly assigned within a shared setting, gender in this study is confounded with school context. Observed differences may therefore reflect not gender alone, but also differences in school climate, teacher style, prior achievement, classroom culture, or variation in implementation between the two sites. This limitation, already noted in Section 1.8.

Within this constraint, the safest interpretation of the quantitative result is that female participants in this sample obtained significantly higher post-test writing scores than male participants; this study's design does not permit a comparison of pre-to-post gain scores by gender, as no such analysis was conducted, nor does it permit attributing this difference to gender independent of school context. Both groups showed statistically significant improvement from pre-test to post-test (Section 4.3), which is an important qualification: it indicates that ASIM was associated with meaningful writing gains for male participants as well, and that the gender difference observed reflects a difference in post-test performance level rather than evidence that male students did not benefit from the intervention.

The qualitative data offer a possible, though interpretive rather than confirmed explanation for this pattern. As discussed in relation to Research Question 5, journal and interview data suggested that female participants more frequently engaged in explicit goal-setting and self-monitoring behaviors, behaviors closely aligned with the self-directed learning processes ASIM is designed to support. This qualitative pattern is offered as one plausible contributing factor among several, rather than as a proven mechanism, and should be read alongside the confound noted above rather than as an alternative explanation that resolves it.

In this regard, Jmaiel et al. (2025) reported that female students showed larger gains than male students in their study of ChatGPT-assisted writing instruction, and suggested this pattern warrants more systematic investigation rather than being treated as an isolated finding. The parallel between Jmaiel et al.'s (2025) university-level Arab student sample and the present study's Palestinian secondary-school participants suggests that gender-related patterns in response to AI-assisted writing instruction may extend across different educational levels and contexts, although further research across a broader range of settings would be needed to establish this more firmly.

Based on the cooperative teachers' reflective journals, qualitative results provide some degree of possible explanation for this gender difference. The two teachers have observed that when engaging in self-diagnosis and goal-setting during Phase 1, female students produced more specific, more detailed, and accurately calibrated self-diagnosis reports about their linguistic gaps (i.e., identifying specific grammar rules, areas of specific vocabulary, and certain types of writing structures they wanted to improve). The fact that female students were better metacognitively prepared to use AI writing feedback tools in Phase 2 (i.e., being able to set specific goals) is demonstrated by the fact that females with clear goals could evaluate and respond to AI-generated feedback substantially better than students who did not have clearly defined goals.

This explanation matches with what Zimmerman (2002) discusses in his self-regulation model of learning (SRL), which identifies goal setting and strategic planning as the most significant determinants of performance when using technology in education (i.e., learners need to set specific goals before entering the performance phase). Female students who accessed the

performance phase with self-generated and specific goals were able to use AI-generated feedback more purposefully and cumulatively than students without specific, self-generated goals, who tend to passively receive feedback and process each response as an isolated event with no reference to a more general developmental path. Further support for this mechanism comes from an experimental study carried out by Alfares (2025) at Umm Al-Qura University, which showed that improvement in writing ability does not occur simply as a result of abstract autonomous learning, but instead from deliberately scaffolded autonomous learning experiences (i.e., through goal-setting and self-monitoring) functioning within a theoretical instructional framework - exactly the behaviors that the female students in the present study consistently engaged with from Phase 1.

However, it is important to point out that males exhibited equal or higher levels of engagement on another aspect within ASIM of the creativity and spontaneity of their use of AI and prompt engineering during Phase two. This difference was documented in teachers' journals and indicates that the difference in writing results between boys and girls is not based on one group being at an advantage to the other, but rather a difference in the different engagement patterns exhibited by each group based on the cognitive requirements of academic writing assessment. The metacognitive orientation for systematic self-monitoring of grammatical accuracy and structural integrity, which are two primary facets of the writing instrument, fit better for girls than for boys, as girls exhibited a more reflective style of engagement in their ASIM response than boys.

Discussion of the Results of the Third Question: What is the effectiveness of ASIM in developing eleventh-grade students' English speaking skills?

Results of the third research question showed statistically significant improvements in English-speaking scores from pre-test to post-test across the combined sample of female and male participants, for both Task One (Responding to a Prompt) and Task Two (Interactive Conversation), $t(49) = -13.18$ and $t(49) = -16.47$ respectively, $p < .001$. This finding indicates that ASIM was associated with meaningful development in students' oral communication, with gains observed across fluency, content coverage, vocabulary, and grammatical accuracy. That significant improvement was observed across both tasks, which differ in structure and cognitive demand, strengthens the interpretation that the observed gains reflect genuine development in speaking ability rather than an artifact of a single assessment format.

These results echo those of Fathi et al. (2024), whose controlled mixed-methods study of 65 EFL learners found that AI-mediated interactive speaking activities produced greater gains in speaking skills, fluency, coherence, grammatical accuracy, and willingness to communicate relative to peer-based activities. Drawing on qualitative data, Fathi et al. (2024) suggested that learners attributed these gains in part to the interactive and personalized nature of AI chatbot-mediated interaction, an affective dimension also reflected in the interview findings of the

present study, where students frequently identified the non-judgmental nature of AI conversation as a key factor supporting their speaking development. Tai and Chen (2024) similarly reported substantial speaking gains among EFL learners engaged in structured AI chatbot interaction, across both individual and paired conditions, with student interviews identifying reduced anxiety and increased engagement as the primary perceived benefits, themes that similarly emerge in the qualitative findings discussed in relation to Research Question 5.

Evidence from Qassrawi et al. (2024) within the Palestinian context illustrates that university-level Palestinian EFL students exhibited significant gains across all measured speaking skills following AI-assisted speaking activities: fluency, interaction, pronunciation, vocabulary, and spoken grammar. Qassrawi et al. (2024) further reported that, while the conditions supporting speaking development through AI-assisted methods exist among Palestinian EFL learners, these conditions are not currently being realized in practice through theoretically grounded instructional scaffolding, a gap that ASIM was designed to address through the structured, sequential engagement embedded in its second and third phases.

In the secondary school context, three weeks of supervised interaction with an AI chatbot produced statistically significant gains in speaking among Saudi secondary EFL learners (Alenezi & Alenezi, 2025). The parallels between this context and the Palestinian one - both are with low-resources and geographically isolated - further affirm that these two contexts provide very limited opportunities to practice authentic English speaking and hence are each conducive to developing speaking through AI-assisted methods. Alenezi and Alenezi (2025) also conclude that consistent and prolonged engagement with structured activities involving AI is critical to measuring speaking improvement, thus affirming that ASIM's multi-week structure is a necessary feature of effectiveness rather than an incidental one. Al-Obaydi et al. (2025) corroborated these results through a phenomenographic study of Iraqi and Czech EFL college students. Results of that study indicated that 75% of participants stated that AI interaction is necessary to develop their speaking skills, and the authors concluded that engaging in practice across multiple sessions provides the best opportunity to develop language through ChatGPT.

The theoretical foundation of ASIM's speaking gains may be explained by Canale and Swain's (1980) model of communicative competence. During Phase 2, students received opportunities for extended, low-stakes conversational practice across a range of contexts and registers, addressing sociolinguistic and discourse competence. Through Phase 3, students in the role of teacher had a specific opportunity to develop strategic competence, that is, the ability to manage communication, repair breakdowns, and use compensatory strategies independently.

Knowles' (1975) theory of self-directed learning further illuminates this process. Across Phases 2 and 3, responsibility for selecting speaking topics, evaluating performance, and

identifying gaps progressively shifts from teacher to student, creating the conditions for the initiative-taking and self-assessment that Knowles (1975) identifies as prerequisites for autonomous language development. In the Palestinian EFL context, where both Murad and Jalambo (2019) and the British Council and LearnJam (2025) document a near-complete absence of opportunities for authentic English speaking outside the classroom, the AI's role as a surrogate language environment takes on particular pedagogical importance; it provides a volume and variety of oral practice that neither the classroom structure nor the broader social environment is currently able to offer.

Discussion of the Results of the Fourth Question 4: Are there statistically significant differences between eleventh graders' mean scores on the post-test of English-speaking skills attributable to gender?

Results of the fourth research question showed statistically significant differences in post-test speaking scores between female and male students, with female students obtaining higher mean scores on both Task One (Responding to a Prompt) and Task Two (Interactive Conversation). This pattern parallels the gender difference observed in post-test writing scores (Research Question 2) and is subject to the same methodological caveat: because female and male participants were drawn from two separate single-sex schools, gender is confounded with school context in this study, and the observed difference cannot be attributed to gender independent of that context (see Section 5.3 and Section 1.8).

Within this constraint, the convergence of this finding with the writing results is noteworthy. That a similar pattern emerged independently across two distinct productive skills strengthens the case that this is a genuine feature of this sample rather than an artifact specific to one assessment; however, this convergence should be read as descriptive corroboration within this study, not as evidence generalizable beyond it. A comparable gender pattern was also reported by Jmaiel et al. (2025) in their study of ChatGPT-assisted writing instruction, suggesting the pattern observed here is not without precedent in the broader literature on AI-assisted language learning, though this single parallel is insufficient to establish a generalizable claim about gender differences in AI-assisted instruction.

Journal data offer a possible, interpretive explanation for this pattern. Cooperating teachers' journals recorded that female students more frequently engaged in structured goal-setting, identifying specific vocabulary, register features, or fluency targets to work on, and in verbal self-assessment at the end of sessions to evaluate progress and remaining gaps. This pattern of goal-setting and self-monitoring aligns with what Knowles (1975) identifies as a central behavior of the self-directed learner, and may plausibly support the cumulative, session-to-session development of oral proficiency that ASIM's structure is designed to foster. By contrast, journal entries described male students as demonstrating comparable or greater

creativity and spontaneity in AI interaction, including the construction of original prompts and topics, but as engaging less consistently in structured self-reflection and self-assessment.

This pattern should be treated as one plausible contributing factor drawn from qualitative observation, not as a confirmed mechanism, and is offered alongside, rather than as a resolution of, the school-context confound noted above.

Hamdan et al. (2025), in a qualitative study of Palestinian EFL university students, reported concern that ChatGPT use might enhance students' confidence and autonomy while simultaneously fostering over-reliance on the technology, a concern expressed particularly by students using it without structured pedagogical guidance. In the present study, both male and female students showed statistically significant improvement in speaking scores from pre-test to post-test despite the gender difference observed at post-test, suggesting that ASIM's structured design supported measurable development across differing engagement styles. The post-test difference between groups nonetheless suggests a possible area for refinement: ASIM's self-monitoring components may benefit from adaptations that better support students with a more spontaneous or exploratory engagement style.

Discussion of the Results of the Fifth Research Question: What Are Eleventh-Grade Students' Experiences and Perceptions of Engaging with the Three Phases of ASIM, Including Their Development of Self-Directed Learning Behaviors?

Quantitative Trajectory of Self-Directed Learning across Sessions

Through the structured observation checklist administered across six sessions of ASIM, there was a clear and statistically significant upward trend in self-directed learning (SDL) behaviors throughout the intervention. Each session's mean total scores showed gradual increases from Session 1 through Session 6, reflecting advancement across all three dimensions of interest, AI interaction skills, SDL skills, and language skill development. This trajectory, established at six separate points in time, provides evidence that SDL development under ASIM followed a developmental path rather than emerging through isolated, unconnected gains; each phase of the intervention appears to build upon skills established in the preceding phase.

The most rapid rate of SDL growth occurred between Sessions 3 and 5, coinciding with students' transition from Phase 1 preparation into Phase 2 active practice and the beginning of Phase 3. The timing of this growth period is theoretically significant, suggesting that observed gains in SDL corresponded with the point at which students began to assume greater cognitive responsibility for directing their own AI interactions. Li et al. (2024) provide a relevant theoretical framework for this developmental trajectory, describing iterative self-planning and adaptive goal-setting as constructs of SDL specific to generative-AI-mediated learning, behaviors closely aligned with those captured in Part Two of the observation checklist. Li, Wang, and Bonk (2024) further demonstrated that self-monitoring is an essential element of

L2 learners' engagement with AI, attributing this in part to AI's flexibility and personalization, affordances intentionally embedded in ASIM's Phase 2 design through students' ability to select their own discussion topics based on their own diagnosed needs.

The progressive nature of these gains is also consistent with the Cognitive Apprenticeship framework (Collins, Brown, & Newman, 1989), which posits that the rate of learning increases as scaffolding is progressively withdrawn. The largest gains in language skill development were observed in Sessions 5 and 6, coinciding with students' use of the AI Confusion Strategy in Phase 3, the point in the intervention at which teacher and AI scaffolding was most reduced. In the researcher's assessment, this pattern was more gradual, and less linear across individual students, than the theoretical framework alone would predict; the framework describes scaffolded withdrawal as a relatively orderly process, whereas the session-by-session data and observation notes suggest a more uneven trajectory shaped by individual proficiency, confidence, and the disrupted school calendar described in Section 3.6. This underscores a practical distinction between how scaffolded autonomy is theorized and how it unfolds in a politically constrained classroom setting.

Alfares (2025) offers additional support for this interpretation, finding that gains in autonomous learning behavior were attributable not to unstructured freedom of choice, but to scaffolding of specific autonomy behaviors, that is, structured choice, embedded within a theoretically grounded instructional framework. This finding responds directly to concerns raised by Hamamra et al. (2024), Alhur et al. (2025), and Hamdan et al. (2025) regarding unstructured AI use fostering dependency and disengagement rather than authentic SDL development in Palestinian educational contexts.

The qualitative observation notes discussed in Section 4.7.2 corroborate this quantitative trajectory: the documented shift from passive to active AI engagement parallels the upward trend in Part One scores, while the increasing specificity of students' self-diagnosis reports across sessions reflects the SDL gains documented in Part Two. The researcher's Week 5 journal entry, noting that students appeared to experience the model as coherent rather than merely procedural, may indicate a developmental threshold, observed in both schools between Weeks 4 and 6, at which students had sufficiently internalized ASIM's underlying logic to engage with it independently. This threshold is noted here as a pattern specific to this sample rather than a generalizable claim about intervention duration, a question that would require replication across a larger and more varied sample to address.

Students' Experiences and Perceptions of ASIM's Three Phases

Qualitative data from semi-structured interviews with six purposively selected participants indicated that, despite variation in engagement patterns and perceived gains associated with proficiency level and gender, all six participants described ASIM as fundamentally different from, and more cognitively demanding than, traditional EFL instruction.

The most immediate and recurrent impression was that ASIM required a new kind of learning. This aligns with Ouyang and Jiao's (2021) framework describing "AI-directed learning," in which the learner holds cognitive authority over the technology, as a configuration that remains underrepresented in empirical studies of AI in education. Students' accounts of Phase 3, in which they took on the role of teacher and directed the AI as learner, reflect the mechanisms described in Fiorella and Mayer's (2015) Learning by Teaching model, including knowledge reorganization, gap identification, and the exercise of evaluative authority. One advanced-level female student's account of teaching the AI producing a deeper understanding than any other school task closely illustrates this theoretical mechanism.

All six interviewed students identified the AI Confusion Strategy as the most distinctive and memorable feature of ASIM, describing it as shifting their perception of grammar from a set of memorized rules toward a principle-based system. This finding is consistent with Chi et al.'s (1989) work showing that self-explanation produces deeper and more transferable understanding than passive observation of correct models. Students' progression from identifying surface-level errors in early sessions to more sophisticated metalinguistic explanation by Phase 3 is consistent with the growth in metacognitive monitoring described by Flavell (1979). Hamdan et al. (2025) similarly found that Palestinian EFL students viewed ChatGPT as supporting independent learning and self-efficacy, while cautioning that this benefit depended on pedagogical structure; the AI Confusion Strategy addresses this concern directly by requiring students to critically evaluate AI output before accepting it. From the researcher's perspective, the consistency of this particular theme across all six interviews, more consistent than any other theme in the data, suggests that the AI Confusion Strategy may be doing more pedagogical work within ASIM than its position as a single component of Phase 3 might initially suggest.

Students most frequently described improvement in speaking ability in affective rather than purely linguistic terms, reporting reduced anxiety, increased confidence, and greater willingness to attempt spoken English. This is consistent with Horwitz, Horwitz, and Cope's (1986) finding that reducing evaluative threat is central to encouraging student speech, and with Fathi et al. (2024) and Tai and Chen (2024), who reported reduced speaking anxiety among students interacting with AI chatbots. In the Palestinian context specifically, Qaddumi et al. (2023) describe speaking anxiety as shaped by large class sizes, high-stakes examinations, and limited authentic speaking opportunities, suggesting that the affective gains associated with ASIM may be as pedagogically significant as the linguistic gains themselves.

Interview data also surfaced technical and linguistic challenges. All six students described difficulty formulating pedagogically effective prompts, a difficulty more pronounced among lower-proficiency students. This is consistent with Khlaif and Salha (2022), who found that access to educational technology alone does not ensure effective use, and with Qaddumi et al. (2021), who found Palestinian students rated the instructional influence of educational technology as moderate. This suggests that, although prompt-engineering training was

incorporated into ASIM's Phase 1, additional individualized support may be needed for lower-proficiency learners.

Development of Self-Directed Learning

Interview data indicated a consistent progression from teacher-directed dependence toward self-directed engagement, moving through initial disorientation, growing confidence in AI interaction, and eventual internalization of ASIM as a self-sustaining learning framework. This qualitative pattern corresponds with the quantitative increase in SDL scores documented above.

Students' descriptions of this progression, from waiting for teacher direction in early sessions to independently planning and executing Phase 3 tasks, align closely with Knowles' (1975) account of the self-directed learner as one who proactively defines learning needs, sets goals, identifies resources, and applies learning strategies. One student's description of becoming "his own teacher and his own student" reflects the full cycle described by Candy (1991), in which the learner simultaneously acts as designer, practitioner, and evaluator of their own learning, a cycle mirrored in the structure of ASIM's Phase 3.

Li, Bonk, Wang, and Kou (2024) describe iterative self-planning and adaptive goal-setting as constructs specific to SDL in the generative-AI era, operating at both a local level (individual learners' personalized interaction with AI) and a global level (learners' evolving understanding of AI's social and technological context). At the local level, students described adjusting prompts to address self-diagnosed gaps, consistent with adaptive self-planning. At the global level, several students demonstrated critical AI literacy by recognizing the technology's limitations, including formulaic responses and the absence of session memory.

Several students explicitly attributed their SDL development to ASIM's structured three-phase design rather than to AI access alone. This is consistent with Hamamra et al. (2024), Alhur et al. (2025), and Hamdan et al. (2025), who describe unstructured AI use as fostering dependency and cognitive offloading rather than deep engagement. One student's comment that "the structure helped me learn how to learn alone" illustrates how ASIM appears to have balanced the provision of structure with the cultivation of independence.

Journal entries further indicated gender-differentiated pathways within this developmental pattern, discussed in relation to Research Question 2 and Research Question 4: female students more consistently engaged in structured goal-setting and detailed self-assessment, while male students demonstrated greater spontaneity and creative prompt engineering. Both pathways were associated with meaningful SDL growth. This is consistent with Zimmerman's (2002) model of self-regulated learning as a cycle of forethought, performance, and reflection that may take different forms depending on learners' prior experience, culture, and cognitive style. In the researcher's view, this divergence was not anticipated as clearly at the outset as it appears in hindsight; the original design of ASIM's self-monitoring components was not

differentiated by engagement style, and the gender pattern observed here suggests this may be a design limitation rather than a finding to note only in passing. This is addressed further in the recommendations.

Discussion of the Results of the Sixth Research Question: How Do the Researcher and Cooperating Teachers Perceive the Implementation of ASIM, in Terms of Students' Engagement Patterns and the Contextual Challenges Encountered during the Intervention?

Findings from the reflective journals addressing this question provide an insider perspective on ASIM's implementation, complementing the student perspectives discussed above. Three aspects of the implementation experience recurred across the journals of the researcher and the two cooperating teachers: the transition from teacher-directed to self-directed learning, engagement with the AI Confusion Strategy, and gender-differentiated patterns of student initiative.

The most consistently documented implementation challenge was students' initial difficulty adjusting from teacher-directed to self-directed participation, a pattern the researcher attributed in part to the broader instructional culture described by the British Council and LearnJam (2025), in which Palestinian EFL teachers rely predominantly on lecture-based instruction and serve as the primary source of correction and task direction. Alhur et al. (2025) similarly identified low academic self-belief, reported by 60% of their participants, as a contributor to AI dependence; a related pattern may explain the initial dependency and disorientation documented across both journals in the early weeks of the intervention. Journal entries indicate a transition point between Weeks 3 and 5, after which students increasingly self-initiated tasks and directed their own AI engagement without prompting, suggesting that while this cultural pattern is real, it appears responsive to ASIM's phased structure and progressive scaffolding. In reflecting on this period, the researcher notes that the duration of this initial adjustment, roughly one-quarter of the total intervention, was longer than had been anticipated when designing the intervention timeline, a consideration with practical implications for how much preparatory time future implementations of ASIM should allocate before expecting independent engagement.

The gender-differentiated engagement patterns documented in the journals, with female students showing greater strength in reflective and metacognitive dimensions and male students showing greater strength in creative, exploratory AI interaction, carry implications for future ASIM design. Rather than treating this difference as a deficit to correct, it may represent a complementary strength that structured peer-learning activities could leverage, consistent with Vygotsky's (1978) account of peer interaction as an effective scaffolding mechanism between learners at adjacent proficiency levels.

The contextual challenges documented across all three journals, the three-day school week, extended remote learning due to access restrictions, and the use of community-hall sessions, illustrate the practical realities of conducting educational research within a politically and socially unstable environment, consistent with the British Council and LearnJam's (2025) account of how political conditions shape Palestinian education more broadly. Adaptations developed in response to these constraints, including a standardized re-engagement routine to address ChatGPT's lack of session memory, and the unexpectedly high level of self-directed engagement observed during remote and community-hall sessions, may be transferable to other resource-constrained EFL contexts facing similar disruption. The researcher acknowledges that these adaptations were reactive rather than planned in advance, and that a more deliberate contingency design, prepared before rather than during implementation, would likely have reduced some of the disruption documented in the journals.

Omar, Shaqour, and Khlaif (2024) found that Palestinian university faculty reported moderate confidence in integrating AI technology and recommended structured institutional training to build confidence further. This is consistent with both cooperating teachers' accounts of growing more comfortable in a facilitator role over the course of the intervention, a shift that parallels students' own development of self-directed learning behavior. This reinforces the recommendation, discussed further in Section 5.6, that any expansion of ASIM beyond the two participating schools be accompanied by structured professional development for participating teachers, addressing both the technical use of AI and its pedagogical facilitation.

5.2 Convergences of evidence: synthesis across quantitative and qualitative findings

The research questions in this study were designed to be complementary, such that quantitative findings could be interpreted with reference to corresponding qualitative data, and qualitative themes could be examined against quantitative patterns. The synthesis presented below draws these strands together; it should be read as an integrated interpretation of the study's own findings rather than as evidence generalizable beyond this sample, given the qualitative data's basis in a subsample of six students.

The first area of convergence lies between the statistically significant gains in writing and speaking scores and the qualitative accounts of the learning process. Pre-test/post-test gains alone are open to alternative explanations, such as rote practice, familiarity with the assessment format, or motivation unrelated to genuine skill development. Interview data speak to these alternative explanations: students consistently described ASIM as demanding greater cognitive effort than conventional instruction, and attributed their language gains to specific mechanisms, namely the iterative writing-revision process in Phase 2, the metacognitive demands of the AI Confusion Strategy in Phase 3, and the developing habit of self-monitoring.

These are the same mechanisms the theoretical framework in Chapter Two predicts would support durable language development, lending qualitative support to the interpretation that the quantitative gains reflect genuine skill development rather than test familiarity alone; this remains an inference drawn from a six-student subsample, however, rather than a claim established for the full sample of fifty participants.

The second convergence lies between the progressive upward trajectory of self-directed learning documented in the observation checklist and the qualitative accounts of developing autonomy. The observation data alone establish that SDL behaviors increased significantly and consistently across six sessions; the qualitative data offer a plausible account of how this occurred. Students' descriptions of moving from waiting for teacher direction to independently structuring Phase 3 sessions provide an experiential narrative consistent with the developmental pattern in the observation scores. Cooperating teachers' documentation of a transition point between Weeks 3 and 5, during which students began initiating tasks without prompting, corresponds with the period of most rapid gain recorded on the checklist, suggesting that both data sources are tracking a related underlying process through different methods.

The third convergence, and arguably the one with the most direct practical relevance, concerns the gender difference in post-test scores and the qualitative documentation of gender-differentiated engagement patterns. In isolation, the finding that female participants outperformed male participants on both productive-skill assessments is descriptive and confounded with school context. Read alongside the journal data, however, a plausible contributing explanation becomes available: female participants were more frequently observed setting specific goals, engaging in detailed self-assessment, and tracking their own progress. Because journal data also indicated that male participants demonstrated comparable or greater creativity and spontaneity in their AI interactions, this pattern is treated here as a difference in engagement style with implications for ASIM's design, rather than as evidence of unequal capacity to benefit from the intervention.

The fourth convergence concerns the relationship between the context in which the intervention was implemented and its outcomes. Test scores alone do not capture how the researcher and cooperating teachers adapted to disruptions, including a standardized re-engagement routine developed to address ChatGPT's lack of session memory, or how students themselves responded to interruptions in continuity. The unexpectedly high level of engagement observed by both cooperating teachers during remote and community-hall sessions, attributed to reduced institutional pressure and the novelty of the setting, suggests that contextual factors outside the traditional classroom may have supported self-directed engagement in ways this study was not designed to formally measure or compare against a traditional-classroom baseline. This pattern emerges only through the combination of journal and observation data and would not be visible from either source alone.

Taken together, these four areas of convergence suggest a conclusion that extends beyond any single research question: within this study, ASIM's association with positive outcomes does not appear attributable to any single component in isolation, such as AI access, the feedback loop, or the AI Confusion Strategy, but to the coherence of the three-phase structure as a whole, through which cognitive responsibility was progressively transferred from teacher to student. This interpretation is consistent with the theoretical framework presented in Chapter Two and distinguishes ASIM's structured approach from the unstructured AI use that the Palestinian literature associates with dependency rather than genuine skill development (Hamamra et al., 2024; Alhur et al., 2025; Hamdan et al., 2025). This synthesis reflects the patterns observed within this study's specific sample and context and is offered as a basis for the recommendations in Section 5.6, rather than as a claim of effectiveness established beyond this study.

Figure 5.1 presents a visual summary of this convergence, illustrating how the four quantitative outcomes, writing gains, speaking gains, the upward SDL trajectory, and the gender difference, relate to the qualitative themes of agency, the AI Confusion Strategy as a metacognitive turning point, growing confidence, and gradual autonomy, as different reflections of ASIM's three-phase structure rather than as independent findings.

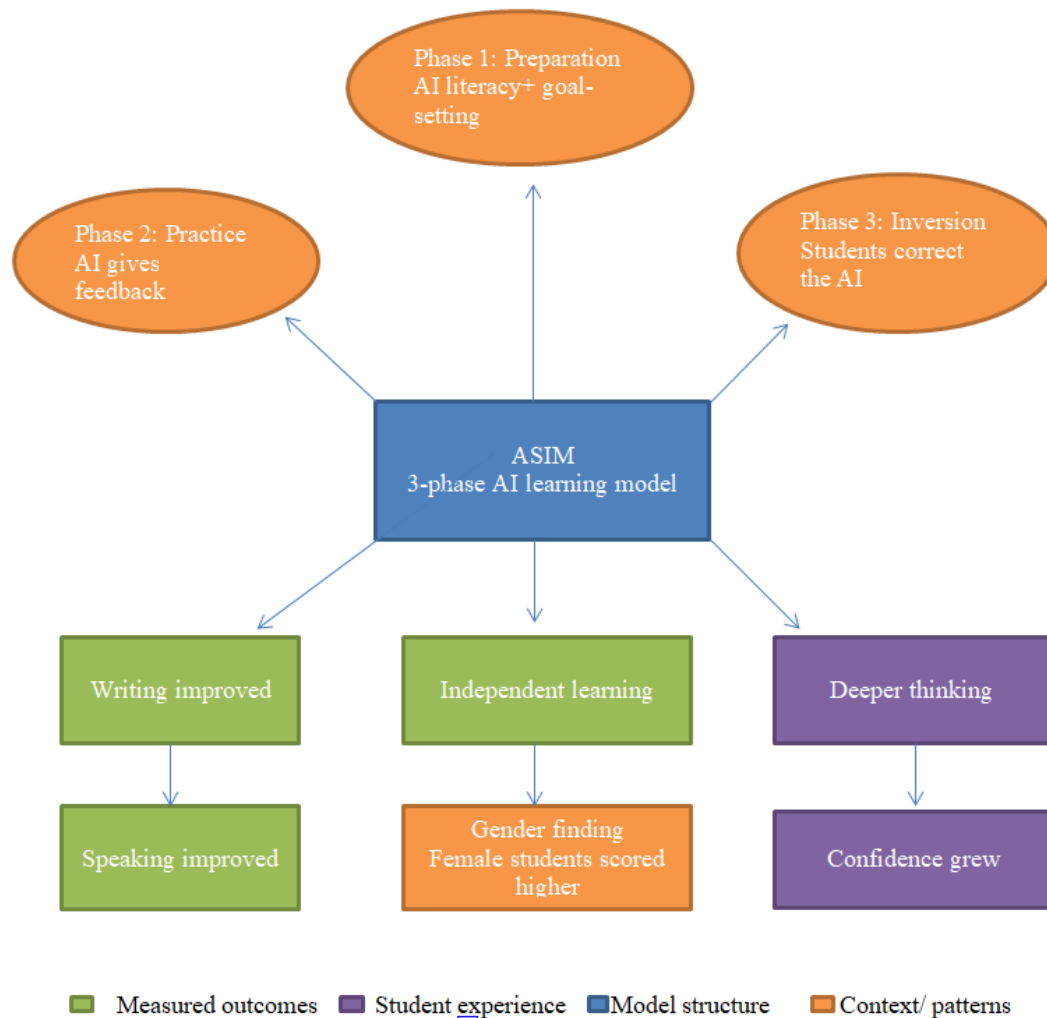


Figure 5.1: Integrated Synthesis of Quantitative and Qualitative Findings across ASIM's Three Phases

5.3 Recommendations

Based on the findings of this study, the following recommendations are proposed

For EFL Teachers

- Teachers in large classes where individualized feedback is difficult to provide may consider piloting ASIM's three-phase structure as a supplementary framework for developing students' academic writing and speaking skills.
- Teachers should provide short pre-activity training sessions in prompt engineering before any AI-assisted writing or speaking task. Students unfamiliar with generative AI

may benefit from this preparation to engage more productively with AI-assisted activities.

- The AI Confusion Strategy can be integrated into individual EFL lessons as a standalone metalinguistic activity, independent of ASIM's full three-phase structure, given its potential to strengthen grammatical awareness. Teachers adopting this strategy on its own should continue to provide guidance during its use, so that students apply it as an analytical task rather than misinterpreting AI-generated errors as reliable models.
- Students transitioning to self-directed, AI-assisted learning are likely to require an adjustment period before engaging independently; in this study, most students required approximately three to five weeks before consistently self-initiating tasks (Section 5.2). Teachers should plan explicit scaffolding activities for this initial period rather than expecting immediate independent engagement, while recognizing that the length of this adjustment period may vary across other settings and has not been established as a fixed timeframe.

For Curriculum Developers:

- Curriculum developers for secondary EFL learners are encouraged to include prompt engineering and AI literacy as explicit skills within the curriculum, given their role in enabling productive, AI-assisted language learning.
- Writing and speaking tasks in the eleventh-grade EFL curriculum could be redesigned to include iterative cycles of feedback, allowing students to revise and improve their work through AI-assisted feedback rather than relying on single-draft submission. This approach may be particularly valuable in contexts such as Palestinian public schools, where large class sizes limit the frequency of individualized teacher feedback (Section 1.2).
- Given that ASIM's design is grounded in positioning students to direct and evaluate AI output rather than passively receive it, curriculum developers designing AI-enhanced instructional materials for secondary education are encouraged to build this principle into task design, for example, by requiring students to justify, question, or correct AI-generated responses rather than accept them without evaluation.

For Future Researchers:

- Future research should examine whether ASIM can be effectively implemented at other educational levels, such as lower secondary grades or higher education, to determine the range of educational levels and proficiency levels across which the model's three-phase structure remains effective, rather than assuming applicability across all ages and contexts without further testing.
- Longitudinal research is needed to determine whether the writing and speaking gains observed in this study are maintained after the intervention ends, and whether they transfer to authentic real-world contexts beyond the classroom. To provide stronger

causal evidence of ASIM's effectiveness relative to traditional instruction, researchers should replicate this study using a control group in settings where this is logistically feasible.

- Because the AI Confusion Strategy consistently emerged across the quantitative, observational, and interview data as the most impactful cognitive element of the ASIM intervention, it warrants dedicated empirical investigation as an independent pedagogical tool, separate from its role within ASIM's broader three-phase structure.

References

- Abu Alyan, A. (2013). Oral communication problems encountering English major students: Perspectives of learners and teachers in Palestinian EFL university context. *Arab World English Journal*, 4(3), 226–248.
- Abu Hashish, M., & Al-Halimi, R. (2023). The effect of digital transformation on teacher performance evaluation in Palestinian schools. *Frontiers in Education*.
- Ahmed, A. M., Troudi, S., & Riley, S. (Eds.). (2020). *Feedback in L2 English writing in the Arab world: Inside the black box*. Palgrave Macmillan.
- Al-Atrash, A. N. F. (2024). *Exploring EFL students' perspectives of the role of ChatGPT-aided personalized learning in enhancing autonomous learning in Palestinian higher education [Master's thesis, An-Najah National University]*.
- Al-Mahrooqi, R., & Denman, C. J. (2020). Assessing students' critical thinking skills in the humanities and sciences colleges of a Middle Eastern university. *International Journal of Instruction*, 13(1), 783–796.
- Al-Obaydi, L. H., Pikhart, M., & Hossain, M. K. (2025). ChatGPT and the development of core language skills: An exploratory study of EFL college students. *Contemporary Educational Technology*, 17(4), Article ep591.
- Al-Ureimi, K. K. A. (2021). The difficulties that students face in learning English language skills: A descriptive survey study in the light of the results of research and studies. *Arab Journal for Humanities and Social Studies*, 9(3), 319–362.
- Al-Yahyaie, R. Z., Muhamad, M. M., & Alkharusi, H. A. (2022). Barriers to teaching higher order thinking skills to EFL school learners: A systematic review. *International Journal of Academic Research in Progressive Education and Development*, 11(2), 784–801.
- Alenezi, A., & Alenezi, A. (2025). Evaluating the effectiveness of chatbot-assisted learning in enhancing English conversational skills among secondary school students. *Education Sciences*, 15(9), Article 1136.
- Alfares, N. (2025). Effectiveness of a scaffolded autonomous learning scheme in enhancing learners' motivation in EFL classes. *Learning and Motivation*, 89, 102084.
- Alhur, A. A., Khlaif, Z. N., Hamamra, B., & Hussein, E. (2025). Paradox of AI in higher education: Qualitative inquiry into AI dependency among educators in Palestine. *JMIR Medical Education*, 11, Article e74947.
- Ali, F., Choy, D., Divaharan, S., Tay, H. Y., & Chen, W. (2023). Supporting self-directed learning and self-assessment using TeacherGAIA, a generative AI chatbot application: Learning approaches and prompt engineering. *Learning: Research and Practice*, 9(2), 135–147.
- AlTwijri, L., & Alghizzi, T. M. (2024). Investigating the integration of artificial intelligence in English as a foreign language classes for enhancing learners' affective factors: A systematic review. *Heliyon*, 10(10), e31053.
- Anderson, L. W., & Block, J. H. (1975). *Mastery learning in classroom instruction*. Macmillan.

- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Anderson, N. J. (2002). *The role of metacognition in second language teaching and learning*. ERIC Digest. Washington, DC: ERIC Clearinghouse on Languages and Linguistics.
- Aqel, F. (2002). English language teachers' implementation of effective teaching principles in Nablus schools. *An-Najah University Journal for Research (Humanities)*, 16(2), 441–468.
- Badah, A. M., Mafarjah, F. M. N., Abuzaid, A. M. A., & Shehada, S. M. A. (2024). Behind the screens: Evaluating the effectiveness of Zoom in EFL learning and teaching in higher education. *Electronic Journal of e-Learning*, 22(9), 163–175.
- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the dangers of stochastic parrots: Can language models be too big?* *FACCT '21*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Benson, P. (2001). *Teaching and researching autonomy in language learning*. Longman.
- Benson, P. (2013). *Teaching and researching autonomy in language learning (2nd ed.)*. Pearson Education.
- Benware, C. A., & Deci, E. L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal*, 21(4), 755–765. <https://doi.org/10.3102/00028312021004755>
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Lawrence Erlbaum Associates.
- Bergmann, J., & Sams, A. (2012). Flip your classroom: Reach every student in every class every day. *International Society for Technology in Education*.
- Biesta, G. J. J. (2009). *Good education in an age of measurement: On the need to reconnect with the question of purpose in education*. Educational Assessment, Evaluation and Accountability, 21(1).
- Biesta, G. J. J. (2022). *World-centred education: A view for the present*. Routledge.
- Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university (4th ed.)*. Open University Press.
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). MIT Press.
- Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment*, 1(2), 1–12.
- Bommasani, R., Hudson, D. A., Adelli, E., et al. (2021). *On the opportunities and risks of foundation models*. arXiv preprint arXiv:2108.07258.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. National Academy Press.
- British Council & LearnJam. (2025). An evaluation of primary and secondary English teaching and learning in Palestine. *British Council Palestine*.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.
- Candy, P. C. (1991). *Self-direction for lifelong learning*. Jossey-Bass.

- Carr, W., & Kemmis, S. (1986). *Becoming critical: Education, knowledge and action research*. Falmer Press.
- Celce-Murcia, M. (2008). Rethinking the role of communicative competence. In E. Alcón Soler & M. P. Safont Jordà (Eds.), *Intercultural language use and language learning* (pp. 41–57). Springer.
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145–182.
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education (8th ed.)*. Routledge.
- Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Lawrence Erlbaum Associates.
- Connor, U. (2011). *Intercultural rhetoric in the writing classroom*. University of Michigan Press.
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239.
- Creemers, B. P. M., & Kyriakides, L. (2015). Educational effectiveness. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences (2nd ed., Vol. 7, pp. 149–155)*. Elsevier.
- Cumming, A. (2016). Theoretical orientations to L2 writing. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 65–88). De Gruyter Mouton.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222–251.
- Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E. D., & Marzuki. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, 11(1), Article 2290342.
- Dewey, J. (1938). *Experience and education*. Macmillan.
- Duran, D. (2017). Learning-by-teaching: Evidence and implications as a pedagogical mechanism. *Innovations in Education and Teaching International*, 54(5), 476–484.
- Duran, D., & Topping, K. J. (2017). *Learning by teaching: Evidence-based strategies to enhance learning in the classroom*. Routledge.
- Ellis, R. (1997). *Second language acquisition*. Oxford University Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, 121, Article 103254.
- Ferris, D. R. (2011). *Treatment of error in second language student writing (2nd ed.)*. University of Michigan Press.
- Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, 38(4), 281–288.
- Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity: Eight learning strategies that promote understanding*. Cambridge University Press.

- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911.
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387.
- Freire, P. (1970). *Pedagogy of the oppressed*. Continuum.
- Gay, L. R., Mills, G. E., & Airasian, P. (2012). *Educational research: Competencies for analysis and applications* (10th ed.). Pearson.
- Guglielmino, L. M. (2013). The case for promoting self-directed learning in formal educational institutions. *SA-eDUC Journal*, 10(2), 1–18.
- Guo, K., & Wang, D. (2023). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, 29(7), 8435–8463.
- Guo, K., & Wang, D. (2024). Exploring the impact of AI chatbots on L2 students' writing performance and their perceptions. *Computer Assisted Language Learning*.
- Hamamra, B., Khlaif, Z. N., & Mayaleh, A. (2025). Between promise and precarity: Palestinian university educators' experiences with AI integration in higher education. *Interactive Learning Environments*.
- Hamamra, B., Khlaif, Z. N., Mayaleh, A., & Abu Baker, A. (2025). A phenomenological examination of educators' experiences with AI integration in Palestinian higher education. *Cogent Education*, 12(1).
- Hamamra, B., Mayaleh, A., & Khlaif, Z. N. (2024). Between tech and text: The use of generative AI in Palestinian universities — A ChatGPT case study. *Cogent Education*, 11(1), Article 2380622.
- Hamdan, M., Al-Atrash, A., & Dweikat, K. (2025). Qualitative insights into EFL students' use of ChatGPT-assisted autonomous learning for enhancing self-directed learning in Palestinian higher education. *Discover Global Society*.
- Hammad, E. A. (2016a). Palestinian university students' problems with EFL essay writing in an instructional setting. In A. Ahmed & H. Abouabdelkader (Eds.), *Teaching EFL writing in the 21st century Arab world* (pp. 91–118). Palgrave Macmillan.
- Hammad, E. A. (2016b). Palestinian EFL students' writing errors: A case study. *Journal of Language Teaching and Research*, 7(3).
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Holec, H. (1981). *Autonomy and foreign language learning*. Pergamon Press.
- Holmes, W., & Porayska-Pomsta, K. (2023). *The ethics of artificial intelligence in education: Practices, challenges, and debates*. Routledge.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132.
- Hutchins, E. (1995). *Cognition in the wild*. MIT Press.
- Hwang, G. J., & Chang, C. Y. (2021). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*.
- Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press.

- Ji, Z., Lee, N., Frieske, R., et al. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1–38.
- Jmaiel, H. A., Abukhait, R. O., Mohamed, A. M., & Shaaban, T. S. (2025). The role of ChatGPT in enhancing EFL students' ESP writing skills: An experimental study of gender and major differences. *Discover Education*, 4, Article 240.
- Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Interaction Book Company.
- Joyce, B., Weil, M., & Calhoun, E. (2015). *Models of teaching (9th ed.)*. Pearson.
- Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal*, 99(1), 1–18.
- Kaplan, R. B. (1966). *Cultural thought patterns in intercultural education*. Language Learning, 16(1–2), 1–20.
- Khlaif, Z. N., & Salha, S. (2022). Exploring the factors influencing mobile technology integration in higher education. *Technology, Pedagogy and Education*.
- Knowles, M. S. (1975). *Self-directed learning: A guide for learners and teachers*. Association Press.
- Kohnke, L., & Moorhouse, B. L. (2023). Facilitating synchronous online language learning with Zoom. *Knowledge Management & E-Learning*, 15(2).
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Lai, C., & Zheng, D. (2018). Self-directed use of mobile devices for language learning beyond the classroom. *ReCALL*, 30(3), 299–318.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- Lee, S., Choe, H., Zou, D., & Jeon, J. (2025). Generative AI (GenAI) in the language classroom: A systematic review. *Interactive Learning Environments*.
- Li, B., Bonk, C. J., Wang, C., & Kou, X. (2024). Reconceptualizing self-directed learning in the era of generative AI: An exploratory analysis of language learning. *IEEE Transactions on Learning Technologies*, 17, 1515–1529.
- Li, B., Tan, Y. L., Wang, C., & Lowell, V. (2025). Two years of innovation: A systematic review of empirical generative AI research in language learning and teaching. *Computers and Education: Artificial Intelligence*, 9, Article 100445.
- Li, Z., Wang, C., & Bonk, C. J. (2024). Exploring the utility of ChatGPT for self-directed online language learning. *Online Learning*, 28(3), 157–180.
- Liang, W., & Wu, Y. (2024). Exploring the use of ChatGPT to foster EFL learners' critical thinking skills from a post-humanist perspective. *Thinking Skills and Creativity*, 54, Article 101645.
- Little, D. (1991). Learner autonomy 1: Definitions, issues and problems. *Authentik*.
- Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S.-Y. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments*, 11, Article 50.
- Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching*, 51(3), 285–329.
- Lombardi, M. M. (2007). Authentic learning for the 21st century: An overview. *Educause Learning Initiative*, 1(2007), 1–12.

- Long, D., & Magerko, B. (2020). *What is AI literacy? Competencies and design considerations*. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–16.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.
- Lyu, B., et al. (2025). Effectiveness of chatbots in improving language learning: A meta-analysis of comparative studies. *International Journal of Applied Linguistics*.
- Magerko, B., & Long, D. (2020). *What is AI literacy? Competencies and design considerations*. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–16.
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11, Article 9.
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054.
- Murad, A. I., & Jalambo, M. O. (2019). EFL students' reluctance in participating in English speaking activities at University College of Applied Sciences: Challenges and solutions. *Open Journal of Modern Linguistics*, 9(1), 1–17.
- Murase, F. (2015). Measuring language learner autonomy: Problems and possibilities. In C. J. Everhard & L. Murphy (Eds.), *Assessment and autonomy in language learning*. Palgrave Macmillan.
- Nation, I. S. P., & Newton, J. (2021). *Teaching ESL/EFL listening and speaking (2nd ed.)*. Routledge.
- Navigli, R., Conia, S., & Ross, B. (2023). Biases in large language models: Origins, inventory, and discussion. *ACM Journal of Data and Information Quality*, 15(2), 1–24.
- Nestojko, J. F., Bui, D. C., Kornell, N., & Bjork, E. L. (2014). Expecting to teach enhances learning and organization of knowledge in free recall of text passages. *Memory & Cognition*, 42(7), 1038–1048.
- Nguyen, T. M. T., & Nguyen, T. T. L. (2017). Influence of explicit higher-order thinking skills instruction on students' learning of linguistics. *Thinking Skills and Creativity*, 26, 113–127.
- Omar, A., Shaqour, A. Z., & Khlaif, Z. N. (2024). Attitudes of faculty members in Palestinian universities toward employing artificial intelligence applications in higher education: Opportunities and challenges. *Frontiers in Education*, 9.
- Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, 2, Article 100020.
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 422.
- Paul, R., & Elder, L. (2007). Critical thinking: The art of Socratic questioning. *Journal of Developmental Education*, 31(1), 36–37.
- Piaget, J. (1952). *The origins of intelligence in children*. International Universities Press.
- Piaget, J. (1954). *The construction of reality in the child*. Basic Books.
- Piaget, J. (1970). *Science of education and the psychology of the child (D. Coltman, Trans.)*. Orion Press. (Original work published 1969)
- Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. Viking Press.

- Poláková, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1), Article 2410101.
- Qaddumi, H. A., Smith, M., Alawneh, Y., Shawamreh, N., Bakeer, A., & Itemizeh, M. (2023). Palestinian undergraduate learners' foreign language classroom anxiety in online environments. *Studies in English Language and Education*, 10(3), 1358–1378.
- Qaddumi, H., Bartram, B., & Qashmar, A. L. (2021). Evaluating the impact of ICT on teaching and learning: A study of Palestinian students' and teachers' perceptions. *Education and Information Technologies*, 26(2), 1865–1876.
- Qaddumi, H., Smith, M., Masd, K., Bakeer, A., & Abu-Ulbeh, W. (2023). Investigating Palestinian in-service teachers' beliefs about the integration of information and communication technology (ICT) into teaching English. *Education and Information Technologies*, 28, 12785–12805.
- Qassrawi, R. M., ElMashharawi, A., Itmeizeh, M., & Tamimi, M. H. M. (2024). AI-powered applications for improving EFL students' speaking proficiency in higher education. *Forum for Linguistic Studies*, 6(5), 535–549.
- Reynolds, L., & McDonell, K. (2021). *Prompt programming for large language models: Beyond the few-shot paradigm*. CHI EA '21, Article 314.
- Richards, J. C. (2008). *Teaching listening and speaking: From theory to practice*. Cambridge University Press.
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching (3rd ed.)*. Cambridge University Press.
- Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249–255.
- Roehr, K. (2007). Metalinguistic knowledge and language ability in university-level L2 learners. *Applied Linguistics*, 29(2), 173–199. <https://doi.org/10.1093/applin/amm037>
- Roehr-Brackin, K. (2018). *Metalinguistic awareness and second language acquisition*. Routledge. <https://doi.org/10.4324/9781315661001>
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press.
- Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, 77(4), 534–574.
- Rosenberg, J. M., & Koehler, M. J. (2015). Context and technological pedagogical content knowledge (TPACK): A systematic review. *Journal of Research on Technology in Education*, 47(3), 186–210.
- Roy, S. S., & Gandhimathi, S. N. S. (2025). Self-directed learning for optimizing sustainable language learning via mobile assisted language learning: A systematic review. *Frontiers in Education*, 9, 1463721.
- Sapan, M., & Uzun, L. (2024). *The effect of ChatGPT-integrated English teaching on high school EFL learners' writing skills and vocabulary development*. *International Journal of Education in Mathematics, Science and Technology*, 12(6), 1679–1699.
- Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology*, 60, 101832.
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Shehadeh, A., & Dwaik, R. (2013). The practice of teaching English to young learners in the Palestinian context: A multidimensional analysis. *Arab World English Journal*, 4(3), 63–76.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, Article 1260843.
- Southworth, J., Migliaccio, K., Glover, J., et al. (2023). Developing a model for AI across the curriculum. *Computers and Education: Artificial Intelligence*, 4, 100127.
- Strobl, C., Ailhaud, E., Benetos, K., Devitt, A., Kruse, O., Proske, A., & Rapp, C. (2019). Digital support for academic writing: A review of technologies and pedagogies. *Computers & Education*, 131, 33–48.
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6(1), 31–40.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.
- Swain, M. (2006). Linguaging, agency and collaboration in advanced second language learning. In H. Byrnes (Ed.), *Advanced language learning: The contributions of Halliday and Vygotsky* (pp. 95–108). Continuum.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285.
- Szabó, F., & Szoke, J. (2024). *How does generative AI promote autonomy and inclusivity in language teaching?* *ELT Journal*, 78(4), 478–488. <https://doi.org/10.1093/elt/ccae052>
- Taguchi, N. (2015). Instructed pragmatics at a glance: Where instructional studies were, are, and should be going. *Language Teaching*, 48(1), 1–50.
- Tai, T.-Y., & Chen, H. H.-J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers & Education*, 220, Article 105112.
- Thaher, M. (2005). The effects of large class on EFL students at An-Najah National University. *An-Najah University Journal for Research (Humanities)*, 19(3), 1047–1096.
- Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
- Vattøy, K.-D., & Gamlem, S. M. (2023). Students' experiences of peer feedback practices as related to awareness raising of learning goals, self-monitoring, self-efficacy, anxiety, and enjoyment in teaching EFL and mathematics. *Scandinavian Journal of Educational Research*, 67(6), 904–918.
- Von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. The Falmer Press.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

- Warschauer, M., Tseng, W., Yim, S., et al. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing*, 62, 101071.
- Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, 14, Article 1261955.
- Wenden, A. L. (1998). Metacognitive knowledge and language learning. *Applied Linguistics*, 19(4), 515–537.
- Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Harvard University Press.
- Whitman, N. A. (1988). *Peer teaching: To teach is to learn twice*. ASHE-ERIC Higher Education Report No. 4. Association for the Study of Higher Education.
- Widdowson, H. G. (1978). *Teaching language as communication*. Oxford University Press.
- Wiley. (2024). *AI has hurt academic integrity in college courses but can also enhance learning, say instructors, students [Press release]*. Wiley Newsroom.
- Wilson, J., & Roscoe, R. D. (2020). Automated writing evaluation and feedback: Multiple metrics of efficacy. *Journal of Educational Computing Research*, 58(1), 87–125.
- Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.
- Xiao, F., Zhu, S., & Xin, W. (2025). Exploring the landscape of generative AI (ChatGPT)-powered writing instruction in English as a foreign language education: A scoping review. *SAGE Open*.
- Yaseen, K. N., & Dweikat, K. (2025). The role of using ChatGPT in improving the creative writing of English language students from the perspectives of instructors and postgraduate students. In M. Sanmugam, N. M. Mohd Barkhaya, S. Edwards, & Z. Khlaif (Eds.), *Fostering inclusive education with AI and emerging technologies* (pp. 301–348). IGI Global.
- Zhang, S., Shan, C., Lee, J. S. Y., Che, S., & Kim, J. H. (2023). Effect of chatbot-assisted language learning: A meta-analysis. *Education and Information Technologies*, 28(11), 15223–15243.
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70.
- Zimmerman, B. J., & Schunk, D. H. (2011). Self-regulated learning and performance: An introduction and an overview. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 1–12). Routledge.
- Zou, D., Xie, H., & Wang, F. L. (2023). Effects of technology-enhanced peer, teacher and self-feedback on students' collaborative writing, critical thinking tendency and engagement in learning. *Journal of Computing in Higher Education*, 35(1), 166–185.

- Özkan, E. K., Erdemir, N., & Coşkun, D. (2024). A systematic review of EFL teachers' perspectives on artificial intelligence technologies. *Ihlara Journal of Educational Research*, 9(3).

Appendices

Appendix 1

The AI-Student Inversion Model (ASIM)

A Researcher Created, Theoretical-Based Instructional Model For the use of GAI in EFL Education

Part I: Overview and Conceptual Background

The AI-Student Inversion Model (ASIM) is a researcher created and theoretical based instructional model to integrate Generative Artificial Intelligence (GAI) tools such as ChatGPT into a well-defined self-directed EFL learning cycle consisting of three phases: 1) AI assisted discovery; 2) Collaborative co-creation; and 3) Student as teacher demonstrating mastery. The ASIM repositions the role of all educational actors (the AI, teacher and student) in these three sequential phases, such that the typical education roles are systematically and purposely inverted.

The ASIM's conceptual genesis stems from the researcher's observation of a persistent and poorly addressed gap in the Palestinian EFL secondary education sector – while a growing number of AI systems are accessible for learners, current instructional frameworks generally consider students to be consumers of AI-generated materials, not agents who can be in charge of directing, assessing and critically engaging with how AI behaves. Therefore, ASIM was specifically designed to bridge this gap through the repositioning of learners as the cognitive authority within the human-AI relationship.

The model's name describes the central feature of the innovation, the inversion. In the conventional manner of instruction using AI, the flow of the relationship follows the following: AI + Teacher → Student. In these three phases of the ASIM model, this relationship is inverted; Phase 3 then has the Student → AI. This final inversion, the student teaching the AI, is the primary characteristic that sets ASIM apart from all existing AI-assisted models of second language acquisition; this is the most original contribution of ASIM as a pedagogical model.

In terms of scope of implementation, ASIM is applied selectively to lessons in the eleventh grade EFL curriculum that have productive language goals — that is, academic writing and speaking lessons. It is not applied to all the lessons in each unit, but rather to the two writing lessons and two speaking lessons contained in each unit. Over the course of 8 to 12 weeks of instruction, ASIM is implemented over three instructional units, which results in a total of

twelve ASIM cycles. Each cycle consists of four 40-minute sessions (one for each phase, except that Phase 2 takes two lessons), so each cycle requires a total of 160 minutes.

Part II: Theoretical Foundations

ASIM is grounded in six interconnected theoretical frameworks, each contributing a distinct dimension to the model's design. The following section presents each theory's philosophical orientation, educational objectives, core components, and the specific strategies that ASIM derives from it. A full elaboration of each framework is provided in Chapter Two of the doctoral thesis from which this model is drawn; the present section is intended for readers who encounter ASIM as a standalone document.

Constructivism (Piaget, 1954; Vygotsky, 1978)

Philosophy: Knowledge is developed rather than transmitted. Learners develop through interaction with their environment, not through passive reception of information. The foundation of ASIM's design and its three phases is influenced by Piaget's cognitive constructivism (which emphasizes individual mental construction through assimilation and accommodation) and Vygotsky's social constructivism (which emphasizes the role of cultural tools/mediation and social interaction in cognitive growth).

Educational Objectives: The goal is to create learners who build, test, and improve their own understanding, to create a context in which learning is meaningful and relevant, and to view the errors made in this process as opportunities for cognitive growth.

Constructivist strategies employed by ASIM:

- **Developing Active Knowledge:** All three phases require the production of knowledge rather than merely receiving knowledge from another person or scenario (both through AI-directed writing drafts, grammar explanations and through person-to-person teaching interactions).
- **Contextualized Learning:** All tasks are presented to students in relevant and meaningful real-life situations as opposed to simply presented in a pure academic context (decontextualized exercise).
- **Learning through Error:** Both the AI-associated revision cycle in Phase 2 and the AI Confusion Strategy in Phase 3 treat identifying and correcting errors as the primary and fundamental way to deepen an individual's understanding of a particular concept.
- **Reflective Abstraction (Piaget, 1977):** The self-diagnosis/assessment report provided to students in Phase 1 and the final reflection they complete in Phase 3 require students to reflect on their own cognitive processes and are therefore considered the most advanced level of constructivist engagement.

Self-Directed Learning (Knowles, 1975)

Philosophy: Adult learners are able to take full responsibility for their own learning (i.e., they are capable of diagnosing their needs, developing goals, identifying resources, and evaluating results) and they benefit from doing so. SDL is more than simply engaging in self-directed study; rather, it is an intentional and self-aware (metacognitive) decision to own your own learning pathway.

Educational Goals: The goals of SDL are to create learners who will be capable of independently planning, monitoring, and evaluating their own learning; facilitating an internal motivation and sense of agency; and developing the metacognitive strategies needed to support lifelong independent learning.

Strategies derived from SDL by ASIM:

- Learner-directed AI interaction: Students will initiate and configure their interactions with AI using prompt engineering throughout all three phases of ASIM; therefore, students will determine what questions they will ask, how to ask them, and how they will utilize the responses.
- Self-diagnosis and goal setting: The self-diagnosis report activity in Phase 1 will require students to evaluate their own understanding, identify any gaps, and develop questions to ask the teacher.
- Self-monitoring and self-assessment: The observation protocol and reflective journal will show students how to develop self-monitoring behaviors which serve as indicators of SDL.
- Autonomous pre-class preparation: Phase 1 is intended to occur asynchronously and individually in relation to the classroom; therefore, its structure and activity form a direct embodiment of the SDL principle that learning occurs beyond the time the teacher is present in the classroom.

Learning by Teaching Model (Duran, 2017; Whitman, 1988).

Philosophy: The act of teaching is considered one of the forms of learning at its highest level. When a variety of students become the teacher, they must take the knowledge that has been given to them, and create, combine, and evaluate their knowledge, at levels that they could not achieve if they passively received information.

Educational Objectives: To develop the most in-depth form of mastery of subject matter that a student can achieve; to develop pedagogical reasoning and metalinguistic awareness; and to assist a learner in increasing their levels of confidence through a successful experience with being the teacher.

Strategies derived from Learning by Teaching:

- Student-as-Teacher Mastery Phase 3: The entire third phase is built upon this premise, by having the students teach the computer grammar concepts; respond to its questions; correct intentional errors made by the computer; and evaluate the computer's performance.
- AI Confusion Strategy: When students instruct the AI to make intentional errors, they develop metalinguistic and analytical skills by identifying, explaining, and correcting those errors.
- Developing Pedagogical Reasoning: In Activity 1 of Phase 3 the students are required to select examples, create explanations, and modify their teaching style based on the questions asked by the AI.
- Peer-discussion in Phase 2: Activity 1 of Phase 2 requires that the students share what they discovered from their experience with the AI and then teach their peers what they have found. This is an example of "learning through teaching" in a peer-mediated environment.

Mastery Learning (Bloom, 1968; Anderson & Block, 1975):

Philosophy - Every student can reach high levels of achievement if given sufficient time, suitable instructional environments, and chances to revisit the content are available until they display true mastery. Mastery is the expected outcome of all learners, with Instructional Design supporting individual learners through appropriate materials.

Learning Objectives: All learners must demonstrate that they are prepared to continue to the next defined level of competency; must receive continued formative feedback to help identify and correct mistakes; and help build assurance through success.

Strategies derived from Mastery Learning

- Phase Progression as a Mastery Gate - Students will complete the self-diagnosis prior to advancing to Phase 2 and will complete the Phase 2 Writing and Speaking assignments prior to entering into Phase 3. Each phase of learning will result in a check for mastery as well as contribute to learning.
- Formative Feedback Loop - AI delivers immediate, individualized corrective feedback as students complete their activities, allowing them to identify errors and receive explanations prior to moving on.
- Individualized Pacing for Phase 1 - Through the asynchronous format and self-directed learning of Phase 1, each learner will engage with curricular material at an appropriate rate to develop a real understanding of the material.
- Teacher Use of Self-diagnosis Data - The teacher examines the students' Phase 1 Diagnostic Data before designing the Class Discussion in Phase 2 in order to gather data-driven, gap-focused instruction.

Cognitive Apprenticeship (Collins, Brown, & Newman, 1989)

Philosophy: The cognitive abilities of the expert can be maximized through the use of a defined sequence of observational experiences, support-practice and independent completion of tasks, similar to the traditional model of apprenticeship. The cognitive processes of the expert must be observable and transferrable to learners, rather than being implicit.

Educational Goals: To create a visible representation of the cognitive and language processes of the expert for learners; to provide a systematic means of allowing the learner to assume responsibility for the task gradually; and to train learners to be able to complete the tasks of an expert independently and to articulate his/her thought process.

Strategy Alignment and ASIM Phases

CA Mechanism	ASIM Phase	Who?	How?
Modeling	Phase 1	AI model expert language	AI shows vocabulary/grammar/text structure through dialogic interactivity
Coaching	Phase 2	AI + Teacher provide dual coaching	AI coaches student accuracy in language/ teacher coaches student conceptual depth and argumentation
Scaffolding	Phase 2	AI + Teacher scaffold writing and speaking	AI provides structured prompts due to annotations and teacher discussion
Fading	Phase 3	Student must perform independently	No scaffolding used; only student can teach AI without teacher or AI initiation.

S-TPACK: Student Technological Pedagogical Content Knowledge (Researcher's Extension of Mishra & Koehler, 2006)

Philosophy: The original TPACK framework (Mishra & Koehler, 2006) illustrates how teachers should effectively integrate technology by demonstrating the body of knowledge a

teacher needs (a combination of Content Knowledge (CK), Pedagogical Knowledge (PK), and Technological Knowledge (TK)). This researcher proposes a version of that same framework which inverts the three areas of Knowledge away from the teacher and towards the student (referred to as Student-TPACK [S-TPACK]), with the student integrating content, pedagogy, and technology at the same time as they are learning.

In S-TPACK, students determine what to teach (use CK), how to explain it (use PK), and which technological tools and prompting strategies they can use in order to support their explanation (use TK). This three-way integration of CK, PK, and TK (initiated by the learner rather than the teacher) is the most sophisticated means of integrating technology into learning during the ASIM cycle.

S-TPACK is operationalized during the ASIM Phase 3.

- CK: Students use their knowledge of grammar and linguistic concepts to explain the concept(s) to the AI.
- PK: Students use appropriate examples, analogies, and scaffolding techniques to assist the AI in understanding the concepts being taught.
- TK: Students create clear and concise prompts in order to configure how the AI should behave as a learner; this is a form of computational thinking and AI literacy (Long & Magerko, 2020; Wing, 2006) that does not exist in any current instructional model.

Part III: Alignment with Bloom's Revised Taxonomy

ASIM's three-phase progression is explicitly designed to move students through the full cognitive spectrum of Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001), from foundational knowledge recall in Phase 1 to the highest levels of evaluation and creation in Phase 3. This deliberate cognitive progression distinguishes ASIM from conventional AI-assisted models that tend to cluster student activity at the lower cognitive levels.

ASIM Phase	Bloom's Level	Cognitive Demand	Exemplary Student Action
Phase 1: AI-Guided Discovery	Remember / Understand	Retrieving prior knowledge; developing initial conceptualization with AI support	Student answers AI-generated comprehension questions and identifies vocabulary gaps
Phase 2: Human-AI	Apply / Analyze	Applying knowledge in authentic production; analyzing AI	Student writes a formal letter, submits to AI for

Co-Creation		feedback and revising	linguistic feedback, and revises based on analysis of errors
Phase 3: Student-as-Teacher Mastery	Evaluate / Create	Evaluating AI-generated content for accuracy; creating explanations and complex questions	Student teaches grammar to AI, corrects intentional AI errors with justification, and creates complex evaluative questions

Part IV: What Makes ASIM Different from Existing Models

ASIM departs from all existing AI-assisted language learning models in three fundamental and interconnected ways:

1. Role Inversion as the Core Design Principle

Model Type	Information Flow	Student Role
Traditional instruction	Teacher → Content → Student	Passive recipient
Flipped classroom	Content → Student (pre-class) → Teacher (in-class)	Active pre-class learner
Conventional AI-assisted instruction	AI + Teacher → Student	Enhanced recipient
ASIM	Phase 1: AI → Student Phase 2: AI ↔ Teacher ↔ Student Phase 3: Student → AI	Agent, co-creator, and ultimately teacher

The fourth and most distinct arrow of ASIM's information flow (Student → AI, Phase 3) is the defining innovation of the model, which removes the role of AI from an educator or a source

of expertise, and provides the student with full epistemic authority (or learning authority) over the AI. No current AI-assisted Model of Instruction identifies this as the principal role of the student.

2. The AI Confusion Strategy

In conventional language learning, students are presented with grammatical or spelling errors in isolation (decontextualized) and guess the missing word or correct spelling, a very low-level task that requires little cognitive effort or metalinguistic reasoning. The AI confusion strategy introduces a different type of challenge for students in phase 3 of the ASIM method.

Conventional Error Correction	AI Confusion Strategy (ASIM)
Errors are decontextualized and isolated	Errors are embedded within meaningful, authentic texts
Student identifies the error — surface recognition only	Student identifies, explains, and corrects the error with justification
Minimal cognitive load — recall task	High cognitive load — evaluation and metalinguistic analysis
Student has no control over error type	Student instructs AI on the type of errors to produce — agentic control
Tests recall and recognition	Tests metalinguistic knowledge, analytical competence, and evaluative reasoning

3. Prompt Engineering as Core Language Literacy

ASIM identifies prompt construction as a core language literacy rather than a technical skill; enabling students to use structured and purposeful written communication to further develop computational thinking (Wing, 2006) and AI literacy (Long & Magerko, 2020). Through the process of writing prompt responses, students learn how to create clear instructions on what to do, express what outcome they want from the response, identify the limitations of AI, and revise the prompts they write based on the results they receive. This component of ASIM does not exist in any of the current EFL instructional models, and reflects the researcher’s belief that to prepare students for an AI-dominated future, there is a need to develop technology literacy as an agent, rather than as a consumer.

Part V: Model Architecture and Phase Overview

ASIM is structured as a three-phase instructional cycle applied selectively to lessons within the eleventh-grade EFL curriculum that contain explicit productive language objectives — specifically academic writing and speaking activities. Within each instructional unit, ASIM is applied to two writing activities and two speaking activities, yielding four ASIM cycles per unit. Over the 8-to-12-week study period, ASIM was implemented across three instructional units, producing a total of twelve full cycles.

Each ASIM cycle consists of four sessions of 40 minutes each, totaling 160 minutes per cycle: one session for Phase 1 (asynchronous, pre-class), two sessions for Phase 2 (in-class), and one session for Phase 3 (in-class).

Regarding activities: The activities that are suggested within each phase provide an indication of the relevant pedagogical principles and operational logic in the implementation of each phase. The suggested activities are not intended to be prescriptive, nor are they designed to be followed strictly. Both teachers and learners should modify the activity types along with content, prompts, and tasks to address the unique subject matter, learner ability and context in each specific implementation. The table below delineates the major functions of these phases and Figure 3.1 gives an illustration of how they relate to the full ASIM structure.

Phase	Name	Sessions	Mode	Work Type	Bloom's Level
Phase 1	AI-Guided Discovery	1 × 40 min per cycle	Asynchronous Pre-class	Individual	Remember Understand
Phase 2	Human-AI Co-Creation	2 × 40 min per cycle	Synchronous In-class	Individual + Group (Act.1)	Apply Analyze
Phase 3	Student-as-Teacher Mastery	1 × 40 min per cycle	Synchronous In-class	Individual	Evaluate Create

AI-Student Inversion Model

Researcher-proposed instructional model EFL Academic Writing and Speaking

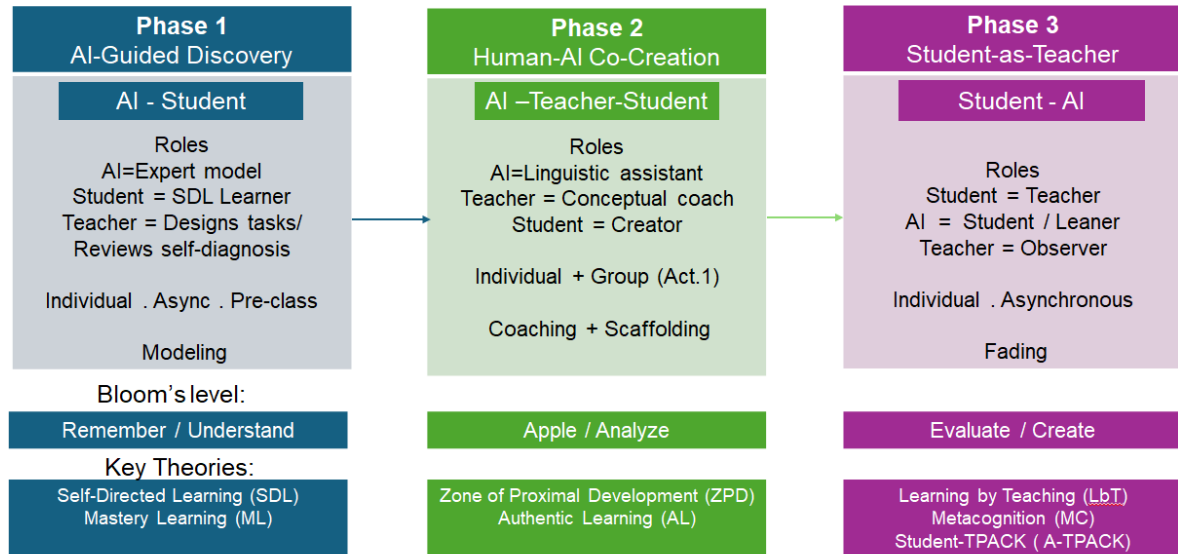


Figure 3.1. The AI-Student Inversion Model (ASIM): Three-phase instructional cycle showing role distribution, Bloom's cognitive levels, and theoretical foundations.

Part VI: Phase Descriptions

Phase 1: AI-Guided Discovery

Phase Parameters	
Duration	One 40-minute session per cycle, completed asynchronously before the in-class session
Mode	Individual, asynchronous, pre-class — student works independently with AI
Student Role	Active self-directed learner — initiates, directs, and evaluates AI interaction
AI Role	Expert model and adaptive tutor — provides examples, generates questions, offers feedback

Teacher Role	Absent from Phase 1 activity; reviews self-diagnosis reports before Phase 2
Primary Theory	Self-Directed Learning (Knowles, 1975); Mastery Learning (Bloom, 1968); SRL (Zimmerman, 2002)
CA Mechanism	Modeling — AI demonstrates expert language use; student observes and engages

Phase 1 Rationale

Phase 1 establishes a personalized knowledge base through AI-assisted discovery prior to formal classroom instruction, implementing a flipped learning approach (Bergmann & Sams, 2012). Because the student works alone with the AI at their own pace, each learner engages with the material at the depth and speed their individual proficiency requires — operationalizing the mastery learning principle that instruction should adapt to the learner, not the reverse. The self-diagnosis report produced at the end of Phase 1 serves as a diagnostic instrument that informs the teacher's design of the Phase 2 class discussion, ensuring that in-class time is targeted at genuine gaps rather than wasted on material students already understand.

Phase 1 Activities — Illustrative Examples

The following activities illustrate how Phase 1 principles operate in practice. The content prompts, and specific tasks should be adapted by the teacher and students to reflect the relevant lesson topic and targeted language skills.

Activity	Student Actions	Expected Outcomes
Activity 1 Reading with AI (10 min)	Student reads assigned text independently Asks AI to clarify unfamiliar vocabulary and implicit meanings Answers AI-generated comprehension questions at literal, inferential, and evaluative levels Pedagogical strategy: Socratic questioning	Vocabulary list with definitions and examples Answers to comprehension questions 3–5 questions for teacher discussion

	(Paul & Elder, 2007)	
Activity 2 Grammar Exploration (15 min)	Reviews grammar concept from textbook Asks AI to explain concept differently with localized examples Completes AI-generated practice exercises Receives immediate corrective feedback from AI Pedagogical strategy: Diagnosis and applied scaffolding	Personalized grammar notes with examples Completed practice exercises with corrections Specific areas to discuss with teacher
Activity 3 Vocabulary in Context (10 min)	Requests AI to create contextual examples of new vocabulary Determines formal vs. informal usage Creates original sentences and submits to AI for feedback Pedagogical strategy: Contextualized vocabulary development	Vocabulary words used in context 3–5 original student sentences AI feedback on accuracy and appropriateness
Activity 4 Self-Diagnosis Report (5 min)	Reflects on session: what was understood vs. unclear Formulates specific questions for teacher Uses AI to organize reflections into a structured report Submits report to teacher before Phase 2 Pedagogical strategy: Metacognitive self-regulation (Zimmerman, 2002)	Structured self-assessment document Prioritized list of clarification needs Evidence of metacognitive monitoring

Phase 2: Human-AI Co-Creation

Phase Parameters	
Duration	Two 40-minute in-class sessions per cycle
Mode	Synchronous, in-class — Activity 1 is whole-group discussion; Activities 2, 3, and 4 are individual
Student Role	Creator and analyst — produces authentic texts and speaking tasks, analyses AI feedback, and revises
AI Role	Linguistic assistant — provides immediate feedback on language accuracy only; does not rewrite student work
Teacher Role	Strategic facilitator and conceptual coach — addresses identified gaps, observes, and provides content feedback
Primary Theory	Zone of Proximal Development (Vygotsky, 1978); Authentic Learning (Lombardi, 2007); Distributed Cognition (Hutchins, 1995)
CA Mechanism	Coaching and Scaffolding — AI and teacher provide complementary, dual-source support

Phase 2 Rationale

Phase 2 constitutes the co-creation studio of ASIM: a space in which students produce authentic language (writing and speaking) with the support of two complementary scaffolding sources. The AI handles the mechanical dimension of linguistic feedback (grammar, vocabulary, fluency), freeing the teacher to provide the higher-order conceptual coaching that AI cannot supply — argumentation, evidence use, rhetorical structure, and motivational support. This division of labor is the operational expression of Vygotsky's ZPD principle (1978), in which targeted support from expert sources enables students to perform at levels beyond their independent capability.

Phase 2 Activities — Illustrative Examples

Activity	Procedure	Expected Outcomes
Activity 1 Targeted Class Discussion (15 min) Whole Group	Teacher collects and analyzes Phase 1 self-diagnosis reports Identifies shared gaps and misconceptions across students Leads targeted discussion addressing common difficulties only Students share insights from AI sessions and teach peers Note: This is not re-teaching — it is gap-targeted instruction based on students' own diagnosed needs	Clarification of shared conceptual difficulties Collective knowledge construction Peer teaching and reinforcement
Activity 2 Writing with AI Support (Session 1: 25 min) Individual	Step 1 (10 min): Student writes first draft independently — focus on ideas and organization, no self-editing Step 2 (10 min): Student submits draft to AI for linguistic feedback (grammar, vocabulary, clarity only — AI does not rewrite) Step 3 (5 min): Student revises draft using AI feedback Step 4: Teacher reviews revised draft and provides content feedback on argumentation and organization	Linguistically accurate text (AI-supported) Conceptually developed text (teacher-supported) Student understands revision as an iterative process
Activity 3 Conversational Practice (Session 2: 30 min)	Step 1: Teacher assigns realistic speaking scenario (e.g., job interview, university application) Step 2 (15 min): Student engages in AI	Individual speaking practice (10–15 min per student) Personalized AI feedback on

min) Individual	conversation simulation — AI plays assigned role and follows up on student responses Step 3: AI provides post-conversation feedback on pronunciation, grammar, vocabulary, and fluency Step 4: Teacher observes, notes common errors, and provides brief whole-group feedback (5 min)	oral production Teacher understanding of students' pragmatic and strategic communication abilities
Activity 4 Peer Sharing (Session 2: 10 min) Individual sharing	Students share innovative AI interaction strategies with peers Group identifies most successful approaches Discussion of solutions to common problems Builds a community of AI-literate learners	Expanded repertoire of effective AI interaction techniques Student ownership of learning tools Peer-to-peer learning and inspiration

Phase 3: Student-as-Teacher Mastery

Phase Parameters	
Duration	One 40-minute in-class session per cycle
Mode	Synchronous, in-class — all activities are individual
Student Role	Teacher and evaluator — instructs the AI, corrects its errors, and assesses its responses
AI Role	Student and learner — acts as a novice, asks questions, makes intentional errors, and receives correction

Teacher Role	Observer and assessor — reviews transcripts of student-AI teaching interactions; provides additional support where needed
Primary Theory	Learning by Teaching (Duran, 2017); Metacognition (Flavell, 1979); S-TPACK (Researcher's extension of Mishra & Koehler, 2006)
CA Mechanism	Fading — all scaffolding is withdrawn; student performs at expert level independently

Phase 3 Rationale

Phase 3 represents the culmination of the ASIM cycle and the fullest expression of its defining inversion: the student, having moved through AI-guided discovery and collaborative co-creation, now assumes the role of teacher while the AI assumes the role of student. The cognitive demands of this role are the highest in the entire cycle — to teach, the student must organize and synthesize their knowledge, formulate clear explanations, respond to questions, and detect and justify corrections. These demands correspond precisely to the Evaluate and Create levels of Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001) and to the Fading stage of the Cognitive Apprenticeship model (Collins, Brown & Newman, 1989).

Phase 3 Activities — Illustrative Examples

Activity	Procedure	Expected Outcomes
Activity 1 Teach Grammar to AI (15 min) Individual	Student configures AI as a novice student who knows nothing about the grammar/language topic Student explains the concept, provides examples, and responds to AI questions AI challenges the student with deep, probing questions and asks for clarification AI evaluates the student's teaching at the end: clarity, examples, and responsiveness	Confirmation of conceptual mastery Development of pedagogical reasoning Recognition of persistent misconceptions

	Teacher reviews transcript to assess conceptual depth and identify residual gaps	
Activity 2 AI Confusion Strategy (15 min) Individual	Student instructs AI to produce a text containing deliberate errors of specific types Student identifies all errors in the AI-generated text Student explains why each identified item is an error Student provides the correct form with justification AI confirms or challenges the student's corrections Teacher assesses: number of errors identified, quality of explanation, accuracy of correction	Demonstrated analytical and metalinguistic competence Identification of grammatical rules through application Transfer of error detection skills to student's own writing
Activity 3 Complex Questioning (5 min) Individual	Student formulates 3–5 complex, high-order questions about the lesson topic AI responds and then asks the student to evaluate and correct the response Teacher assesses question quality on a cognitive complexity scale	Evidence of comprehensive content understanding Ability to create evaluative and analytical questions Increased awareness of knowledge gaps
Activity 4 Final Reflection (5 min) Individual	Student reflects on their biggest challenge and how they overcame it Student describes specific improvements in writing or	Metacognitive awareness of learning process Understanding of personal learning trajectory Future learning goal-setting

	speaking Student sets goals for the next cycle Student rates confidence on the topic (1–10) Teacher uses reflections to monitor engagement and inform future instruction	
--	---	--

Part VII: Implementation Timeline (8–12 Weeks)

ASIM is implemented over a period of 8 to 12 weeks across three instructional units of the eleventh-grade EFL curriculum. Within each unit, ASIM is applied selectively to two writing activities and two speaking activities, yielding four ASIM cycles per unit and twelve cycles in total. Each cycle consists of four sessions (one per phase, with Phase 2 occupying two sessions), totaling 160 minutes per cycle. The following table presents a representative 10-week implementation schedule.

Week(s)	Phase	Unit Activity /	Focus	Assessment
1–2	Preparatory	Training sessions	Familiarization with ChatGPT; prompt engineering fundamentals; Technology Familiarity Questionnaire	Questionnaire; Pre-test (writing + speaking)
3	Unit 1 Phase 1	Writing Activity 1 + Speaking Activity 1	AI-Guided Discovery — two parallel cycles (one for writing, one for speaking) from Unit 1 selected lessons	Self-diagnosis reports
4	Unit 1	Writing Activity 1 +	Human-AI Co-Creation — targeted class	Writing draft + teacher feedback; Speaking

	Phase 2	Speaking Activity 1	discussion, writing with AI support, conversational practice	transcript
4	Unit 1 Phase 3	Writing Activity 1 + Speaking Activity 1	Student-as-Teacher Mastery — teach AI, AI Confusion Strategy, complex questioning, final reflection	Teaching transcript; Error correction; Reflection
5–6	Unit 1 All Phases	Writing Activity 2 + Speaking Activity 2	Second pair of cycles from Unit 1 — same three-phase structure with adapted content	Ongoing formative assessment
7–8	Unit 2 All Phases	All activities 4	Four cycles from Unit 2 — writing Activities 1 and 2, speaking Activities 1 and 2	Ongoing formative assessment
9–10	Unit 3 All Phases	All activities 4	Four cycles from Unit 3 — highest cognitive demand; increasing student autonomy across all cycles	Post-test (writing + speaking); Interviews; Final reflections
11–12	Post-intervention	Data collection	Semi-structured interviews; Observation review; Journal analysis	Qualitative data collection and analysis

Part VIII of the ASIM Study: Key Terms and Definitions

AI-Student Inversion: The basic premise of ASIM that defines the change in the conventional education relationship of AI and teacher directing the student into an inversion in Phase 3 of the student directing the AI. The inversion occurs when the student becomes the teacher and the AI becomes the student.

S-TPACK (Student Technological Pedagogical Content Knowledge): A new version of Mishra and Koehler's 2006 TPACK framework, proposed by the present researcher, based on the student being the one who integrates their content knowledge (what they teach), how they explain things (pedagogical reasoning) and how they create prompts (technical competency), all at the same time. The full implementation of S-TPACK occurs in Phase 3.

AI Confusion Strategy: A new form of assessment and teaching that has students instructs the AI to create pre-planned errors of a certain type and then identify, describe and correct the pre-planned error based on justification. This strategy operationalizes the Evaluate level of Bloom's Revised Taxonomy of teaching and assists students in developing metalinguistic awareness in a manner in which traditional error correction tasks cannot.

Prompt Engineering: The process of producing clearly structured and purposeful natural language instructions for AI tools so as to get back the specific and high-quality response that you are attempting to elicit. In ASIM, prompt engineering is essential in developing language literacy and is taught throughout all three phases.

Distributed Scaffolding: The ASIM principle of design whereby multiple complementary support sources (AI = linguistic accuracy; teacher = conceptual understanding; and peer = strategy development) provide distinct and targeted assistance to enable students to perform at levels beyond their independent capabilities.

Metacognitive Reflection: The purposeful, structured process of examining one's learning, including identifying what you have learnt; identifying what you do not yet know; and identifying what was effective for your learning. ASIM incorporates metacognitive reflection as part of every phase through the completion of self-diagnosis reports and final reflections.

Cognitive Apprenticeship: A framework for teaching that uses a structured process (i.e., modeling, coaching, scaffolding and faded support) to develop transferability and visibility of expert cognitive skills (Collins, Brown, & Newman, 1989). The three phases of ASIM map directly to this framework, with the AI acting as the expert initial model and the student performing as an independent expert in Phase 3.

Conclusion

The AI-Student Inversion Model (ASIM) represents a deliberate and theoretically-grounded departure from the prevailing paradigm of AI-assisted language education, in which students are positioned as recipients of AI-generated content and feedback. By systematically inverting the roles of all educational actors across three sequential phases — from AI-guided discovery, through collaborative co-creation, to student-as-teacher mastery — ASIM develops the full spectrum of cognitive competencies identified by Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001), cultivates the self-directed learning skills demanded by contemporary

education (Knowles, 1975), and operationalizes the principle that technology literacy as agency is fundamentally more valuable than technology literacy as consumption.

Three features distinguish ASIM from all existing AI-assisted instructional models: the role inversion that culminates in the student teaching the AI; the AI Confusion Strategy that replaces passive error recognition with active metalinguistic analysis; and the treatment of prompt engineering as core language literacy rather than a technical skill. Together, these features constitute not merely a new instructional sequence, but a new epistemological orientation toward AI in education — one in which the student, rather than the technology, is the cognitive authority.

In terms of practical implementation, ASIM was applied across three instructional units of the eleventh-grade EFL curriculum, targeting the productive language activities — writing and speaking — contained within each unit. Twelve full ASIM cycles were completed over the 8-to-12-week intervention period, each consisting of four sessions of 40 minutes totaling 160 minutes per cycle, across two school groups (male and female) in the Ramallah and Al-Bireh Governorate, Palestine.

Appendix 2

List of Expert Judges for Content Validity of Proposed Model

Name	Specialization/ Job Title	Affiliation
Prof. Khaled Dwiekat	English Language Curriculum and instruction	Al-Quds Open University Part-timer at the Faculty of Graduate Studies at An-Najah National University / Faculty of Graduate Studies at The Arab American University _Palestine
Prof. Enass Nasser	Science Curriculum and Teaching Methods	Al-Quds University
Prof. Omar Abuhumos	Assistant Professor English Language Teaching (ELT) Methods	Al-Quds University
Prof. Hussam Qadomi	Professor of English Language Teaching (ELT) Methods and Dean of the Faculty of Education	Al- Isteqlal University
Prof Mahmoud Itmeizeh	Professor of English and Dean of the Faculty of Arts	Palestine Ahliya University
Dr. Ayman Khalifah	Assistant Professor Master of Arts in Teaching Program, Director	Al-Quds Bard College for Arts and Science
Dr. Tahany Bsharat	Doctor of Philosophy in Education (English studies)	Part-timer at An-Najah University and Al-Quds Open University
Dr. Faisal Khamayseh	Associate Professor in	Palestine Polytechnic

	Computer Science	University
Dr. Majdi Alhinawi	Educational Technology	Al-Quds Open University
Dr. Islam A.Ismail	Assistant Professor at RWTC Academic AI Specialist	Editor-in-Chief at Dibon Journal of Education, UK

Appendix 3

University Letter of Facilitation for Research Implementation

Al-Quds University
Faculty of Educational Sciences
Dean's Office

جامعة القدس
كلية العلوم التربوية
مكتب العميد

2026/6/13
26/10/1111

حضرة السادة وزارة التربية والتعليم المحترمين

الموضوع: تسهيل مهمة

تحية طيبة وبعد،،

تقوم الطالبة شيماه عبد الكريم عبد الرحمن علي ورقمها الجامعي (22311022)، من تخصص دكتوراه الفلسفة في المناهج وأساليب التدريس بإعداد دراسة ، بعنوان :

" فعالية نموذج التعلم الذاتي المقترح الذي يستخدم الذكاء الاصطناعي التوليدي في مهاراتي الكتابية والتحدث باللغة الإنجليزية لدى طلاب الصف الحادي عشر، وتصوراتهم حول استخدام النموذج."

لذا يرجى من حضرتكم تسهيل مهمة الطالبة المذكورة في الحصول على المعلومات المطلوبة ولتطبيق الدراسة خلال العام الحالي .

وتفضلوا بقبول فائق الاحترام،،

أ.د. محمود أبو سمرة
عميد كلية العلوم التربوية

نسخة: الملف

Tel: 02-2794913 - Jerusalem P.O. Box 20002
تلفون: 02-2794913 - القدس ج. ب 20002

Appendix 4

Pre- and Post-Test Assessments with Scoring Rubrics

English Language Writing and Speaking Skills — Grade 11

D.1: Pre-Test: English Language Skills Grade 11 / Writing and Speaking Assessment	D.2: Post-Test: English Language Skills Grade 11 / Writing and Speaking Assessment
--	---

Student Name:	
Code:	
Total Time Allowed:	60 minutes
Total Score:	_____/ 100

Part A: Writing Skills

50 points / 40 minutes

Paragraph Writing Task (50 points)

Instructions: Choose one topic below. Write a well- organized paragraph of 80-100 words.

Your paragraph must include: a clear topic sentence, 2-3 supporting ideas with details, and a concluding sentence.

Topic 1: Health and Lifestyle How do you exert control over your environment in order to live a healthy lifestyle? Give specific examples to support your answer.	Topic 2: Family and Community What is the value of the family environment in terms of personality development? Give specific examples to support your answer.
---	---

My chosen topic: Topic 1 Topic 2

Your paragraph

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Writing Scoring Rubric

Criterion	Points	Score
Topic sentence: should be clear, and states the main idea.	10 pts	
Supporting sentences: two or more relevant ideas with examples or details.	20 pts	
Concluding sentences: summarizing the topic.	10 pts	
Grammar and sentence structure: correct verb tenses and sentence formation.	6 pts	
Vocabulary: use suitable words that related to the topic	4 pts	
Part A Total	50 pts	

Writing Score Level Descriptors

Supporting ideas (20 pts) / Descriptor

Score	Performance Descriptor
18-20	More than two separate concepts are expressed and explained; identified with relative specificity. Each idea gives relevant support to the thesis statement.
14-17	Two ideas with adequate explanation of supporting evidence. Relationships between ideas may or may not be clearly identified.

10-13	One or two concepts are presented, though they do not provide sufficient clarity in relation to the primary concept.
6-9	Only one idea offered and lacks substantive detail; it is also difficult to understand in relation to the primary concept.
0-5	No explicitly supporting concept is provided; overall response is limited in terms of responding to the assignment.

Grammar and Vocabulary (10 pts) / Descriptor

Score	Performance Descriptor
9-10	No grammar mistakes. Varied vocabulary and used.
7-8	Some errors exist, but they don't impact meaning; vocabulary is varied but sometimes misused.
5-6	Many errors, some affecting the way the reader processes information; little vocabulary variety and many repeated terms.
3-4	Errors exist consistently; very little vocabulary variety.
0-2	Overall; high level of error; text is very difficult for the audience to understand.

Part B: Speaking Skills

50 points / 10-12 minutes per student

Task one: Responding to a prompt (25 pts / 3 min)

Instructions for teacher: Show the prompt below to the student. Give one minute to think, and then ask them to speak for 3 minutes.

Health and daily life	Family and Society
Many people say that being healthy is one of the most important things in life. Talk about: One habit of health you think is very important and why. What young people can do to keep healthy with a busy lifestyle today? What habit of health you do yourself and how it impacts you.	Families play a big part in what kind of people we are and how we live our lives. Talk about: The most important value or tradition you have in your family. How families in Palestine have changed over the years. Give your opinion on what is the most important thing a family can provide for their child?

Selected prompt: Prompt 1 / Prompt 2

Task one: scoring rubric

Criterion	Points	score
Ability to speak fluently and pronounce words correctly.	8 pts	
Ability to provide all 3 items requested in the prompt with appropriate detail.	8 pts	
Ability to use a large number of correct words related to each part of the prompt.	5 pts	
Ability to use all of the correct verb tenses and sentence structures for the prompt.	4 pts	

Task one total	25 pts	
----------------	--------	--

Task one / Performance Descriptors

Fluency and pronunciation (8 pts)

Score	Performance Descriptors
7-8	Speech is fluid (very few unnatural pauses). Pronunciation is clear and easy to follow throughout.
5-6	Some pauses or hesitations (occasional), but do not considerably interfere with communication. Pronunciation generally clear.
3-4	Frequent pauses /slow delivery. In some cases, parts are new, but some errors cause parts to be difficult to follow.
0-2	Very halting or disjointed speech (long pauses) and a lot of mispronounced words, which severely impair the listener's ability to understand.

Content Coverage (8 pts)

Score	Performance Descriptors
7-8	All three points are covered with relevant and specific information fully developed with sufficient detail.
5-6	All three points are described; one or two of the points may not have sufficient detail or development.
3-4	There may be at least two points covered out of the three listed, but only one of those three points covered in detail.
0-2	At least one point has been described effectively.

Vocabulary Range (5 pts)

Score	Performance Descriptors
5	Utilizes broad selection of relevant vocabulary with accuracy. Effectively,

	reducing repetitive use of vocabulary.
3-4	Adequate vocabulary with some variety. Several instances of repeating words or using incorrect words.
1-2	Limited vocabulary. Frequent repetition, or use of only very basic words.
0	Not enough vocabulary to evaluate or cannot express ideas.

Grammar Accuracy (4 pts)

Score	Performance Descriptors
4	Grammatically correct most of the time. There are very few errors that would affect meaning.
3	There are grammatical errors in what the speaker has said, but generally the meaning is clear.
1-2	A lot of grammatical errors. Sometimes listeners have trouble understanding what is being said.
0	There are enough errors throughout that it is hard for the listener to understand what is being said.

Task Two: Interactive conversation (25 points / 5 minutes)

Teacher Instructions: Pick ONE question set, then speak to the student in a comfortable manner, using 5 complete natural questions and making sure that the student answers completely before proceeding to the next question.

Question Set 1- Health and Lifestyle

Question No	The questions	Score out of 5
1	What do you do to stay healthy? Describe your daily routine.	
2	Do you believe young people in Palestine are concerned about their health? Why or why not?	

3	What do you believe is more important for having a healthy life; eating better or exercising? Provide your thoughts.	
4	What are the ways that stress affects the health of young people, and what ways can they relieve their stress?	
5	If you could change one unhealthy habit within your community, what would it be and why?	

Question Set 2: Community and Family

Question No	The questions	Score out of 5
1	Describe your family. Is there something about your family that makes it unique or special?	
2	Who has made the biggest impact in your life and why - a family member or someone else?	
3	How have the roles of men and women changed in Palestinian families? Is this a good change?	
4	What kind of responsibilities do children have towards their parents and elderly relatives?	
5	How important is the extended family (grandparents, aunts, uncles) in Palestinian culture today?	

Selected set: Set 1____ / Set 2____

Task two total ____/ 25

Task two: Scoring Rubric

Score	Performance Descriptor
5	Appropriately answer the question with enough detail or examples. Very fluent and confident while speaking.
4	Generally answer the question with fair detail, just minor gaps in detail. Clear speaking.

3	Answered the question but was too brief or did not provide any detail. Hesitated in the spoken response because of no or minor grammar problems.
2	Gave partial comprehension to the question asked. Did provide vague response to the original question or did not stay on topic. Clear speaking was somewhat limited due to communication error(s).
1	Finding it difficult to understand the question even after being given the question again; therefore providing a severely limited answer to the question.
0	Response did not respond to the question asked or the spoken response completely did not have anything relating to the question asked.

Part A(Writing)	Part B (Speaking)	Total score
_____/50	_____/50	_____/100

Post-TEST: English language skills
Grade 11: Writing and speaking Assessment

Student Name:	
Code:	
Total Time Allowed:	60 minutes
Total Score:	_____/ 100

Part A: Writing Skills

50 points / 40 minutes

Paragraph Writing Task (50 points)

Instructions: Choose one topic below. Write a well- organized paragraph of 80-100 words.

Your paragraph must include: a clear topic sentence, 2-3 supporting ideas with details, and a concluding sentence.

Topic one: Environment and society	Topic two: Media and Technology
Discuss how individuals and their local communities are working to preserve the environment. Give specific examples (actions) and explain the situation's importance.	In what ways has technology changed how we access news and information? Please provide examples of ways technology has assisted the public and society gain access to news and information.

My chosen topic: Topic 1 Topic 2

Your paragraph

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Writing Scoring Rubric

Criterion	Points	Score
Topic sentence: should be clear, and states the main idea.	10 pts	
Supporting sentences: two or more relevant ideas with examples or details.	20 pts	
Concluding sentences: summarizing the topic.	10 pts	
Grammar and sentence structure: correct verb tenses and sentence formation.	6 pts	
Vocabulary: use suitable words that related to the topic	4 pts	
Part A Total	50 pts	

Writing Score Level Descriptors

Supporting ideas (20 pts) / Descriptor

Score	Performance Descriptor
18-20	More than two separate concepts are expressed and explained; identified with relative specificity. Each idea gives relevant support to the thesis statement.
14-17	Two ideas with adequate explanation of supporting evidence. Relationships between ideas may or may not be clearly identified.
10-13	One or two concepts are presented, though they do not provide sufficient clarity in relation to the primary concept.

6-9	Only one idea offered and lacks substantive detail; it is also difficult to understand in relation to the primary concept.
0-5	No explicitly supporting concept is provided; overall response is limited in terms of responding to the assignment.

Grammar and Vocabulary (10 pts) / Descriptor

Score	Performance Descriptor
9-10	No grammar mistakes. Varied vocabulary and used.
7-8	Some errors exist, but they don't impact meaning; vocabulary is varied but sometimes misused.
5-6	Many errors, some affecting the way the reader processes information; little vocabulary variety and many repeated terms.
3-4	Errors exist consistently; very little vocabulary variety.
0-2	Overall; high level of error; text is very difficult for the audience to understand.

Part B: Speaking Skills

50 points / 10-12 minutes per student

Task one: Responding to a prompt (25 pts / 3 min)

Instructions for teacher: Show the prompt below to the student. Give one minute to think, and then ask them to speak for 3 minutes.

Environmental issues	Food and Culture
Environmental issues are one of the major areas of concern for many people around the world at this time. Discuss in detail: What is one environmental issue you know about and why is it an issue? What can people or governments do to solve this issue? Give an example of something that you do or plan to do to help the environment.	Traditional food is integral part of almost every culture around the globe. Discuss in detail: What is a traditional food related to Palestinian culture and how is it made? How does traditional food relate to the preservation of culture and identity? Do you believe that traditional foods are becoming more or less popular and why?

Selected prompt: Prompt 1 / Prompt 2

Task one: scoring rubric

Criterion	Points	Score
Ability to speak fluently and pronounce words correctly.	8 pts	
Ability to provide all 3 items requested in the prompt with appropriate detail.	8 pts	
Ability to use a large number of correct words related to each part of the prompt.	5 pts	
Ability to use all of the correct verb tenses and	4 pts	

sentence structures for the prompt.		
Task one total	25 pts	

Task one / Performance Descriptors

Fluency and pronunciation (8 pts)

Score	Performance Descriptors
7-8	Speech is fluid (very few unnatural pauses). Pronunciation is clear and easy to follow throughout.
5-6	Some pauses or hesitations (occasional), but do not considerably interfere with communication. Pronunciation generally clear.
3-4	Frequent pauses /slow delivery. In some cases, parts are new, but some errors cause parts to be difficult to follow.
0-2	Very halting or disjointed speech (long pauses) and a lot of mispronounced words, which severely impair the listener's ability to understand.

Content Coverage (8 pts)

Score	Performance Descriptors
7-8	All three points are covered with relevant and specific information fully developed with sufficient detail.
5-6	All three points are described; one or two of the points may not have sufficient detail or development.
3-4	There may be at least two points covered out of the three listed, but only one of those three points covered in detail.
0-2	At least one point has been described effectively.

Vocabulary Range (5 pts)

Score	Performance Descriptors
5	Utilizes broad selection of relevant vocabulary with accuracy. Effectively, reducing repetitive use of vocabulary.
3-4	Adequate vocabulary with some variety. Several instances of repeating words or using incorrect words.
1-2	Limited vocabulary. Frequent repetition, or use of only very basic words.
0	Not enough vocabulary to evaluate or cannot express ideas.

Grammar Accuracy (4 pts)

Score	Performance Descriptors
4	Grammatically correct most of the time. There are very few errors that would affect meaning.
3	There are grammatical errors in what the speaker has said, but generally the meaning is clear.
1-2	A lot of grammatical errors. Sometimes listeners have trouble understanding what is being said.
0	There are enough errors throughout that it is hard for the listener to understand what is being said.

Task Two: Interactive conversation (25 points / 5 minutes)

Teacher Instructions: Pick ONE question set, and then speak to the student in a comfortable manner, using 5 complete natural questions and making sure that the student answers completely before proceeding to the next question.

Question Set 1- Education and future goals

Question No	The questions	Score out of 5
-------------	---------------	----------------

1	What field(s) are you hoping to study? Why?	
2	Which skill(s), do you believe, are necessary to have a successful career?	
3	In what ways will a good education support you in achieving your overall goals?	
4	What do you see as the greatest challenges faced by students today?	
5	What recommendations do you have for developing better skills when speaking and/or writing English?	

Question Set 2: Social Media and Information

Question No	The questions	Score out of 5
1	What is your primary source of news and information; how do you prefer to receive it?	
2	What is one method for determining if a news story is real versus fake? Provide one example.	
3	Do you believe that social media is a good source for getting your news? Explain.	
4	Why do you think it's crucial for people to keep up with the news/current events?	
5	Describe a news story you have recently seen. What was the story and what is your opinion of it?	

Selected set: Set 1____ / Set 2____

Task two total ____/ 25

Task two: Scoring Rubric

Score	Performance Descriptor
-------	------------------------

5	Appropriately answer the question with enough detail or examples. Very fluent and confident while speaking.
4	Generally answer the question with fair detail, just minor gaps in detail. Clear speaking.
3	Answered the question but was too brief or did not provide any detail. Hesitated in the spoken response because of no or minor grammar problems.
2	Gave partial comprehension to the question asked. Did provide vague response to the original question or did not stay on topic. Clear speaking was somewhat limited due to communication error(s).
1	Finding it difficult to understand the question even after being given the question again; therefore providing a severely limited answer to the question.
0	Response did not respond to the question asked or the spoken response completely did not have anything relating to the question asked.

Part A(Writing)	Part B (Speaking)	Total score
_____/50	_____/50	_____/100

Appendix 5

List of Expert Judges for Content Validity of Study Instruments

Name	Specialization	Affiliation
Prof. Khaled Dwiekat	English Language Curriculum and instruction	Part-timer at the Faculty of Graduate Studies at An-Najah National University / Faculty of Graduate Studies at The Arab American University _Palestine and Alquds Open University
Maysa Abu Ne'mah	English Language Teacher / PhD Candidate/ Curriculum and Teaching Methods	Al-Quds University
Nihad Abu Rumaileh	English Language Teacher / PhD Candidate/ Curriculum and Teaching Methods	Al-Quds University
Suha Nafe'	Educational Supervisor/ MA in English Language Teaching Methods	Ramallah and Al-Bireh Directorate of Education
Hanan Abu Al-Roos	Educational Supervisor/ MA in English Language Teaching Methods	Ramallah and Al-Bireh Directorate of Education
Bayan Salawdeh	English Language Teacher / MA in English Language Teaching Methods	Ramallah and Al-Bireh Directorate of Education
Basel Dmidi	Educational Supervisor/ MA in English Language Teaching Methods	Nablus Directorate of Education
Wasim Abu Mer'i	Educational Supervisor/ BA in English Language and	Nablus Directorate of Education

	Literature	
Rana Qadous	BA in English Language and Literature	Nablus Directorate of Education
Abdelmen'em Hakawati	BA in English Language and Literature	Nablus Directorate of Education

Appendix 6

Pre-Intervention Technology Familiarity Questionnaire

AI-Student Inversion Model (ASIM) Study (2025-2026)

Student No	
Group	Male () / Female ()
School	
Date	

Dear Student

This brief questionnaire is designed to gather information about your previous experiences with technology and artificial intelligence (AI) tools prior to beginning this study. You may provide any answer you wish - there are no correct or wrong answers — please answer each question honestly based on your experience. Your responses will be kept confidential, and will not influence your academic performance in any manner.

Completion time estimate: 5 to 7 minutes.

Section 1: General technology use

1. What type of devices do you regularly use? Tick all that apply.

Smartphone Laptop / Computer Tablet None of them

2. Approximately, how much time each day do you spend using the internet?

Less than 1 1–2 hours 3–4 hours More than 4 hours

3. Which digital tools or platforms do you use on a regular basis, Tick all that apply.

Social Media (e.g. Instagram, TikTok, etc.) YouTube Google Search Online Learning Platforms Microsoft Office or Google Docs None of them

4. What is your overall digital technology skill level?

Very Unskilled Unskilled Moderately Skilled Skilled Very Skilled

Section 2: familiarity with Artificial Intelligence AI (Chat GPT)

5. Before taking part in this study, do you know ChatGPT?

☐ Yes ☐ No

6. Have you ever used ChatGPT?

☐ Yes ☐ No

7. If yes. Please answer Questions 7-8. If no. Please go straight to Question 9.

If you have used ChatGPT, what for? (Tick all that apply)

- ☐ Academic tasks/homework
- ☐ English language learning/practicing
- ☐ Personal use/entertainment
- ☐ Translation
- ☐ Obtaining information or doing research
- ☐ Other: _____

8. Approximately how frequently did you use ChatGPT prior to this study?

- ☐ Daily
- ☐ Once in a while or twice per week
- ☐ Once or twice per month
- ☐ Only one or two times

9. How competent do you feel when using ChatGPT as a learning aid? (Please circle or tick the number corresponding to your feelings.

- 1 - No experience
- 2 – Beginner
- 3 - Have some experience with ChatGPT
- 4 - Have experience using ChatGPT
- 5 - Very experienced with using ChatGPT for learning.
- ☐ ☐ ☐ ☐ ☐

10. Have you used an AI tool such as ChatGPT or others for the purpose of studying or practicing English?

☐ Yes ☐ No

If yes, please explain what tools you've used and how you've used them:

For researcher use only (Participant classification)		
Beginner (level 1)	Intermediate (level 2)	Advanced (level 3)
Q5: NO (not heard of ChatGPT) Q6: No (never used ChatGPT) Q9: Score 1-2 Q10: No (no English AI use)	Q5: Yes (heard of ChatGPT) Q6: Yes (used occasionally) Q9: Score 2-3 Q8: Once or twice per month or less	Q6: Yes (regular user) Q9: Score 4-5 Q8: Daily or once in a while or twice per week Q 10: Yes (used AI for English)

Participant Code:	
Assigned Level:	
Researcher's Initials:	

Appendix 7

Semi Structured Interviews

After the implementation stage

Eleventh Graders' Perceptions of the AI-Based Self-Learning Model

Interview Information

Time duration	30-40 minutes
No of students	6 (3 female and 3 male)
Selection criteria	From different levels (Excellent, Good, Average)
Recording	Audio recording

Before beginning an interview, the researcher planned to provide potential participants with guidelines on what to expect. The Informed Consent text reads as follows:

Dear Student,

The first step in participating in this study is the interview. By agreeing to this interview, you are providing your voluntary participation in this research study.

If you choose not to answer any particular question or withdraw from the study at any time, there is no penalty; your information will be maintained as confidential.

During the interview, the audio recording will only be used for research purposes, and your personal information will never be included on any publications.

Your responses to this study will have no bearing on your academic grade or evaluation.

Your answers to all questions are hoped to be honest and reliable since the current study aims to explore your actual experiences and genuine opinions.

Domain	Key Questions
General Impression	How would you compare the experience of using the AI self-learning model vs. old school methods? Can you give me an example?
Reverse Chain Technique	How was the experience of “teaching” the AI vs. being taught

	by it? What did you run into while trying to do that? How has your ability to come up with questions for an AI improved? Tell me one specific time you had success with that.
Impact on Writing Skills	Have you found that your experience using ChatGPT has in turn improved your writing skills? If so, in what areas has it improved? For example: Grammar? Vocabulary? Organization of ideas. Please show us a few pieces you have written both pre and post use of our tools
Impact on speaking Skills	Has using ChatGPT made you a better speaker? How have you used ChatGPT for speaking practice? Do you feel more confidence speaking after using it?
Independence and Self-Learning	To what degree are you more self-sufficient in your own process of learning as a result of your experiences with this model? What is your method of evaluating your own growth from your work with it? Are you relying less heavily on your instructor because of your experiences using AI? How do you know that your skills have improved because of it?
Positive Aspects	What are the positive aspects of this model based on your experience? Were there any aspects that pleasantly surprised you?
Challenges and Suggestions for Improvement	What did you find challenging? Is there anything you would change about the model? Was this a technical challenge, a language challenge, or a challenge utilizing the tool?
Future Use and Recommendations	Will you use AI tools for your future learning? What would you advise another student just starting out using these tools?

Appendix 8

Student Observation Checklist

Self-Directed Learning with (GAI) — ASIM Implementation

Section A: Session Information

Observer's Name		Student Code	
Date		Session No.	1 2 3 4 5 6
ASIM Phase	Phase 1 Phase 2 Phase 3	Target Skill	Writing Speaking
Session Duration	From: _____ To: _____	Intervention Period	Beginning Middle End

Section B: Rating Scale Guide

Part 1: AI Interaction and Language Skills		
Score	Level	Descriptor
3	Demonstrated reliably and purposefully daily	Student demonstrates the behavior independently, accurately and with appropriate intent throughout the session - no prompting needed.
2	Occasionally demonstrated	Student demonstrates the behavior with hesitation or inconsistency, and/or is only occasionally prompted - may be present but not regularly employed.
1	.Not demonstrated	Behavior is not present at all, or too isolated to provide appropriate evidence against the criterion.

Part 2: Self-Directed Learning Skills		
Score	Level	Descriptor
4	Observed consistently and independently	Student shows the SDL behavior independently, appropriately, and with noticeable metacognitive intention throughout the entire session.
3	Observed frequently with minor lapses	Student presents the behavior for the most part, with only minor lapses or rare prompting needed.
2	Partially observed	Student shows the behavior sometimes or only when cued — presents but not happening or self-initiated.
1	Not observed	SDL behavior is absent or too minimal and transient to be credible evidence of the criterion.

Part 1: AI INTERVENTION SKILLS / 15 points (5 criteria ×3)

No	Criterion	Level Descriptors	3	2	1
1	Creates clear and specific prompts. Constructs prompts that clearly communicate the desired output or information needed from the AI, allowing for focused AI responses that are less likely to go astray or respond vaguely or off-target. Uses crisp phrasing when possible, while also being mindful of context and scope	3—Prompting are consistently clear and specific, and requests for information result in relevant, useful AI responses. 2—Prompting are clear but occasionally vague/excessively general. AI responses are partially useful but inconsistent in nervousness. 1—Accepts whatever the AI says without purposeful prompting.			

2	Uses the AI Confusion Strategy well Student directs the AI to make errors of a manageable type, then independently identifies, explains, and corrects those errors using explicit linguistic reasoning — the crux of the metacognitive activity of ASIM Phase 3	3 — Student directs AI precisely, identifies all or most errors, gives clear explanations, and rectifies with accurate justification. 2 — Student attempts the strategy but misses some errors, provides incomplete explanations, or needs prompting to justify corrections. 1 — Student does not use the strategy, or uses the strategy without identifying/explaining errors — accepts AI errors on faith			
3	Critically engages with AI outputs Student critically evaluates the content generated by AI (e.g., asks questions about the accuracy, relevance, and quality of input or output), verifies content, and uses or rejects AI provided suggestions selectively	3 — The student consistently questions, verifies, and assesses the validity of AI outputs and demonstrates sufficient formation of independent critical judgment in determining what data to integrate vs. retain. 2 — Student demonstrates partial critical engagement when reviewing AI outputs but occasionally accepts outputs uncritically or without evaluating the validity. 1 — Student demonstrates full acceptance of all AI outputs without evaluating or skepticism.			
4	Student rephrases and retries prompts as needed Student iteratively rephrases prompts in response to poor, nonsensical, or incomplete AI output - demonstrating adaptive thinking, persistence, and understanding of how the wording of a prompt affects results.	3 - Student autonomously detects bad output and rephrases prompts in strategic and astute ways to produce better results. 2 - Student attempts to rephrase prompts but inconsistently, or after teacher prompting. 1 - Student prompts are never or almost never rephrased, either accepting poor outputs or abandoning the whole thing when the AI doesn't get it right			

5	Prompt engineering It is a method of systematically creating prompts so that the artificial intelligence (AI) responds in a way that meets a particular learning objective of the student. The definition of the levels of proficiency in using prompt engineering is as follows:	3 - The student applies prompt engineering techniques with strategic intent at all times, and their prompts reflect a high-level understanding of how to properly create prompts in order to achieve specific and clearly defined learning objectives. 2 - The student applies some prompt engineering techniques, but they are not always done with strategic intent. 1 - Student does not use prompt engineering; therefore, their prompts are simply general requests to the AI without direction on how the AI should respond.			
---	---	---	--	--	--

Part 2: SELF-DIRECTED SKILLS / 20 points (5criteria × 4)

No	Criterion	Level Descriptors	4	3	2	1
1	Sets clear goals and plans work steps Student sets out what they will learn this session, and what they'll do for it, before they start — showing some proactive self-regulation rather than waiting for the teacher to tell them what to do	4 — Student states clear goals for session before starting, plans out a structured sequence to pursue, and comes back to the goals to check how it's going as they progress. 3 — Student states general goals and plans approach, but fairly vague, or wanders away from them slightly during the session. 2 — Student shows minimum goal-setting — begin work without purpose, or states only when prompted by the observer. 1 — No evidence of goal-setting or planning — just dives in, or waits for a telling.				

2	Works independently without teacher initiation The student participates in and continues learning activities independently with little or no active prompting or direction from the teacher during the session. The student's self-initiation (initiating own learning activities) is evident throughout the entire session.	4 - The student consistently begins all tasks independently, is engaged in tasks for the duration of the session, and has no teacher prompting when transitioning between activities. 3 - The student begins most tasks independently, with only a few exceptions requiring light teacher prompting. 2 - The student requires teacher prompting to begin or continue tasks on more than one occasion, indicating a reliance on teacher-led initiation. 1 - The student relies on the teacher's direction when beginning or continuing any activity.				
3	Self-Monitoring/Self-Assessment of Progress Active student tracking of understanding; observes when his or her work does not meet their goals and make any necessary adjustments according to the self-evaluation. Independent of teacher-provided feedback as to the need for change.	4—Actively tracks understanding, realizes their work isn't achieving their goal, and changes their work according to the self-assessment; doesn't rely on teacher prompting to know when to change. 3—Watches own progress, knows when to change, and changing of approach is self-initiated. 2—Basic self-monitoring; major changes of approach are 'teacher initiated' feedback. 1—Shows no monitoring/self-assessing; goes back and forth making mistakes and misses explaining himself to teacher				
4	Recognizes personal strengths & weaknesses A student demonstrates metacognitive awareness	4 — Student accurately identifies their own strengths and weaknesses in relation to the goals of this session, uses this knowledge to guide their subsequent				

	through accurate identification of their strengths and areas for improvement in this session, as well as demonstrating communicative awareness of this ability with the AI or observer.	learning choices and articulates their self-awareness of their strengths and weaknesses clearly. 3 — Student has identified general strengths/weaknesses but with little precision, and did not connect this to future learning modifications. 2 — Student does have some metacognitive awareness; can identify strengths or weaknesses from prompts, but does not demonstrate any form of independent use for guiding future learning. 1 — Student has no documented evidence of metacognitive self-awareness; student cannot (or does not) identify their own strengths or weaknesses.				
5	Shows persistence when facing challenges Student continues to approach difficult tasks, generating alternative strategies when initiatives fail, and doesn't disengage or seek teacher rescue in the face of difficulty — demonstrating the resilience at the heart of SDL.	4 — Student independently persists in working through challenges, independently generates alternative strategies when initial attempts fail, and remains fully engaged even through the most challenging tasks. 3 — Student shows good persistence most of the time, making a reasonable effort to engage with challenge before soliciting support, occasionally. 2 — Student engages some of the time, but pulls away or asks the teacher for help relatively quickly when faced with difficulty. 1 — Disengages or asks and requires teacher help at the first sign of difficulty — no evidence of independent				

		persistence				
--	--	-------------	--	--	--	--

Part3: LANGAUGE SKILLS DEVELOPMENT / 15 points (5 criteria × 3)

No	Criterion	Level Descriptors	3	2	1
1	Use of AI to enhance language precision Students submit their oral or written assignments for linguistic evaluation by an AI tool and then use the evaluation results to carry out specific corrective actions on their work instead of passively receiving the results of the evaluation without taking any action.	3 - The student has a regular pattern of seeking AI evaluations of their work, taking the time to read the evaluation results thoroughly, and making specific targeted corrections to their assignments based upon the AI's recommendations. 2 - The student has a sporadic pattern of seeking AI evaluations, making some changes to their submitted work but ignoring or incorrectly utilizing others. 1 - The student does not evaluate their language accuracy using AI or evaluates their language using AI but takes no significant action as a result of the evaluation.			
2	Looking for Reasoning Behind Mistakes When a student makes a mistake, they don't just fix it, but they also ask why the mistake happened, so they won't make that same mistake again in the future.	3- Student always asks for reasoning behind mistakes, understands the reasoning and uses that understanding so their future work is significantly better than the original work. 2- Occasionally student asks for reasoning behind errors, but doesn't regularly use this understanding to improve future work. 1-Student does not attempt to obtain reasoning before they correct their mistakes, or does not attempt to correct any errors.			
3	For Writing: Meaningfully revises / For Speaking: Practices dialogue	3 -Writing: Revisions represent a genuine effort to consolidate feedback and improve content, structure, and accuracy. Speaking: Practice is extended and shows clear development from session to session in			

	For writing sessions: Student revises their own writing based on the feedback we give, not simply re-submitting “for compliance.” For speaking sessions: Student has done more than the minimum or used a short practice dialogue that shows development in fluency, confidence, or oral accuracy.	fluency and confidence. 2-Writing: Revisions are surface-level, or incomplete. Speaking: Dialogue practice is present, though minimal, and the growth in fluency or overall development is less present. 1- Writing: Little or no substantial revision, weakness in bright marks. Speaking: No extended practice, little growth evident.			
4	Integrates personal ideas with AI suggestions Student retains their own voice, authorship, and cognitive ownership by putting their thinking, analysis, and creativity into the AI output — not replacing their ideas with AI suggestions or copying AI verbatim.	3-Student clearly retains their voice, and their own ideas, uses AI suggestions to enhance their thinking and work is their own. 2- Student attempts to integrate their own ideas but leans more on AI than learning productively allows — their voice is somewhat lost. 1-Student copies AI or works closely in line with AI — their voice and cognitive ownership are absent.			
5	Be able to demonstrate a measurable level of language progress in the session. The language output from the beginning of this session to the end has shown observable and measurable improvement: accuracy, fluency, vocabulary, organization and/or ability to communicate effectively.	3 - An observable improvement was made in all areas of the language output by the end of the session compared to the start, AND the student's language output is significantly better in at least 2 areas than at the beginning of the session. 2 - Language output improvement was made in only 1 of the language areas. 1 - No language output improvement was noted at the end of the session compared to at the beginning of the session.			

Section C: OVERALL ASSESSMENT

Section	Max Score	Score Obtained	Weight
Part 1: AL Interaction Skills	15		30%
Part 2: Self-Directed Learning Skills	20		40%
Part 3: Language Skills Development	15		30%
Total	50		100%

Rate	Score Range	Percentage	Select
Excellent	45-50 pts	90-100%	
Very good	37.5-44.5 pts	75-89%	
Good	30-37 pts	60-74%	
Acceptable	25-29.5 pts	50-59%	
Needs Improvement	Below 25 pts	Below 50%	

Section D: Qualitative Notes

Level of Student's Engagement Overall	☐ Low ☐ Medium ☐ High ☐ Very High
Key Behaviors Observed	
Progress Made Since Last Session	

Language Observations (Errors, Breakthroughs, Recurring Patterns)	
Recommendations and Follow Up	

Appendix 9

Researcher's Reflective Journal — Weekly Entry

Entry Information

Item	Details
Entry Number	
Date	
Week of Implementation	1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 / 11 / 12
ASIM Phase	Phase 1 Phase 2 Phase 3
Groups Observed	Female Male Both

Section 1 — Model Application and Implementation

Outline the progress of the ASIM intervention during the week. Was the session delivery impacted? What logistical obstacles emerged?

.....
.....
.....
.....
.....

What changes did you make during implementation? Why?

.....
.....
.....
.....
.....

Section 2 — Student Performance and Behavior

What significant changes in student performance or behavior were noticed this week?

.....
.....
.....
.....
.....

Were there discrepancies between the performance of the male and female groups? Describe

.....
.....
.....
.....
.....

Any individual students who need special attention or follow-up? (Using student codes, if applicable)

.....
.....
.....
.....
.....

Section 3 — Connecting Quantitative and Qualitative Observations

Do the patterns you're noticing seem qualitatively aligned with what you expect student progress to look like? Any comments that might point to anticipated quantitative results?

.....
.....
.....
.....
.....

Section 4 — Emerging Themes and Unexpected Phenomena

What are some recurring patterns or surprises that you've found this week with regard to your research questions?

.....
.....
.....
.....

Do you need to edit your research process or tools based on this?

Yes — what? /No

Section 5 — Researcher Reflexivity

How are your own assumptions, expectations, or prior knowledge potentially influencing your observations and interpretations this week?

.....
.....
.....

What will you do differently next week?

.....
.....
.....

Section 6 — Summary Keywords (3–5 words that capture this week's entry)

_____ / _____ / _____ / _____ / _____

فعالية نموذج مقترح للتعلّم الموجه ذاتيًا باستخدام الذكاء الاصطناعي التوليدي في تنمية مهارات الكتابة والتحدث باللغة الإنجليزية لدى طلبة الصف الحادي عشر، وتصوراتهم نحو

استخدام النموذج

الملخص

هدفت هذه الدراسة إلى تصميم وتطبيق وتقييم فعالية نموذج التأمل الطلابي والذكاء الاصطناعي (ASIM) في تنمية مهارتي الكتابة الأكاديمية والتحدث باللغة الإنجليزية لدى طلبة الصف الحادي عشر في فلسطين. اتبعت الدراسة المنهج التجريبي بتصميم شبه تجريبي ومنهج مختلط (كمي، نوعي). تكوّنت العينة من 50 طالبًا وطالبة (28 إناث و22 ذكور) من مدرستين حكوميتين في محافظة رام الله والبيرة خلال العام الدراسي 2025-2026. ولتحقيق أهداف الدراسة، تم تطوير اختبارين في الكتابة الأكاديمية والتحدث (اختبار قبلي واختبار بعدي)، إلى جانب قائمة ملاحظة، ودليل مقابلة شبه منظمة، ويوميات تأملية للباحثة والمعلمين المتعاونين. وتم التحقق من صدق وثبات هذه الأدوات. أظهرت النتائج فروقًا ذات دلالة إحصائية ($\alpha \leq 0.05$) بين درجات الاختبار القبلي والبعدي في مهارتي الكتابة الأكاديمية والتحدث، لصالح الاختبار البعدي، لدى كل من مجموعتي الإناث والذكور اللواتي تلقين تدخل ASIM. علاوة على ذلك، تفوقت الطالبات على الطلاب في الاختبار البعدي في كلتا المهارتين. وكشفت بيانات الملاحظة عن زيادة مطردة في سلوكيات التعلم الذاتي عبر الجلسات الست، حيث ارتفع المتوسط من (31.67) في الجلسة الأولى إلى (46.50) في الجلسة السادسة. وكشفت النتائج النوعية المستمدة من المقابلات شبه المنظمة واليوميات التأملية أن الطلبة أدركوا نموذج ASIM بوصفه نهجًا تعليميًا مختلفًا بشكل ملحوظ وأكثر جذبًا، بدا أنه يدعم تطور سلوكياتهم في التعلم الذاتي، ووعيهم المعرفي، وثقتهم في الكتابة الأكاديمية، رغم التحديات الأولية المرتبطة بالقيود التقنية والمتطلبات المعرفية للمرحلة الثالثة من النموذج. وأوصت الدراسة باعتماد نموذج ASIM إطارًا منظمًا لتدريس مهارتي الكتابة والتحدث باللغة الإنجليزية على جميع مستويات المرحلة الثانوية.

الكلمات المفتاحية: الذكاء الاصطناعي التوليدي، التعلم الذاتي، مهارة الكتابة الأكاديمية باللغة الإنجليزية، مهارة التحدث باللغة الإنجليزية، نموذج انعكاس الطالب والذكاء الاصطناعي، الفعالية.