

**MULTIPLE-CASE STUDIES OF THE COMPLEXITY OF EFL TEACHERS' WRITING
ASSESSMENT BELIEFS AND PRACTICES IN A PREPARATORY YEAR PROGRAM
AT A SAUDI UNIVERSITY**

Wid Altalhi

A Dissertation Submitted to the Faculty of Graduate Studies in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

Graduate Program in Linguistic and Applied Linguistics

York University

Toronto, Ontario

May 2024

© Wid Altalhi, 2024

Abstract

The relationship between teachers' beliefs, practices, and contextual factors has been characterized as complex. Consequently, Complexity Theory (CT) has recently been instrumental in dissecting this nuanced relationship. However, its application to teachers' beliefs regarding second language (L2) writing assessment remains an underexplored area in the literature. This qualitative study employed CT to explore the interplay of beliefs, practices, and contextual factors influencing teachers' assessment of second language (L2) writing within a Saudi university's Preparatory Year Program (PYP). The multiple-case study design involved five Saudi English as a Foreign Language (EFL) teachers, using interviews, observations, think-aloud protocols, and document analysis to uncover the nuanced interplay between belief systems and actual practices within the ecological systems of the macrosystem, exosystem, and microsystem. Findings indicated that EFL teachers held heterogeneous and interactive beliefs, encompassing beliefs concerning teaching and learning L2 writing, assessment purposes, methods, evaluation criteria, and assessment processes and tools. Teachers' core beliefs about teaching and learning writing, shaped by their personal learning histories (microsystem), varied significantly among them and influenced their peripheral beliefs and practices. Tensions between beliefs and practices primarily emerged from exosystemic external factors, such as fixed assessment policies, curricular requirements, and teacher autonomy limitations. However, a harmony between beliefs and practices was noted where teachers exercised greater autonomy, especially in providing feedback. Factors from the macrosystem showed no direct influence on beliefs or practices. The study highlights the complexity of teachers' realities and the need for development programs and policy reforms attuned to teachers' beliefs and contextual challenges in Saudi higher education.

Dedication

To my steadfast and loving family—this work is a tribute to each of you.

To my father, Matar, whose prayers and thoughtful support have been a cornerstone of my resolve—I carry your wisdom in every step.

To my mother, Saliha, whose endless encouragement and unwavering belief in my potential have been the winds beneath my wings—your support has been a guiding star through the darkest nights of this journey.

To Kholoud and Raghad, my two wonderful sisters, whose unwavering love and motivation have enveloped me with strength and joy—your solidarity is the warmth in my heart.

To Mansour, my incredibly supportive brother, whose steadfast presence has been my anchor in turbulent seas—I am indebted to your brotherly love and faith in me.

Your collective presence and boundless support have buoyed me through this remarkable journey. This accomplishment bears your imprint as much as mine; the path would have been shadowed without your light. With all my love, I extend my deepest gratitude and dedicate this milestone to you.

Acknowledgments

First and foremost, I am grateful to Allah Almighty (God) for granting me good health, patience, the ability to think, and the strength to complete this dissertation successfully.

Secondly, I would like to express my sincere gratitude to Dr. Khaled Barkaoui, my advisor and mentor. Thank you for your unwavering support and cooperation whenever it was needed. I am truly thankful for your consistent guidance, invaluable support, dedicated time, and insightful feedback you provided throughout every stage of my dissertation. Your expertise and broad knowledge have been instrumental in nurturing my thinking. Without your invaluable contributions, this dissertation would not have come to fruition.

I extend my sincere gratitude to my advisory committee members, Drs. Antonella Valeo and Geoff Lawrence, for their constructive feedback and warm encouragement. Additionally, I am extremely thankful to the other examining committee members: Drs. Beverly Baker, and Connie Mayer. Their profound knowledge and insightful comments transformed my Ph.D. defense into an enriching learning experience.

Heartfelt thanks to Rose Frezza-Edgecombe, the graduate program assistant at the DLLL, for her continuous guidance and support. Rose, as the cornerstone of the program, your efforts have significantly eased my academic journey, transforming complex processes into manageable steps. Your counsel has been invaluable, and for that, I am truly thankful.

Last, but not least, I would like to express my sincere gratitude to the Ministry of Higher Education in my country, the Kingdom of Saudi Arabia, for granting me this scholarship and offering me the opportunity to pursue my dream of obtaining a Ph.D. degree. I am honored and proud to be a citizen of a nation that truly supports its people. Now is the time for me to invest all my acquired knowledge in contributing to our advancement.

TABLE OF CONTENTS

Abstract	ii
Dedication	iii
Acknowledgments	iv
TABLE OF CONTENTS	v
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ACRONYMS	xi
CHAPTER 1: INTRODUCTION	1
1.1 Research Questions	7
1.2 Significance of the Study	8
1.3 Overview of the Dissertation Chapters	9
CHAPTER 2: LITERATURE REVIEW	11
2.1 Language Teachers' Cognition	11
2.2 Emerging Themes from Language Teachers' Beliefs Research	12
2.2.1 Definition of Terms	13
2.2.2 Methodological Challenges	17
2.2.3 Factors Affecting Language Teachers' Beliefs	18
2.3 Significance of Teachers' Cognition Research	19
2.3.1 Relationship Between Teachers' Beliefs and Practices	22
2.4 L2 Writing Assessment	26
2.5 A Paradigm Shift in L2 Writing Assessment	27
2.5.1 Testing Culture Vs. Assessment Culture	29
2.6 Language Teachers' Assessment Beliefs and Practices	30
2.7 Theoretical Framework: Complexity Theory (CT)	35
2.8 CT and Teacher Cognition Research	40
2.8.1 Significance of Using CT in Teacher Cognition Research	46
2.8.2 CT as a Metaphor	48
2.8.3 Ecological View of Context	49
CHAPTER 3: CONTEXT OF THE STUDY	53
3.1 Education and ELT in KSA	53
3.2 PYP in KSA	54
3.2.1 Teaching and Assessing Writing in PYPs	56
3.3 The Context of the Study	59
3.3.1 PYP in a University for Health Sciences in KSA	59
3.3.2 English Teachers in the PYP	61
3.3.3 PYP Writing Courses and the Role of Writing Teachers	62
CHAPTER 4: METHODS	74
4.1 Qualitative Research	74

4.1.1 Multiple-Case Study Design	75
4.2 Sampling Procedures	77
4.2.1 Recruitment of Participants.....	79
4.3 Data Collection Methods	81
4.3.1 Initial Interviews	85
4.3.2 Classroom Observations	89
4.3.3 Think-aloud Protocols.....	93
4.3.4 Follow-up Interviews.....	99
4.3.5 Documents	100
4.4 Data Analysis	101
4.4.1 Considerations for Analyzing Complex Adaptive Systems.....	102
4.4.2 Transcription, TAPs Initial Data Analysis, and Translation.....	104
4.4.3 Within-Case Analysis	110
4.4.4 Document Analysis.....	115
4.4.5 Cross-Case Analysis	116
4.5 Trustworthiness.....	117
4.6 Ethical Considerations	120
CHAPTER 5: THE CASE OF LEENA: Skill-based Approach and Assessment of Grammar	
Accuracy	123
5.1 Background for Leena.....	123
5.2 Teaching and Learning Writing: A Skill-based Approach	124
5.3 Writing Assessment Purposes: Direct Assessment Methods of Writing	126
5.4 Criteria for Evaluating Writing: Assessing Grammar Accuracy	127
5.5 Feedback: An Indirect Approach to Feedback.....	129
5.6 Grading and Rubric Use: The Significance of Objective and Fair Assessment	131
5.7 Ecological Factors Influencing Leena’s Writing Assessment Beliefs and Practices....	132
5.7.1 The Influence of Leena’s Learning Experience on her Writing Assessment Beliefs and Practices	132
5.7.2 Tensions Between Leena’s Writing Assessment Beliefs and Practices.....	135
5.7.3 Harmony between Leena’s Writing Assessment Beliefs and Practices.....	139
5.8 Summary of Leena’s Case Findings	143
CHAPTER 6: THE CASE OF AYA: Assessing Individual Progress	147
6.1 Background for Aya.....	147
6.2 Writing Assessment Purposes: Assessment of Learning	148
6.3 Teaching and Learning Writing: Emphasizing Idea Generation and Content Quality .	149
6.3.1 Criteria for Evaluating Writing: Prioritizing Clarity of Ideas.....	151
6.4 Writing Assessment Methods	152
6.5 Feedback: Preference of In-class Individual Verbal Feedback.....	153
6.6 Grading and Rubric Use: The Importance of Fairness	156
6.7 Ecological Factors Influencing Aya’s Writing Assessment Beliefs and Practices	159

6.7.1 The Influence of Aya’s Learning Experience on her Writing Assessment Beliefs and Practices	159
6.7.2 Tensions Between Aya’s Writing Assessment Beliefs and Practices.....	161
6.7.3 Harmony between Aya’s Writing Assessment Beliefs and Practices.....	165
6.8 Summary of Aya’s Case Findings	168
CHAPTER 7: THE CASE OF BATOOL: The Value of Formative Methods and Self-Assessment	171
7.1 Background for Batool.....	171
7.2 Writing Assessment Purposes: Formative Assessment	172
7.3 Writing Assessment Methods: Formative Methods and Self-Assessment	174
7.4 Criteria for Evaluating Writing: PYP Rubric Influence	178
7.5 Feedback: The Integration of Written and Verbal Feedback.....	179
7.6 Grading and Rubric Use: The Importance of Fairness	182
7.7 Ecological Factors Influencing Batool’s Writing Assessment Beliefs and Practices... ..	183
7.7.1 The Influence of Batool’s Learning Experience on her Writing Assessment Beliefs and Practices	184
7.7.2 Tensions Between Batool’s Writing Assessment Beliefs and Practices.....	187
7.7.3 Harmony between Batool’s Writing Assessment Beliefs and Practices.....	192
7.8 Summary of Batool’s Case Findings	193
CHAPTER 8: THE CASE OF NOOR: Assessing Meaning and Unity in Informal Writing	196
8.1 Background for Noor	196
8.2 Teaching and Learning Writing: Informal Writing through the Integration of a Multimedia Learning Approach.....	196
8.3 Writing Assessment Purposes: Improving Students’ Learning Outcomes	198
8.4 Criteria for Evaluating Writing: Focusing on Meaning Over Mechanics.....	199
8.5 Writing Assessment Methods: Emphasis on Formative Methods	202
8.6 Feedback and Grading	203
8.7 Ecological Factors Influencing Noor’s Writing Assessment Beliefs and Practices	207
8.7.1 The Influence of Noor’s Learning Experience on her Writing Assessment Beliefs	207
8.7.2 Tensions Between Noor’s Writing Assessment Beliefs and Practices	211
8.7.3 Harmony between Noor’s Writing Assessment Beliefs and Practices	215
8.8 Summary of Noor’s Case Findings.....	219
CHAPTER 9: THE CASE OF RANA: An Affective Approach to the Teaching and Assessment of Writing.....	223
9.1 Background for Rana	223
9.2 Teaching and Learning Writing: An Interactive and Student-centered Class	224
9.3 Writing Assessment Purposes: Formative and Summative	226
9.4 Criteria for Evaluating Writing: Ideas and Students’ Psychology.....	228
9.5 Writing Assessment Methods: Direct and Indirect Methods	230

9.6 Feedback: Critique of PYP Feedback Policies	232
9.7 Grading and Rubric Use: Teacher-centered Rating Approach	235
9.8 Ecological Factors Influencing Rana’s Writing Assessment Beliefs and Practices	236
9.8.1 The Influence of Rana’s Teaching and Learning Experience on her Writing Assessment Beliefs and Practices	237
9.8.2 Tensions Between Rana’s Writing Assessment Beliefs and Practices	240
9.8.3 Harmony between Rana’s Writing Assessment Beliefs and Practices	245
9.9 Summary of Rana’s Case Findings.....	246
 CHAPTER 10: CROSS-CASE FINDINGS	250
10.1 Heterogeneity of Teachers’ L2 Writing Assessment Belief System	251
10.1.1 The Content of Teachers’ Belief Systems	251
10.1.2 The Nature of Teachers’ Beliefs	263
10.2 Interactions Between Teachers’ Beliefs, Practices, and Contextual Factors	265
10.2.1 Factors within Ecological Systems	266
10.2.2 The Influences of Macrosystem Factors on Teachers’ Beliefs and Practices	267
10.2.3 The Influences of Exosystem Factors on Teachers’ Beliefs and Practices	268
10.2.4 The Influences of Microsystem Factors on Teachers’ Beliefs and Practices	271
10.3 Summary of Cross-case Findings	274
 CHAPTER 11: DISCUSSION AND CONCLUSIONS	279
11.1 Summary and Discussion of the Findings	279
11.2 Limitations of the Study.....	286
11.3 Implications.....	289
11.3.1 Implications for Teachers	289
11.3.2 Implications for Teacher Education.....	291
11.3.3 Implications for Policy Makers and Curriculum Developers	293
11.3.4 Implications for Future Research.....	295
 References.....	298
Appendices.....	335
Appendix A: York University Ethics Approval	335
Appendix B: Informed Consent Form for Participants	336
Appendix C: Initial Interview Questions Guidelines	339
Appendix D: Informed Consent Form for Students	342
Appendix E: Classroom Observation Guidelines	343
Appendix F: A Sample of an Observation Field Note	345
Appendix G: Think-aloud Protocols Instructions and Training Guidelines.....	348
Appendix H: The Descriptions and Instructions of the English II Two Writing Tasks	349
Appendix I: Follow-up Interview Questions Guidelines.....	352
Appendix J: An Example of MAXQDA Coding	353

LIST OF TABLES

Table 3.1	Brief Description of the Three Writing Courses in the PYP	62
Table 3.2	Summary of the Writing Assessment Methods in English II	69
Table 3.3	Copy of the Writing Paragraph Tasks Rubric.....	71
Table 4.1	Participants' Demographic Data.....	81
Table 4.2	Data Collection Overview	82
Table 4.3	The Data Collection Timeline of Initial Interviews and Observations	83
Table 4.4	A Brief Description of the TAP Sessions Tasks.....	96
Table 4.5	The Data Collection Timeline of TAPs and Follow-up Interviews.....	97
Table 4.6	Transcription Conventions (Adapted from Seedhouse, 2004).....	105
Table 4.7	TAPs Initial Coding (Adapted from Cumming, 2002, p. 88)	107

LIST OF FIGURES

Figure 2.1	Elements and Processes in Language Teacher Cognition (Borg, 2006, p. 283)	18
Figure 2.2	The Nested Ecological Systems and the Interactions Between Multiple Layers (Adapted from Bronfenbrenner, 2005)	52
Figure 3.1	MCQs Example from the Paraphrasing Task	66
Figure 3.2	Copy of an Online MCQs Task	67
Figure 3.3	MCQ Examples from the Quiz	67
Figure 5.1	Example of Leena’s Written Feedback in the First Task	134
Figure 5.2	Example of Leena’s Written Feedback that Focuses on Grammar	141
Figure 5.3	Summary of Leena’s Case Findings	146
Figure 6.1	Example of Aya’s Written Feedback in the First Task	162
Figure 6.2	Example of Aya’s Written Feedback in the Second Task	167
Figure 6.3	Summary of Aya’s Case Findings	170
Figure 7.1	Example of Batool’s Written Feedback in the First Task	191
Figure 7.2	Example of Batool’s Written Feedback in the Second Task	191
Figure 7.3	Summary of Batool’s Case Findings	195
Figure 8.1	Example of Noor’s Written Feedback in the First Task	213
Figure 8.2	Example of Noor’s Written Feedback in the Second Task	216
Figure 8.3	Summary of Noor’s Case Findings	222
Figure 9.1	Example of Rana’s Written Feedback in the First Task	240
Figure 9.2	Summary of Rana’s Case Findings	249

LIST OF ACRONYMS

AaL	Assessment as Learning
AfL	Assessment for Learning
AoL	Assessment of Learning
CAS	Complex Adaptive System
CT	Complexity Theory
EFL	English as a Foreign Language
ELT	English Language Teaching
ETEC	Education and Training Evaluation Commission
FI	Follow-up Interview
II	Initial Interview
KSA	Kingdom of Saudi Arabia
L1	First Language
L2	Second Language
MCQ	Multiple-choice Question
OBS	Observation
PYP	Preparatory Year Program
SCT	Sociocultural Theory
TAP	Think-Aloud Protocol
UC	University Council

CHAPTER 1: INTRODUCTION

This study used Complexity Theory (CT) to explore the L2 writing assessment beliefs and practices of EFL teachers in a Preparatory Year Program (PYP) at a Saudi University, alongside the examination of the interrelations among those beliefs, practices, and contextual factors.

Writing, as a constituent of productive linguistic capabilities, embodies a multifaceted intellectual endeavor necessitating substantial cognitive engagement. The acquisition of writing proficiency is acknowledged as one of the most challenging feats for learners of English as a Second Language (ESL) and EFL (Hayes, 2000; Hedge, 2000; Weigle, 2002). The significance of writing extends into the global arena, where its pedagogy attracts increased examination within the context of global education and the scholarly exploration of L2 (Weigle, 2002). Mastery of L2 writing is not merely an academic pursuit but a pivotal skill for personal, educational, and professional advancement, enabling individuals to exhibit and disseminate knowledge across diverse fields (Adler-Kassner & O'Neill, 2010). Consequently, the imperative of L2 writing education is matched by the necessity for its assessment to be reliable, valid, authentic, and practical, both for immediate pedagogical outcomes and as a predictor of learners' prospective academic and professional achievements (Crusan, 2010; Weigle, 2002).

Writing assessments have long captured considerable attention in literacy education. Research studies and language teachers' experiences across various contexts and educational levels indicate that writing assessments are highly demanding tasks for teachers. This demand can be attributed to the complexity of writing (Hyland, 2003; Weigle, 2007). Researchers contend that assessing learners' writing skills is often a source of frustration for language teachers because it is both time-consuming and demanding (Brown, 2001; Neff-Lippman, 2012),

and teachers are frequently excluded from the assessment creation process (Huot, 1996). Weigle (2007) observed that teachers often view assessment as “a necessary evil” rather than as a central aspect of teaching with the potential to benefit both teachers and students (p. 194).

There are numerous factors that contribute to the complexity and difficulty of writing assessments for teachers. Crusan (2010) posited that teachers have been marginalized from the writing assessment process for an extended period, despite assertions that assessment should be considered “every teacher’s job” (p. 10). Crusan elaborated that historically, writing assessment responsibilities have been shouldered by assessment developers, whereas ideally, such assessments ought to be crafted and executed by teachers at a local level to ensure authenticity. Assessments devised without adequate contextual grounding often spark disputes over what constitutes ‘a fair assessment’ among assessment developers and teachers (White, 2001). The exclusion of teachers from the assessment process has resulted in a deficiency in their understanding of writing assessments and a waning interest in this crucial area (Yancey & Huot, 1997). Furthermore, Huot and O’Neill (2009) contended that preventing teachers from participating in writing assessment processes undermines their preparedness for assessment, leading to perceptions of assessment as punitive for both themselves and their students. However, Calfee and Miller (2013) argued that the capacity for teachers to design assessments that are consistent with their student’s needs, instructional goals, and specific contexts stands as a pivotal criterion for effective writing assessment.

The complexity of writing assessment lies in the fact that its tools and practices may not be deemed useful or effective if removed from their context, given that assessment should be “site-based and locally controlled” (Huot, 2002, p. 19). The conceptualization of writing assessment is contingent upon the definition of writing ability, which varies across different

contexts and situations (Weigle, 2002). A singular definition is insufficient to encompass all writing abilities and purposes (Camp, 1993; White, 1995). For example, the skills required for composing a cause-and-effect essay differ markedly from those needed to write a recipe. Hence, the variety of writing genres, purposes, and contexts dictates the definition of writing ability within a specific environment, contributing to the complexity of the assessment task. Moreover, Hamp-Lyons (2019) highlighted that interpretations of fairness in assessment diverge among academics and practitioners.

Teachers' dilemmas concerning writing instruction and assessment are not different in EFL contexts. In the Kingdom of Saudi Arabia (KSA), where English is a part of the curriculum at all levels of education, writing instruction and assessment present multiple challenges, particularly in the realm of higher education (Alkhazim, 2003; Ashraf, 2018). English Language Teaching (ELT) in higher education is underscored by mandatory PYPs. Offered by most universities, a PYP is a foundational year that concentrates on English language study and research, alongside computer skills, and basic sciences (Borg & Alshumaimeri, 2019). It has been noted that university-level Saudi students face considerable difficulties with writing skills and are often perceived by EFL teachers as lacking proficiency in this area (Ansari, 2012; Mohammad & Hazarika, 2016). These students are frequently identified as lacking essential writing strategies and knowledge, vocabulary depth and breadth, spelling acumen, and the organizational skills required to produce competent written texts (Ahmed, 2018; Maher & Al-Khasawneh, 2010; Ezza, 2017; Ghalib & Al-Hattami, 2015). As a result, writing teachers are tasked with addressing students who resort to techniques such as memorizing paragraphs to pass exams, rather than striving to achieve mastery in writing (Mohammad & Hazarika, 2016).

From my own experience in teaching and assessing writing within a PYP at a Saudi University, it was apparent that there were ongoing concerns regarding the assessment of writing tasks and the standardization of evaluating students' written texts among writing teachers. These concerns span various aspects, including eliciting student writing performance, feedback practices, justification of grades, and interpretations of what characterizes proficient writing. This observation concurs with the findings of Obeid (2017), who reported that EFL writing teachers at Saudi universities face numerous assessment challenges. These include constraints on time available for assessment, students' perceptions of grading unfairness due to grade inconsistencies, exclusion of teachers from assessment design, and a lack of training in rubric application. Williams (2003) identified the issue of inconsistency in assessment as a significant stressor for writing teachers, impacting the equity of evaluations. Similarly, Smith (1992) asserted that an investigation into such inconsistencies in teachers' assessment practices should extend beyond achieving consensus on a grade to understanding the rationales underpinning teachers' evaluative decisions.

The longstanding recognition of teachers' pivotal role in making pedagogical decisions has yielded significant insights into the domain of education, particularly within language teaching. Numerous scholars acknowledge that the mental lives and prior knowledge of teachers are instrumental in providing nuanced and profound insights into classroom teaching and learning dynamics (Borg, 2006, 2019). The growth of this field is attributed to several factors; nevertheless, three central justifications recur: exploring beliefs facilitates a deeper understanding of teachers and enhances teaching quality, propels educational reform (Skott, 2014), and contributes constructively to supporting the professional development of teachers (Borg, 2018). Consequently, research has begun to delve into various facets of EFL teacher

cognition, including writing assessment (Wang et al., 2020), with particular attention to elements of writing assessment such as formative assessment (Guadu & Boersma, 2018; Lee, 2011), written corrective feedback (Li & Barnard, 2011; Min, 2013), peer feedback (Yu, 2013), assessment purposes (Brown, 2004), and grading practices (Kimberly, 2017).

The terms *teacher thinking*, *teacher cognition*, or *teacher beliefs* have been around for fifty years (e.g., Borg, 2006; Clark & Peterson, 1986; Clark & Yinger, 1977), with extensive efforts undertaken to investigate and define them. Recently, Borg (2019) defined teacher cognition research as follows:

Inquiry which seeks, with reference to [teachers'] personal, professional, social, cultural, and historical contexts, to understand teachers' minds and emotions and the role these play in the process of becoming, being and developing as a teacher. (p. 1167)

Acknowledging the cognitive side is crucial and advantageous to capture the entirety of the language teaching process (Burns et al., 2015) and the personal, professional, and contextual factors that affect the relationship between language teachers' beliefs and their practices.

Teachers' beliefs and actions dynamically interact within their specific contexts in a complex way, as suggested by several researchers (Borg, 2019; Burns et al., 2015; Feryok, 2010; Johnson, 2018; Kiss, 2012; Li, 2013; Zheng, 2015). The relationship between *teacher cognition*, "what teachers know, believe and think" (Borg, 2003, p. 81), and their classroom practices is a sophisticated nonlinear relationship where cognition is not only granted as a crucial element in forming classroom experiences but is also seen as a result of classroom experiences (Zheng, 2015).

The importance of teacher cognition and context in writing assessment has been a focal point for many scholars (Calfee & Miller, 2013; Crusan, 2010; Weigle, 2002). Crusan (2010)

elucidated that the concept of writing intertwines with the perceptions of quality writing, competent writers, and equitable writing assessments. This interrelation suggests that teachers' beliefs about writing assessment and writing abilities are pivotal in determining the aims and methods of assessment within a particular setting. Additionally, Alshakhi (2019) posited that without meaningful integration of teachers' beliefs and practices, writing assessments in Saudi higher education are ineffectual and fail to enhance student learning. Hence, the necessity for writing assessments to be contextually relevant is underscored. Lee (2018) emphasized the global concern regarding student writing standards, highlighting the urgency to identify effective instructional strategies for L2 writing. She asserted that a deeper understanding of writing teacher beliefs could significantly enhance the development of teacher preparation programs tailored to specific contexts. Therefore, Lee suggested that the suitability of writing assessment practices is contingent upon the variables of the teaching environment. The appropriateness of these practices hinges on the teachers' beliefs, practices, and the context in which notions of writing proficiency, assessment objectives, and fairness are defined and applied (Crusan, 2010; Hamp-Lyons, 2019). Consequently, any research pertaining to writing assessment, including inquiries into teachers' L2 writing assessment beliefs, must be situated within its contextual parameters.

Feryok (2010) highlighted that experimental research and reviews of teacher cognition reveal that the relationship between teachers' beliefs, their practices, and context is complex, dynamic, and changing that "they can be systems forming unified and cohesive personal or practical theories" (p. 272). Feryok used CT to understand the complexity of the relationship and concluded that CT is a unified theory to investigate teachers' belief systems and subsystems and their interaction with their practices and contexts. Recently, CT has been considered an ideal

approach to social sciences, including applied linguistics (de Bot et al., 2007; Larsen-Freeman & Cameron, 2008; Thelen & Smith, 1996; van Geert, 2007). It is a framework that helps investigate human cognition without separating it from its context and gives cognition a dynamically adapting role within specific contexts (Zheng, 2015). CT elucidates “the interconnectedness of the components of a system in producing the whole, and of the system and its context” (Larsen-Freeman & Cameron, 2008, pp. 38–9). Larsen-Freeman (2011) suggested that CT allows recapturing our objects of concern in a way that highlights how the interacting components of a complex system produce the system’s collective behavior and how such a system concurrently interacts with its external context. Therefore, I used CT as my study framework to address the complexity of the interaction and relationship between teachers’ writing assessment beliefs, practices, and contextual factors.

Moreover, to effectively address the role of contextual factors in affecting teachers’ writing assessment beliefs and practices, this study adopted Bronfenbrenner’s (2005) ecological system model which divides contexts into three levels: (a) microsystem representing teachers’ personal history and classroom context, (b) exosystem representing the institutional context (e.g., university program), and (c) macrosystem, representing the broader societal blueprint of a country and its educational cultures, ideologies, and attitudes.

1.1 Research Questions

This study of the interrelationships between EFL teachers’ writing assessment beliefs, writing assessment practices, and the ecological factors in a PYP at a Saudi University was guided by the following research questions:

1. What are EFL writing teachers’ beliefs regarding the assessment of L2 writing?
2. How do teachers’ writing assessment beliefs relate to their actual practices?

3. What are the interrelationships between EFL teachers' writing assessment beliefs, actual practices, and contextual factors?

A qualitative multiple-case study design was used to address the research questions by including five EFL teachers in a PYP at a Saudi University as participants. The data collection methods used to answer the research questions were as follows: initial interviews, classroom observations; Think-Aloud Protocols (TAP), follow-up interviews, and documents analysis (e.g., course syllabus, samples of students' written texts, assessment tools).

1.2 Significance of the Study

Teachers bring a wealth of individual experiences and beliefs to their writing assessment practices, contributing to notable variability in instructional delivery, student engagement, and evaluation methods. This diversity extends from the introduction of writing tasks to the nuanced interpretation of rating criteria, each influenced by unique contextual factors. By examining teachers' assessment beliefs, the study aims to uncover deeper insights into how EFL teachers perceive writing assessment, and, in turn, shape their evaluation strategies. Such understanding could inform the development of more tailored professional support, ultimately enhancing writing instruction, assessment, and student learning.

Borg (2019) highlighted the importance of expanding teacher cognition research, particularly within international contexts, due to its profound implications. In the context of Saudi higher education, existing literature has investigated various dimensions of EFL teachers' beliefs. Previous studies have addressed beliefs concerning self-regulated learning (Soliman & Alenazi, 2017) and reading strategy instruction (Bamanger & Gashan, 2014), among other areas. Similarly, research on writing assessment has delved into aspects such as marking criteria (Ezza, 2017) and the assessment challenges within PYP (Alshakhi, 2019). Nonetheless, studying Saudi

teachers' beliefs about writing assessments has not been extensively explored. While Obeid (2017) examined teachers' perceptions of writing assessment, the study did not investigate the belief-practice connection nor the historical and social contexts, relying primarily on semi-structured interviews for data collection. Thus, a gap persists in comprehending the relationship between EFL teachers' writing assessment beliefs, practices, and contextual factors within Saudi higher education.

Furthermore, to the best of my knowledge, no research has applied CT to unravel the intricate web of EFL teachers' belief systems related to writing assessment. This study endeavors to fill this gap by examining teachers' writing assessment beliefs, practices, and contextual factors and, importantly, how these components intersect. The findings aim to inform the development of efficacious assessment policies and instruments, as well as tailored teachers' educational and professional development initiatives and measures, catering specifically to the context of PYP in Saudi universities. Such measures are particularly vital as the KSA embarks on higher education reforms geared towards shifting from centralized education to enhanced institutional and teacher autonomy.

1.3 Overview of the Dissertation Chapters

Chapter 1 lays the foundation of the study, presenting the research aims, literature background, questions, and significance. Chapter 2 reviews relevant literature across four segments. The initial segment explores theories and approaches related to teachers' beliefs, defines key terms, and considers methodological challenges within the research on teachers' beliefs. The following sections offer insights into the interplay between language teachers' beliefs, practices, and contextual factors, followed by an examination of L2 writing assessment theories and paradigms. Subsequently, the discussion shifts to research concerning L2 teachers'

assessment beliefs and practices. The concluding segment delineates the theoretical underpinning of this study, CT, and reviews related scholarly works.

Chapter 3 introduces the Saudi educational landscape, with a focus on ELT at the university level, specifically within PYPs. It outlines L2 writing instruction and assessment practices in Saudi universities' PYPs and elucidates the context of this study, including details about the University and PYP administration structure, policies, and regulations. Moreover, it describes the specific writing course under study, English II, detailing its objectives, materials, assessment methods, tools, and the role of writing teachers within the PYP framework. Chapter 4 describes the qualitative research design, sampling procedures, and approaches to data collection and analysis. It also discusses the strategies employed for trustworthiness measures and the ethical considerations adhered to during the study.

Chapters 5 to 9 present the in-depth analysis of the qualitative data of each of the five case studies, encompassing initial interviews, observations, TAPs, follow-up interviews, and document analysis. Each chapter focuses on the teacher's core and peripheral beliefs, actual practices, and ecological factors, as well as their interactions.

Chapter 10 synthesizes the findings across the individual cases, engaging in comparative analysis and discussion of the collective data. Finally, Chapter 11 consolidates the primary findings derived from the qualitative data analysis, relating them to existing literature. It also addresses the study's limitations and presents implications for teachers, policymakers, and future research.

CHAPTER 2: LITERATURE REVIEW

2.1 Language Teachers' Cognition

In psychology, numerous studies on cognition have demonstrated the direct and substantial influence of beliefs and thoughts on human actions (Matlin, 2009). These studies have also provided valuable insights into the field of education, leading to increased interest in exploring the hidden aspects of teaching starting in the 1970s. Since then, there has been a growing recognition that teachers are human beings with rich mental lives, actively engaging in thinking and decision-making processes that significantly shape classroom behaviors (Clark & Yinger, 1979; Green, 1971).

This interest in understanding teacher cognition and beliefs has resulted in its acceptance as a crucial factor for comprehending and explaining the educational behaviors of teachers (Clark & Peterson, 1986; Kagan, 1992; Pajares, 1992). Early research on teacher cognition primarily focused on teacher decision-making (Shavelson et al., 1977, 1979) and teacher thinking (Clark & Yinger, 1977). However, in the 1980s and 1990s, there was a shift in focus towards investigating teachers' beliefs and teachers' knowledge (Freeman & Johnson, 1998; Kagan, 1992; Nespor, 1987; Pajares, 1992) within this line of research.

Eventually, teacher cognition research has infiltrated language teaching, with key publications on language teacher learning and cognition appearing in the mid-1990s (Freeman & Richards, 1996; Woods, 1996). Ever since, a substantial amount of research has been conducted on language teacher cognition (Borg, 2003, 2006; Couper, 2017; England, 2017; Wyatt & Ager, 2017). In language teacher cognition research, there is an emphasis on *teachers' beliefs*, defined broadly as anything that a teacher considers true (Murphy & Mason, 2006).

In early studies, research on language teacher beliefs focused on two important areas of language teaching: literacy and grammar instruction, with a greater emphasis on grammar than reading or writing (Borg, 2003, 2006). In the last ten years, language teacher cognition research has expanded to encompass other domains, including technology integration (e.g., Kim et al., 2013), learner autonomy (Borg & Alshumaimeri, 2019), the use of First Language (L1) in L2 classrooms (Miri et al., 2017), language-teaching methodology (Zhu & Shu, 2017), multilingualism (Haukås, 2016), interactive thinking (Jackson & Cho, 2018), writing assessment literacy (Crusan et al., 2016), and written feedback (Junqueira & Payant, 2015). The growing interest has extended to theoretical debates concerning the nature and future directions of the field (e.g., Borg, 2019; Burns et al., 2015; Kubanyiova & Feryok, 2015). Despite the debates and changes in theoretical perspectives over the years, teacher cognition research has maintained its core principle, highlighting that exploring the hidden dimension of teaching is crucial for understanding the process of becoming, being, and developing as teachers (Borg, 2019).

2.2 Emerging Themes from Language Teachers' Beliefs Research

The extensive study of the relationship between teachers' beliefs and practices has given rise to several issues over the years. Firstly, the in-depth exploration of language teacher cognition has led to the proliferation of numerous terms in the field, such as teacher thinking, knowledge, attitudes, cognition, and beliefs (Pajares, 1992; Borg, 2019). Secondly, cognition is an abstract entity influenced by researchers' theoretical paradigms, resulting in a wide range of methodological approaches employed to investigate this phenomenon (Borg, 2006, 2019). Lastly, teacher cognition is a dynamic entity influenced by various factors, highlighting its complex and context-dependent nature (Feryok, 2010; Zheng, 2015). These aspects of language teacher cognition are considered problematic and challenging by some scholars, as they

contribute to discrepancies in research findings (Borg, 2006, 2019). Therefore, it is crucial to carefully define concepts, select appropriate methodologies, and establish a solid theoretical framework to ensure valid and reliable findings in the specific context under investigation.

2.2.1 Definition of Terms

Borg (2019) emphasized that despite the substantial literature on teacher cognition, significant discrepancies in research findings may be attributed to conceptual challenges. The study of internal influences on teachers has been present for over 40 years, employing different terms. In early studies, *teacher thinking* emerged as the dominant concept (Clark & Peterson, 1986), while other concepts such as *decision-making* (Clark & Yinger, 1977), *teacher knowledge* (Freeman, 2002), and *teachers' attitudes and beliefs* were also present (Munby, 1984). The abundance of terms used in the field has caused confusion among new researchers (Borg, 2006), leading scholars to criticize it as a “messy construct” (Pajares, 1992, p. 307). Zheng (2015) attributed this complexity to the abstract, broad, and intricate nature of the concepts employed in the field. For instance, the terms *belief* and *attitudes* are used in various disciplines, including psychology and anthropology, resulting in different interpretations and connotations. In the realm of social psychology, Ajzen's (1988) theory of planned behavior distinguishes between beliefs and attitudes, highlighting how they collaboratively influence actions. Within this framework, beliefs are understood to be cognitive constructs, representing the knowledge an individual possesses regarding an innovation. Attitudes, contrastingly, are anchored in the situational context and are expressive of an individual's emotional and value-driven judgments (as cited in Lawrence, 2018).

In addition, Li (2017) highlighted that researchers often struggle to distinguish between belief and knowledge. There has been extensive discussion regarding the definitions of

knowledge and belief. Borg (2006) pointed out that since the emergence of knowledge as a focus of study in the 1980s, various classifications of knowledge have been proposed, including personal practical knowledge (Connelly & Clandinin, 1988), case knowledge (Shulman, 1986), and content knowledge (Grossman et al., 1989). Teacher knowledge has been categorized differently in many studies. For instance, Elbaz's (1983) case study of an English teacher identified five categories from interview data to describe teachers' practical knowledge: knowledge of self, the teaching milieu, subject matter, curriculum, and instruction. Another study by Shulman (1986) defined three types of content knowledge: subject matter knowledge, pedagogical knowledge, and curricular knowledge. These varying categorizations have contributed to significant confusion within the field.

Pajares (1992) observed that terms such as *beliefs, values, perceptions, conceptions, attitudes, judgments, implicit theories, personal theories, opinions, and ideologies* are often used interchangeably, making it challenging to differentiate between the characteristics of knowledge and beliefs. It was not until 2003 that Borg introduced the term *teacher cognition* as an umbrella term encompassing the unobservable aspects of teachers' work. This allowed for the incorporation of various constructs in the field, including beliefs, knowledge, and attitudes, each with multiple categories and classifications among researchers. According to Borg (2003), teacher cognition refers to "the unobservable cognitive dimension of teaching – what teachers know, believe, and think" (p. 81). Later, additional dimensions, such as the emotional aspect of teachers' experiences, were incorporated into the concept of teacher cognition too (Golombek & Doran, 2014).

Borg (2019) argued that the term teacher cognition has been crucial for the field as it provides a conceptual framework for an area of inquiry that was previously perceived as

scattered and lacking coherence. He proposed that using a superordinate term allows researchers to comprehensively grasp the complexity of teachers' mental lives rather than attempting to identify and categorize various aspects of teachers' minds, such as beliefs and knowledge. Borg justified this by noting that while beliefs and knowledge are well-established concepts, other concepts such as teacher motivation and teacher identity are not, and thus, they can all be encompassed under the umbrella term of teacher cognition, representing what teachers know, believe, think, and feel.

Borg (2019) further acknowledged that the inclusive nature of the term teacher cognition could present challenges, as it results in fuzzy boundaries, allowing for the inclusion of numerous non-behavioral aspects of teachers' lives without clear delineation. Burns et al. (2015) criticized the term cognition, as it still reflects a cognitive view of the mind that does not fully capture the advancements made in the study of psychology and human education since the 1970s, for example, the recognition of the vital role of affect in teachers' beliefs and practices.

In the field of assessment, the term *teacher conceptions* has been widely used by researchers (Azis, 2015; Brown, 2004, 2010; Davis & Neitzel, 2011; Opre, 2015; Thompson, 1992). Brown (2004) defined *conceptions* as "the organizing framework by which an individual understands, responds to, and interacts with a phenomenon" (p. 303). Brown's research focused on four conceptions of assessment, aiming to examine teachers' agreement or disagreement with them. These conceptions included the improvement of teaching and learning, school accountability, student accountability, and the perception of assessment as irrelevant. Thompson (1992) employed the term conceptions as a broader concept that encompasses "beliefs, meanings, concepts, propositions, rules, mental images, preferences, and the like" (p. 130).

In a recent study, Brown et al. (2019) argued that there is no globally homogeneous construct of teacher conceptions of assessment, as teachers' beliefs develop within the historical, cultural, social, and policy contexts in which they work. Consequently, teacher conceptions of assessment will always be influenced by local expressions. For their study, Brown et al. (2019) defined teachers' conceptions as "the cognitive beliefs about and affective attitudes toward assessment that teachers espouse, presumably in response to the policy and practice environments in which they work" (p. 2).

The existence of numerous terms underscores the lack of consensus on which term to use, highlighting the importance for researchers to select and define their terms in teacher cognition research carefully. Zheng (2015) emphasized the need for researchers to define their terms based on previous research definitions and the specific context of their study since abstract terms such as teacher beliefs are "study bound, culture based, and context specific" (p. 15). Li (2017) noted that research investigating different aspects of teacher cognition often employs terms such as teacher thinking, cognition, beliefs, and decision-making. Similarly, Borg (2019) noted that the terms teacher cognition (e.g., Lim, 2016; Miri et al., 2017; Ngo, 2018; Tajeddin & Aryaeian, 2017; Zhu & Shu, 2017) and teacher beliefs (e.g., Borg & Alshumaimeri, 2019; Wang et al., 2020) are the dominant terms used in recent studies.

In this thesis, I employ the terms *teacher cognition* and *teachers' beliefs* interchangeably, following their common usage in recent literature. I define these terms as the internal components of teachers' knowledge, beliefs, thoughts, decisions, and attitudes regarding L2 writing assessment within their personal context and the context of their workplace, PYP. Additionally, following Zheng's (2015) approach, I consider any beliefs that teachers report

about their practices to be a component of their belief system, termed *belief in practice*, that is distinct from their actual writing assessment practices.

2.2.2 Methodological Challenges

Borg (2019) argued that the discrepancies found in teacher cognition research could be attributed to the wide array of methodological options available for investigating teacher cognition. Therefore, a significant challenge for researchers lies in critically examining the various data collection strategies that can be employed to elicit teacher cognition.

Drawing from Borg's (2006) literature review, data collection strategies can be broadly grouped into four categories: self-reporting instruments (questionnaires, scenario-rating tasks, and tests), verbal commentaries (structured interviews, semi-structured interviews, scenario-based interviews, repertory grids, and TAPs), observations (structured and unstructured observations), and reflective writing (journal writing, biographical accounts, retrospective accounts, and concept maps).

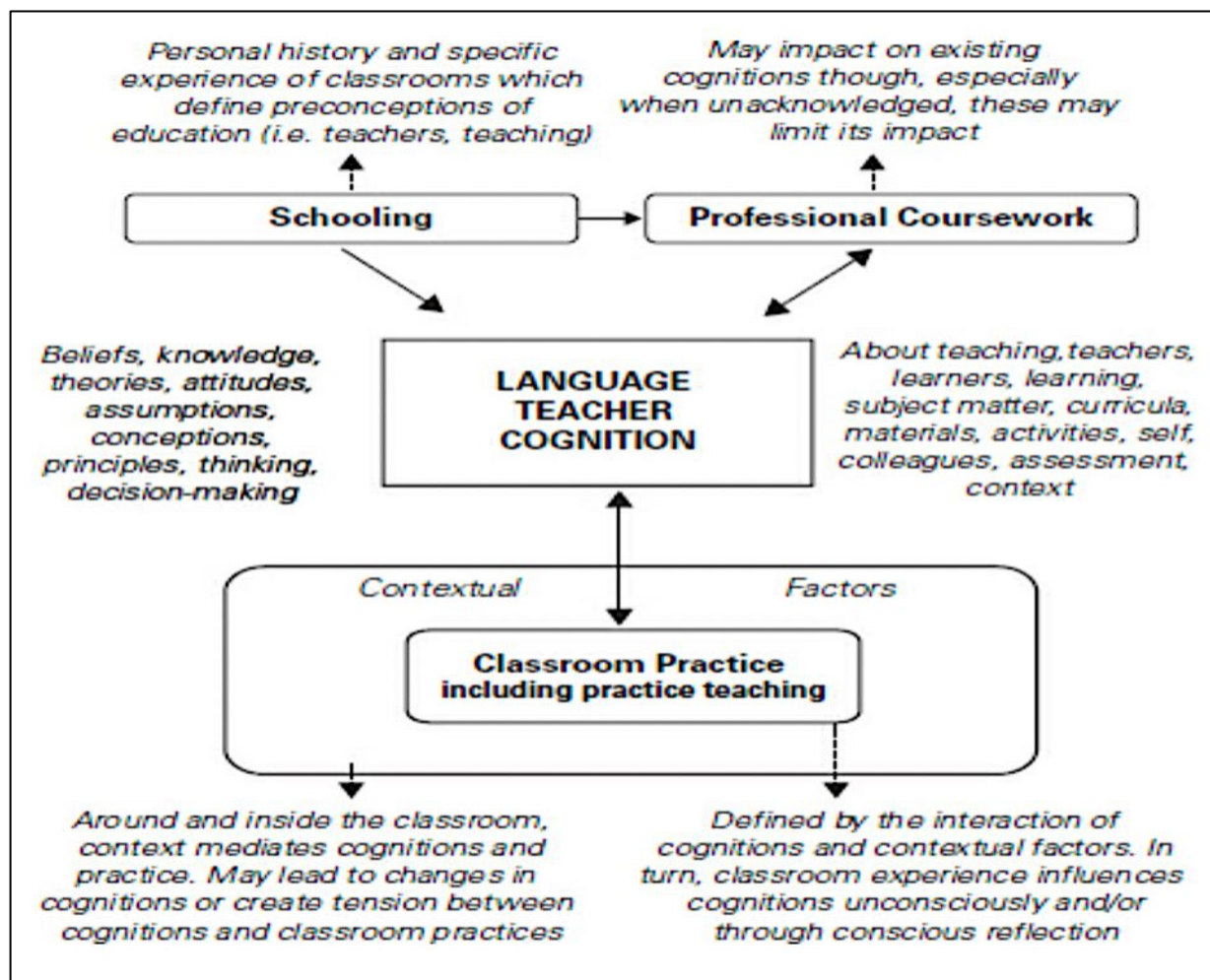
Given the wide range of methods available, including different types of questionnaires that measure various components of teacher cognition, Borg (2019) emphasized the importance of being aware and rational when selecting a data collection strategy or strategies. This is crucial to ensure valid findings and to provide a greater sense of coherence and common purpose in the study of language teacher cognition. Consequently, based on the purpose of my study, I chose to explore the interactions between teachers' writing assessment beliefs, practices, and ecological factors using initial interviews, observations, TAPs, follow-up interviews, and document analysis. The rationale for using each method is discussed in Chapter 4.

2.2.3 Factors Affecting Language Teachers' Beliefs

The studies mentioned earlier in this chapter demonstrate that teacher cognition is influenced by numerous factors, leading to potential discrepancies in research findings (Basturkmen, 2012; Borg, 2018; Brown et al., 2019; Li, 2013). To conceptualize the relationship between cognition and these essential factors in teachers' lives, Borg (2006) developed a schematic model that incorporates teachers' learning experiences, classroom practices, and the specific contexts in which they work. Figure 2.1 presents Borg's (2006) teacher cognition model, summarizing these key factors.

Figure 2.1

Elements and Processes in Language Teacher Cognition (Borg, 2006, p. 283)



The key elements of Borg's (2006) language teacher cognition model encompass teachers as learners and the role of context and classroom practices. Teachers, as learners, are shaped and influenced by their past personal, social, and historical experiences, which in turn impact their cognition and, subsequently, their classroom practices. Contextual factors also play a significant role in shaping and developing language teacher cognition. Borg demonstrated that classroom practices are a part of the broader context, encompassing teachers' interactions with factors both inside and outside the classroom. These contextual factors can consciously and unconsciously influence teacher cognition, often through reflection. Borg (2006) concluded that "teacher cognition and practices are mutually informing, with contextual factors playing an important role in mediating the extent to which teachers are able to implement instruction congruent with their cognition" (p. 284).

Borg's model highlights the significance of teachers' learning experiences and the influence of various contextual factors when studying language teachers' beliefs. In this study, the contextual factors are emphasized by employing Bronfenbrenner's (2005) ecological systems model, which situates a developing person within interactions across different environmental systems. Further details on this model are discussed later in this chapter.

2.3 Significance of Teachers' Cognition Research

Johnson (2006) noted that numerous factors and concepts had influenced the field of understanding L2 teachers' work, "but none is more crucial than the emergence of a substantial body of research, now referred to as teacher cognition" (p. 236). The significance of studying teacher cognition lies in its relevance to understanding various aspects of teaching, including pedagogical decisions, effective teaching, and teacher learning and development. Scholars asserted that teacher cognition has a substantial impact on teachers' judgments and perceptions

of learning and teaching interactions within pedagogical settings, leading to different practices and behaviors in the classroom (Borg, 2019; Li & Walsh, 2011). There is a reported connection between teachers' understanding of pedagogy and their instructional decisions in the classroom (Borg, 2003). Consequently, the importance of studying cognition has been emphasized for a considerable time, with Pajares's (1992) argument that research on teachers' beliefs has been effective in shedding light on how teachers approach and structure their tasks and challenges, serving as "stronger predictors of behavior" (p. 311).

Li (2017) explained the significance of studying teacher cognition based on two key reasons. Firstly, it allows scholars to investigate teaching and learning from the perspective of teachers, providing insights into their understanding of what constitutes effective instruction. Secondly, it enables a deeper exploration of how various aspects of teaching and learning work and interact, such as the selection of materials, task design, classroom interaction, and assessment decisions. By studying teacher cognition, researchers gain valuable insights into the complex nature of teaching and learning processes.

Breen et al. (2001) emphasized that the effectiveness of teaching is closely linked to the use of appropriate methodologies. They argued that researching teacher cognition could lead to developing new language pedagogical frameworks that are directly derived from classroom experiences in diverse teaching contexts. This approach can generate more practical and context-dependent teaching methodologies, offering grounded alternatives to existing language teaching methods. This is particularly relevant in situations where the implementation of a new curriculum or technology is planned.

Speer (2005) suggested that studying teacher cognition can enhance teaching effectiveness in two ways. Firstly, it provides a deeper understanding of the feasibility, potential

obstacles, and possibilities for innovation. Secondly, teachers can develop pedagogies, guidelines, or strategies tailored to specific contexts, thereby facilitating the implementation of innovative approaches. The underlying idea is that teacher cognition is a key factor that influences teachers' decisions about the relevance of knowledge, appropriate teaching routines, desired goals, and important aspects of the social context within the classroom (Speer, 2005).

Research on teacher cognition has significant implications for teacher education and development. Breen et al. (2001) emphasized that teacher cognition research has been a central focus in L2 teacher education, serving as a valuable resource for initial teacher education and fostering reflection in in-service teacher development. Many scholars have explored the impact of pre-service teacher education on teacher cognition (e.g., Borg, 2004, 2005; Peacock, 2001; Urmston, 2003). Evidence of changes in teacher cognition following teacher education courses has been reported in studies investigating the influence of teacher education on pre-service teacher cognition (e.g., Busch, 2010; Mattheoudakis, 2007). Similarly, studies focusing on in-service language teachers have also identified shifts in teacher cognition (e.g., Borg, 2011; Lamie, 2004; Scott & Rodgers, 1995). For instance, Scott and Rodgers (1995) found that teachers' beliefs about writing were highly aligned (89%) with the principles learned in their language teacher education course. Similarly, Borg's (2011) qualitative longitudinal study concluded that an eight-week educational course for six in-service English teachers critically transformed their beliefs. The course facilitated changes in their prior beliefs about various aspects of language teaching and learning, increasing their awareness of classroom practices through reflective processes. Furthermore, other studies have demonstrated that changes in language teacher beliefs can vary depending on the individual and the specific area of belief (e.g., Liu & Fisher, 2006; Murray, 2003). The link between teacher cognition and teacher

education has been a longstanding focus, and language teacher cognition research holds significant implications for the development of teacher education courses and programs (Borg, 2011). Kagan (1992) highlighted that teacher cognition “may be the clearest measure of a teacher’s professional growth” (p. 85).

2.3.1 Relationship Between Teachers’ Beliefs and Practices

The importance of research on teachers’ beliefs has emerged from the recognition of the potentially significant relationship between teacher beliefs and practices (Calderhead, 1996; Hoy et al., 2006; Pajares, 1992; Richardson, 1996). The question of how language teachers’ beliefs influence their practices has prompted extensive research in the field of teacher cognition, specifically focusing on the connection between teachers’ beliefs and their classroom practices (e.g., Borg, 2018; Cundale, 2001; Flores, 2001). Buehl and Beck (2014) have presented four perspectives on the beliefs-practices relationship in teacher cognition research: (a) teachers’ beliefs affect their practices, (b) teachers’ practices affect their beliefs, (c) teachers’ beliefs are disconnected from their practices, and (d) the relationship between teachers’ beliefs and practices is reciprocal and complex.

First, there has been considerable research conducted over the past three decades that demonstrates how teachers’ beliefs influence their instructional practices (Ng & Farrell, 2003), the decisions they make when teaching (Tillema, 2000), and the adoption of new teaching methods, strategies, and pedagogical techniques (Donaghue, 2003). Ng and Farrell (2003) found evidence that teachers’ pedagogical choices and practices in grammar classrooms are shaped by their beliefs. Fives and Buehl (2012) remarked on the different types of teachers’ beliefs and how they affect their teaching. They illuminated that teachers’ beliefs act as “filters for interpretation, frames for defining problems, and guides or standards for action” (p. 79). Moreover, they noted

that a strong connection between teachers' beliefs and practices was crucial for teaching satisfaction and good practice for both pre-service and in-service teachers. Regarding assessment practices, researchers found that teachers' beliefs have a clear influence on the implementation of curriculum and feedback, the conduct of assessments, and the monitoring of students' performance and development in language classrooms (Brown, 2004; Brown et al., 2012; Opre, 2015; Rogers et al., 2007; Skott, 2014; Thoonen et al., 2011).

For example, Rogers et al. (2007) found that the similarities between teachers' beliefs and their assessment and evaluation practices were greater than the differences in three ELT contexts: Canada, Hong Kong, and Beijing. Using 32 questionnaire items, they found that teachers' beliefs had diverse effects on how teachers evaluated students in the classroom, including influences on their planning, implementation, and choice of student assessments. Similarly, Brown (2004) conducted a study in New Zealand examining teachers' agreement or disagreement with four assessment purposes: improving teaching and learning, school accountability, student accountability, and irrelevance to the teaching and learning processes. The findings indicated that participants generally agreed that assessment plays a positive role in improving students' learning outcomes, enhancing the quality of teaching, and promoting accountability at the school level. However, they expressed disagreement with the notion that assessment is irrelevant to the teaching and learning processes.

The second perspective focuses on the notion that teachers' beliefs are formed and developed after engaging in specific teaching actions or practices. This perspective is evident in studies that investigated the impact of professional development on teachers' beliefs and the role of experiences on pre-service teachers' beliefs. This view examines the changes in pre-service teachers' beliefs following their participation in professional development courses or learning

experiences (Borg, 2015). For instance, teachers' learning and teaching beliefs can be influenced by their positive or negative experiences as learners (Holt-Reynolds, 1992). Although Buehl and Beck (2014) identified this view, it is not explicitly stated in the mentioned studies. The effect of teachers' learning experiences is primarily considered as one of the factors influencing the development of their beliefs. Borg (2018) further explained that personal and contextual factors also play a role in shaping teachers' beliefs.

The third perspective focuses on understanding why teachers' beliefs may not always be reflected in their practices. Some research has found inconsistencies in the belief-practice relationship among teachers (Davis & Neitzel, 2011; Lim & Chai, 2008; Lee, 2009; Jorgensen et al., 2010). For instance, Lee (2009) reported discrepancies between writing teachers' beliefs about feedback and their actual practices due to contextual constraints. Similarly, Davis and Neitzel (2011) found that teachers' assessment environments did not align with their understanding of classroom assessments, resulting in inconsistencies between their beliefs and practices. Researchers have attempted to explain the reasons behind these misalignments in the belief-practice relationship using various theoretical frameworks, such as Sociocultural Theory (SCT) (Rashidi & Moghadam, 2015) and CT (Zheng, 2015).

The second and third perspectives highlight that teachers' beliefs and practices are influenced by multiple factors (Borg, 2018; Phipps & Borg, 2009), indicating a more complex relationship between them than a simple cause-and-effect relationship. As a result, the prevailing view is that the relationship between teachers' beliefs and practices is much more intricate than previously assumed (Basturkmen, 2012; Li, 2013). This complexity arises from the recognition that teachers' beliefs and practices have a reciprocal relationship. Borg (2018) argued that the relationship between teachers' beliefs and practices is non-linear. Teachers hold diverse beliefs

that interact with one another and are influenced by external factors such as policies and curricula, which in turn shape their actions. For instance, Alzaanin (2019) utilized data from semi-structured interviews, classroom observations, stimulated recall interviews, and course documents to investigate the intricate interplay between the beliefs and practices of eight EFL writing teachers at two Palestinian universities. The study also examined how teachers' beliefs and practices interacted with their ecological contexts. The in-depth analysis revealed that teachers' beliefs about L2 writing, teaching and learning L2 writing, and their professional roles influenced their instructional methods, curriculum design, and classroom assessment approaches. The findings also indicated that the physical constraints of the classroom, such as large class sizes, limited resources, and spatial restrictions, posed significant challenges to writing instruction. However, Alzaanin argued that teachers' beliefs about their professional roles could mitigate the impact of these ecological constraints.

To explore the impact of these factors, Basturkmen (2012) conducted a literature review on the relationship between teachers' beliefs and practices and found that contextual factors significantly influenced teachers' practices. Moreover, Basturkmen noted that predicting teachers' practices solely based on their beliefs is challenging, as teachers' beliefs do not always align with their instructional practices, and teachers with similar beliefs may exhibit different practices. This relationship was described as "tenuous" (p. 243). For instance, Gao and Liu (2013) surveyed 325 Chinese college EFL teachers and conducted a case study involving four teachers. They reported that the relationship between teachers' beliefs and practices could be both consistent and inconsistent, influenced by various factors indicating complex connections. This phenomenon arose from the existence of multiple belief systems within teacher cognition, where beliefs in one system might conflict with beliefs in another system (Basturkmen, 2012).

Additionally, Phipps and Borg (2009) explained that incomplete correspondences between beliefs and practices can be attributed to the “tension” between core and peripheral beliefs (p. 380). In other words, teachers’ beliefs form a system where certain beliefs are considered core, while others are peripheral (Green, 1971; Pajares, 1992). Core beliefs are more stable and influence behavior more strongly than peripheral beliefs. Phipps and Borg (2009) suggested that exploring discrepancies in the belief-practice relationship can be enhanced by focusing on the differences between these belief subsystems. The review by Buehl and Beck (2014), which presents the four perspectives on the belief-practice relationship, highlights the alignment between adopting CT in teacher cognition research and recent views in the field.

2.4 L2 Writing Assessment

Assessment in teaching is a broad term encompassing various activities undertaken by teachers to gather and interpret evidence regarding their students’ progress, learning needs, and abilities. Brown (2004) defined *assessment* as “any act of interpreting information about student performance, collected through any of a multitude of means or practices” (p. 304). Assessment is considered a fundamental component of the teaching and learning process, playing a crucial role in enhancing the quality of education (Crusan, 2010). Teachers continuously assess their students’ strengths and weaknesses in terms of abilities and knowledge (Hyland, 2003; Taras, 2005), enabling them to determine the extent, nature, and quality of student learning and providing insights for program and institutional accountability (Taras, 2005). Writing assessments, in particular, offer students an opportunity to reflect upon and improve their weaknesses in writing, shaping their curriculum and academic achievements (Crusan, 2010; Hyland, 2003). The significance of assessment extends beyond teachers and students, impacting school administrators and stakeholders involved in language learning (Cumming, 2009).

Weigle (2002) highlighted the importance of defining writing assessment in relation to the specific understanding of writing ability and purpose. Given that different contexts and conditions necessitate distinct writing abilities and purposes, it is not possible to have a single definition of writing assessment that encompasses them all (Weigle, 2002; White, 1995). Writing assessment is inherently connected to the learning process, as any modifications in assessment methods have a direct impact on students' learning and teachers' pedagogical approaches (Weigle, 2002).

Researchers have identified three main purposes of assessment: improving teaching and learning, ensuring student accountability for learning, and ensuring school and teacher accountability (Brown, 2004; Nisbet & Warren, 1999). Hyland (2004) summarized five major purposes for assessing students' writing. These include placement, which provides information to allocate students to appropriate classes; diagnostic, which identifies students' writing strengths and weaknesses and helps determine areas where remedial action is needed; achievement, which enables learners to demonstrate the progress they have made in a course; performance, which assesses students' ability to perform specific writing tasks associated with real-life academic or workplace requirements; and proficiency, which assesses students' general level of competence, often for purposes of certification for employment or university study. In most writing classrooms, assessment primarily focuses on diagnostics and/or achievement, although other purposes may also be included (Hyland, 2004).

2.5 A Paradigm Shift in L2 Writing Assessment

The history of writing assessment is closely intertwined with the history of measurement, writing instruction, and composition theories. Furthermore, the history of L2 writing and writing assessment is deeply connected to the history of their L1 counterparts (Behizadeh & Engelhard,

2011; Crusan, 2010; Matsuda, 2003). By tracing this history, the influence of dominant L2 learning theoretical frameworks, such as behaviorism, cognitivism, constructivism, and SCT, can be observed in the forms and focus of writing assessment (Crusan, 2010; Elliot, 2005; Yancey, 1999).

Many scholars have emphasized that theories of L2 learning are not developed in isolation. Instead, they constantly interact and negotiate with one another, which is reflected in the emergence and evolution of writing assessment theories (Crusan, 2010). These theories have been shaped by various influences and forces, as discussed by Behizadeh and Engelhard (2011), Crusan (2010), Hamp-Lyons (2001), Weigle (2002), and Yancey (1999). Consequently, different writing assessment methods based on diverse theoretical perspectives have emerged and faded over time, only to resurface again (Crusan, 2010). Yancey (1999) aptly described this historical process using the metaphor of “waves” in the field of writing assessment. These waves represent overlapping periods of assessment methods, where each wave builds upon the previous ones without completely replacing them (p. 131). Yancey identified three waves in the history of writing assessment, characterized by the dominant assessment methods used to measure writing competence during each respective period. The three waves of writing assessment are as follows:

During the first wave (1950-1970), writing assessment took the form of objective tests; during the second (1970-1986), it took the form of the holistically scored essay; and during the current wave, the third (1986-present), it has taken the form of portfolio assessment and of programmatic assessment. This is the common history of writing assessment: the one located in method. (p. 484)

Similarly, Hamp-Lyons (2001) provided a framework for understanding the history of writing assessment approaches by categorizing them into four generations based on the

assessment methods employed. These generations include the direct approach using essay tests, the indirect approach using multiple-choice tests, portfolio-based assessment, and the emerging fourth generation characterized by “humanistic and technological” assessment methods (Hamp-Lyons, 2001, p. 117).

2.5.1 Testing Culture Vs. Assessment Culture

Yancy’s (1999) waves illustrate the paradigm shift in the assessment along a continuum spanning from what can be called a *testing culture* to an *assessment culture* (Gipps, 2012). The testing culture prioritizes traditional assessment concepts focusing on standardized testing procedures, while the assessment culture focuses on alternative approaches to assessment as carried out mostly by teachers (Gipps, 2012). Gipps (2012) explained that the shift between the two cultures can be identified in many ways. First, there is a shift in the purpose of assessment from Assessment of Learning (AoL) to Assessment for Learning (AfL), aligned with the summative and formative assessment concepts. AoL focuses on summative assessment, which emphasizes the evaluation of students’ ability in a standardized test environment to gain information about students’ learning outcomes (Gipps, 2008; Harlen, 2005). AfL is based on formative assessment to identify learners’ developing abilities and then attempt to predict what learners will do independently in the future (Taras, 2005).

Second, the shift can be noticed in the differentiation between norm-referenced assessments and criterion-referenced assessments. *Norm-referenced assessments* are emphasized in a testing culture in which students’ abilities are compared to the abilities of other students. In contrast, *criterion-referenced assessments* are present more in an assessment culture, emphasizing the student’s individual learning in relation to specific learning goals (Isaacs et al., 2013). The third shift concerns the relationship between instruction and assessment. In a testing

culture, instruction and assessment are considered separate activities, in which teachers are responsible for class instructions and stakeholders for developing assessments. As such, the *validity* of tests, defined as whether a test accurately measures what it is supposed to measure (Fulcher & Davidson, 2007), is affected because the assessments and tests do not reflect classroom instruction. In contrast, an assessment culture advocates the integration of instruction and assessment in the classroom, where teachers are responsible for both. Finally, in a testing culture, measurement theories have the most significant influence on writing assessment practices, while writing theories have less impact and vice versa (Behizadeh & Engelhard, 2011).

2.6 Language Teachers' Assessment Beliefs and Practices

Rea-Dickins (2004) stated that teachers are “agents of assessment” (p. 249), making choices about their instruction and assessment depending on their knowledge, experience, and intuition. Thus, teacher cognition and writing assessment researchers began to consider the value of exploring teachers' writing assessment beliefs. The concern originates from the awareness that teachers' conceptualization of assessment's purposes and functions profoundly affects how they implement it in their classroom practice (Brown et al., 2019; Yin, 2010). Conversely, it is the need to understand the relationship among beliefs, practices, and context that impels these investigations to exist, specifically when writing assessment is a context-dependent practice (Brown et al., 2019; Weigle, 2002; Crusan, 2010). Therefore, understanding teachers' writing assessment beliefs has great implications for policy and professional development (Brown, 2010; Brown et al., 2019).

Teachers possess varying beliefs about assessment and writing assessment. They have different opinions about the purpose and utilization of assessment (Davis & Neitzel, 2011; Harris & Brown, 2009). For instance, Davis and Neitzel (2011) found that 15 upper-elementary and

middle school teachers had diverse beliefs and understandings of the forms and functions of classroom assessment; however, those understandings did not endorse students' self-regulated learning. Through analyzing semi-structured interviews and observational data, they determined that teachers showed varied beliefs and knowledge of assessment forms (e.g., assessing through tangible products or interaction) and assessment functions for various audiences (e.g., to inform teachers, hold students accountable for their work, and inform extranet audiences such as parents or state-level administrators). Nonetheless, those beliefs did not stimulate the development of students' self-regulation, as teachers also held that they were the controllers and initiators of the assessment forms. They concluded that teachers' assessment beliefs generally affected assessment practices; they could also support or hinder specific instructional practices, such as self-regulated learning.

Brown's (2004) findings are in line with Davis and Neitzel's, in which 525 schoolteachers held different beliefs about the functions or purposes of assessment. However, the degree to which teachers had those beliefs also differed. Specifically, teachers agreed that assessment enhances students' learning and the quality of teaching, and teachers or schools were made accountable through assessment. Conversely, they did not see students' accountability as crucial as teachers' or schools' accountability. Harris and Brown (2009) examined 26 New Zealand teachers' beliefs about assessment for seven purposes, such as teacher use for individualizing learning, organizing group instruction, and externally motivating students. Using a phenomenographic approach, they concluded that teachers have complex beliefs about assessment for various purposes.

Moreover, teachers' writing assessment belief systems may include beliefs about varied areas of assessment, such as feedback (Al-Bakri, 2016; Brown et al., 2012; Lee, 2009; Mori,

2011), rating (Humphry & Heldsinger, 2019; Skar & Jølle, 2017), and a specific type of assessment, such as formative assessment (Guadu & Boersma, 2018). For instance, Crusan et al. (2016) used quantitative data to explore writing teachers' beliefs about writing assessment methods, scoring accuracy, and general writing assessment issues in classes. After surveying 702 teachers from 41 different countries with different writing teaching experiences and qualifications, they found that teachers usually believed in using a wide range of writing tasks, such as out-of-class writing assignments, portfolios, and timed in-class assignments. Furthermore, 53% of the teachers believed that writing assessment was exciting and challenging, whereas 33% believed it was an integral part of their job. As to scoring beliefs, participants had mixed feelings. For instance, more than half agreed that scoring of students' writing was always inaccurate, while, concurrently, 60% disagreed that scoring was subjective. Moreover, half of the participants perceived rubrics as a helpful tool for students to understand why specific grades were assigned, while the other half opined that students did not care about rubrics. Writing assessment issues among teachers were evaluated as challenging, as depicted in difficulty working with a colleague during scoring and achieving inter-rater reliability. Even though Crusan et al. (2016) used only quantitative data that may not explain differences at the individual level, their data showed that teacher cognition usually involves multifaceted beliefs, with different degrees, about varied aspects of writing assessment.

The studies mentioned above did not examine the relationship between teachers' writing assessment beliefs and practices, but other studies did (Guadu & Boersma, 2018; Montgomery & Baker, 2007; Min, 2003; Wang et al., 2020). The belief-practice relationship of teachers' writing assessment also reveals consistency and inconsistency. Using an explanatory mixed-method design, Guadu and Boersma (2018) discovered that 25 EFL writing teachers believed in the

value of formative assessment in enhancing students' learning and gaining insights into teaching and learning. However, those beliefs were not evident in their practices of monitoring students' progress and providing scaffolding. Such mismatch was also observed in Montgomery and Baker's (2007) study of teachers' written feedback, indicating that teachers' training may influence teachers' beliefs but not their practices.

Contrarily, Min's (2013) longitudinal self-case-study tackled changes in written feedback practices attributable to evolving knowledge. Min decided to incorporate peer review training into her writing course curriculum. As a result of consulting literature on peer review training, there was an observable shift in her approach to written feedback, stemming from revised knowledge and beliefs. For instance, Min elaborated that prior to the course, she preferred written feedback that focused on identifying problems and offering specific suggestions over posing questions to understand students' intentions. Min also concluded that the discussion of misalignment between teachers' beliefs and practices in assessment research are not widespread.

Wang et al. (2020) found both belief-practice consistencies and inconsistencies in a much greater group of teachers although inconsistencies were more salient. They studied 136 EFL teachers' beliefs about classroom writing assessment at eight Chinese universities. They examined if teachers' beliefs conform to their self-reported practices and what factors influence writing assessment beliefs and practices. First, using survey data, they found that teachers believed in AfL (i.e., making learning explicit and encouraging learners' autonomy) more than AoL (i.e., performance-based assessment) in supplying opportunities for students to learn writing. Nonetheless, teachers' self-reported assessment practices were not consistent with their beliefs. Teachers regularly practiced AoL, centering on rating performances and error identification, more than on AfL approaches, such as descriptive feedback or giving students

more responsibility in assessment. The second part of their study utilized qualitative data from semi-structured interviews with ten teachers, analyzing the effect of contextual factors (micro, meso, and macro levels) on the belief-practice mismatch. At the micro-level, they discovered that teachers' inadequate language assessment training and beliefs about students' English proficiency and motivation prevented them from implementing AfL practices. For instance, teachers believed in AfL, but they were perplexed about applying it due to a lack of assessment training. At the meso level, institutional factors, such as syllabus, assessment policy, workload, and large class size, affected teachers' implementation of AfL. Specifically, the school assessment policy focused on the high quantity of students' writing products and summative scores rather than writing processes and formative feedback. Therefore, teachers seesaw between exercising their beliefs of AfL or getting the job done on time for school administration. At the macro level, the testing culture in China valued high-stakes tests, leading to teachers' overuse of scores to evaluate students' writing despite teachers' preference for comment-only feedback.

The notable influences of contextual factors uphold the notion that the relationships between teachers' writing assessment beliefs and practices are complex. Many literature reviews have addressed teachers' assessment beliefs (Barnes et al., 2015; Bonner, 2016; Brown, 2016; Fulmer et al., 2015). They unequivocally showed that teachers understand and respond to the tension between using assessment to enhance learning and teaching and hold teachers and schools accountable for students' outcomes by the state administration. Therefore, when teachers are under increasing pressure to focus on scores, they are less likely to implement formative assessment, employing the alternative contemporary practices of assessment (Harris & Brown, 2009). Davis and Neitzel (2011) stated that "external audiences exert tremendous influence on the classroom assessment environment" (p. 210).

Lam (2019) used a mixed-method design, including questionnaires, telephone interviews, and classroom observations to investigate what 66 teachers know and think about classroom-based writing assessment and how they apply it. Lam concluded that teachers had the knowledge to distinguish between AoL and AfL, but not Assessment as Learning (AaL). Lam also acknowledged the value of AfL in enhancing learning more than AoL. Nonetheless, the findings showed that the participants attempted to use alternative writing assessments (i.e., peer assessment and post-writing consolidation activities) but experienced institutional barriers. Lam (2019) reported that teachers' assessment practices were "contextually mediated by meso-level constraining factors, including a lack of collaborative work culture, school support, autonomy, and space for professional development" (p. 85). Harris and Brown (2009) suggested that there was a high tension between "what teachers feel is best for students versus what is deemed necessary for school accountability" (p. 365). Both Lam (2019) and Harris and Brown (2009) recommended that teachers should select assessments by balancing the needs of society and school assessment culture and the students' needs to avoid tension. Furthermore, institutions should provide space, professional support, and autonomy for teachers to implement their alternative writing assessment beliefs.

2.7 Theoretical Framework: Complexity Theory (CT)

According to Feryok (2010), CT is rooted in other associated theories, including chaos theory and complex dynamic systems theory, which emanated from natural science in the mid-20th century and developed from various disciplines such as mathematics, biology, physics, and cybernetics. In the human sciences, these theories have been mainly applied to investigate the development of cognitive abilities and skills among children (e.g., Smith & Thelen, 1993; Thelen & Smith, 1996). In applied linguistics research, they have been used most notably in language

development (van Geert, 2007), L2 acquisition (de Bot et al., 2007; Larsen-Freeman, 2002, 2011), multilingualism (Herdina & Jessner, 2002; Jessner, 2008), teacher education (Davis & Sumara, 1997; Martin & Dismuke, 2018; Cochran-Smith et al., 2014; Opfer & Pedder, 2011), and teacher cognition (Finch, 2010; Gao & Zhou, 2021; Kiss, 2012; Zheng, 2015). Using CT as a theoretical framework in applied linguistics is based on the belief that cognitive systems should be viewed as wholes and examined dynamically in relation to how they interact with internal and external factors (de Bot, 2008; Larsen-Freeman, 2012). In addition, the use of CT in applied linguistics has originated from a perspective that emphasizes the following concepts: context, interactions, reciprocity relationships rather than causal ones, using system analysis units that explain interactions rather than dualistic analysis (Larsen-Freeman & Cameron, 2008).

The term *complexity* does not mean *complicated*; this term has a different meaning when used within CT (Larsen-Freeman, 2013, p. 372). *Complexity* is defined by Waldrop (1992) as “a chaos of behaviors in which the components of the system never quite lock into place, yet never quite dissolve into turbulence either” (p. 293). de Bot et al. (2007) suggested that the terms *chaos* and *complexity* refer to “unpredictability” rather than “randomness” (p. 4). Oxford University Press (n. d.) defines a *system* as “a set of things working together as parts of a mechanism or an interconnecting network.” Cameron (1995) explained that a *system* is defined as a means of describing “a conceptual holism.” (p. 31). A *complex system* is defined by Mitchell (2009) as “a system in which large networks of components with no central control and simple rules of operation give rise to complex collective behavior, sophisticated information processing, and adaptation via learning or evolution” (p. 13). In other words, it is composed of various interconnected and interrelated elements or agents that interact in different, nonlinear, and unpredictable ways. Larsen-Freeman and Cameron (2008) reported that these elements,

themselves, are complex systems containing other systems. Thus, CT, as a theoretical framework, seeks to understand complex systems and to explain how the interacting components of a complex system contribute to its collective behavior and how such a system concurrently interacts with its context. Within the applied linguistics literature, many researchers have analyzed complex systems and described the key properties of those systems possessing all or some of the following characteristics: (a) complex, heterogeneous, and dynamic, (b) sensitive to initial conditions, (c) have a nested structure, (d) open and co-adaptive to influences from contexts and other subsystems, and (e) self-organizing and emergent (Davis & Sumara, 2007; Gleick, 1987; Larsen-Freeman & Cameron, 2008).

First, CT perceives systems as complex and dynamic, meaning that they consist of different agents or elements; they are systems within systems with the same level of complexity that evolve, develop, and change with time (de Bot, 2008; Larsen-Freeman & Cameron, 2008). The complexity of such systems is attributed to their sensitivity to initial conditions, meaning that a minor change in the early state causes a dramatic and unpredictable change to other elements in the system; thus, resulting in different outcomes, hinting that the change is nonlinear (de Bot, 2008; Larsen-Freeman & Cameron, 2008). van Orden (2002) suggested that such nonlinearity can be observed in cognitive systems when a cognitive agent interacts with elements in the environment; thus, they change, and so do their interaction.

Heterogeneity in CT means that “the elements, agents, and/or processes in a complex system are of many different types” (Larsen-Freeman & Cameron 2008, p. 28). Zheng (2015) found that EFL teacher cognition was a whole complex system that comprises other complex belief systems (e.g., different belief systems about language teaching, learning, learners, and self as a teacher) and engages in various nonlinear and heterogeneous interactions.

Second, complex systems are contextualized and co-adaptive. Larsen-Freeman (2011) explained that complex systems are interrelated with other systems, responding to and influencing changes within them, and adapting to new environmental conditions to produce new behaviors that also affect the environment. Larsen-Freeman and Cameron (2008) described *co-adaptation* in complex systems as the process by which a system modifies itself in reaction to changes in its context; such systems are called “complex adaptive systems” (p. 33). A Complex Adaptive System (CAS) lacks centralized control and does not adhere to fixed or permanent states; thus, it adapts to new emergent states (Larsen-Freeman, 2014). This has been evident in teacher cognition studies as teachers’ belief systems are highly contextualized, and their formation is related to teachers’ historical and social contexts (Feryok, 2010; Zheng, 2015).

Third, CASs are open, emergent and self-organizing. CT is a theory “of a process rather than state, of becoming rather than being” (Gleick, 1987, p. 5). This refers to the openness of CASs, suggesting that they are open to interacting with other systems in varied ways (Larsen-Freeman & Cameron, 2008). As they develop, order emerges by interacting with different elements in their environment, such as knowledge, energy, and other systems (Larsen-Freeman, 2014). Although they are open to being “fed by energy coming into the system,” they can create order and stay there (p. 31). Their interactions seem random, but they lead to self-organization and the emergence of new elements, agents, or systems at multiple levels and timeframes.

The concept of *emergence* (Holland, 1998) is highlighted in CT as “the spontaneous occurrence of something new” (van Geert, 2008, p. 182). Larsen-Freeman (2014) clarified that a system exhibits emergence when new systems, behaviors, patterns, and structures dynamically and orderly arise from the interaction of the elements of a system, such as when a flock of birds emerges through the interaction of its individual members. Larsen-Freeman further elucidated

that it is hard to fully comprehend a bird flock phenomenon from studying a single bird; thus, CT allows examining how “complex order emerges” from the interacting elements, which would offer a more prominent and deeper frame for the phenomena (p. 228).

The complex order emerges from the self-organization nature of complex systems, which forms orderly patterns in all systems, whether in the natural world or human minds, cultures, and societies (Heylighen, 2009). Mitchell (2003) referred to *self-organization* in CASs as “any set of processes in which order emerges from the interaction of the component of a system without direction from external factors and without a plan of the order embedded in an individual component” (p. 6). They are systems that appear to organize and generate themselves without external direction, manipulation, or control.

Finally, attractors and basins of attraction are concepts used to describe the interactional behaviors within CASs. Larsen-Freeman (2014) noted that CASs are inherently chaotic and move through a sequence of modes of behaviors or states. Some of these states are stable, in which the system shows the same behavioral patterns for a while. The stable behaviors are called *attractors*, “islands of stability in a sea of chaos” (Lucas, 2004, para.1). In other words, under the influence of their self-organizing nature, the systems usually remain within one of a range of possible, stable behaviors. They tend to “settle into one or a few modes of behavior” that the system favors as opposed to all other possible modes (Thelen & Smith, 1996, p. 56). Such systems would reach a specific attractor based on its initial states; and *a basin of attraction* represents “the set of initial states” leading to each attractor. Basin of attraction refers to the state of being before a mode of behavior is locked into a stable state for a while (Lucas, 2004). A basin of attraction in a system is the space that can be regarded as the leading factor for an

attractor, so an attractor can be effective within a specific area of space, and there can be many various attractors adjoining each other (van der Leeuw & Folke, 2021).

The concept of attractors in complex systems is also related to human cognition. Lucas (2004) noted that *mind attractors* represent the mental categorization of what humans perceive and experience. Lucas stated that the mind has many internal attractors, one for each concept it holds. When individuals interact with an experience or a scene, they gather various information from it trying to identify it. By doing so, the gathered information follows a trajectory toward the strongest mind attractor, an idea, so that the individual will have the natural ability to explain why the phenomena/experience is perceived or learned in one way and not another. Lucas (2004) elaborated that the starting point of forming an idea (a mind attractor) is the basin of attraction and that any idea can later be transferred into a belief when it stays in a stable state for a while. Furthermore, the starting point of any belief is learned in a unique way, depending on the initial conditions and the interaction of other components within the mind or the environment. The path that an idea takes to be in a stable state (belief) is unpredictable and nonlinear.

2.8 CT and Teacher Cognition Research

Although CT is not originally a theory of learning and cognition, it is a broad-based theory pertaining to the evolution and functioning of nonlinear systems. As such, CASs can be applied to many areas, including learning, memory, and cognition (Hase & Kenyon, 2013; Morowitz & Singer, 1995).

Since the 1970s, substantial evidence in teacher cognition research has demonstrated that teacher cognition is complex, systematic, dynamic, changing, co-adaptive, and context-sensitive (Borg, 2006, 2019; Feryok, 2010; Fives & Buehl, 2012; Gao & Zhou, 2021; Li, 2013; Zheng, 2015). Fives and Buehl (2012) referred to the complexity of teachers' beliefs as a significant

theme in reviewing research on teachers' beliefs. The complexity is especially apparent in their discussion of internal and external factors that can support or inhibit teachers' implementation of their beliefs in their classroom practices. They outlined several important internal factors that may affect how teachers act according to their beliefs, such as their personal beliefs, personal knowledge, perceived self-efficiency, and identity. In addition, immediate contextual factors can affect their practices, such as students' attitudes toward their practices, educational policies, and culture. Therefore, many teacher cognition studies consider adopting CT. For example, researchers have investigated the applicability of CT to teacher cognition research and the interactions of several affecting factors (Feryok, 2010), teachers' learning in a teacher education course (Kiss, 2012), the interaction of EFL teachers' belief systems (Zheng, 2015), and language teachers' beliefs about the medium of instruction and actual practices (Gao & Zhou, 2021). In a broad sense, teacher cognition can be perceived as a CAS resulting from the interaction of a collection of elements or agents in a specific way to produce "some overall state or form" at a particular moment in time (Larsen-Freeman & Cameron, 2008, p. 26).

Teacher cognition researchers that have used CT as their theoretical framework moved away from the traditional view of the relationship between teachers' beliefs, practices, and context as a merely cause-and-effect relationship. Accordingly, Feryok (2010) used data from her previously published case study of the practical theory of an Armenian EFL teacher of English (Feryok, 2008, 2005) to explore how CT can be used to explain language teacher cognition. After analyzing data from seven email interviews, two on-site observations, and one oral interview, Feryok concluded that CT as a theoretical framework "fits" language teacher cognition research (p. 277). Characteristics such as complexity, heterogeneity, openness, self-organizing, and co-adaptation were discernible in her data.

One of the key characteristics of both pre-service and in-service language teachers is that their cognitions are complex and encompass various related areas (Andrews, 2003; Breen et al., 2001; Gatbonton, 2008; Johnston & Goettsch, 2000; Zheng, 2015). In CT terms, teachers' cognition is heterogeneous. Like Feryok's (2010) finding, Zheng (2015) found that heterogeneity was evident in EFL teachers' beliefs about EFL in two ways: the diversity of the content and the type of teachers' beliefs. On the one hand, Zheng reported that teachers held various beliefs (e.g., beliefs about the nature of language, language learning, and teaching processes). On the other hand, teachers' beliefs could appear in different forms in their practice (e.g., explicitly or implicitly). The heterogeneity of teacher cognition might be why researchers have focused on exploring varied areas related to language teachers' beliefs to understand teacher cognition.

Moreover, researchers have started to accept the notion that beliefs are complex as they form systems and subsystems. Breen et al. (2001) stated that beliefs within teacher cognition may interconnect, contradict one another, and form complex systems, thereby illustrating the interactive, heterogeneous, and complex nature of belief systems. Teacher cognition does not only consist of networks of various belief systems, but teachers also make active pedagogical decisions based on "complex, practically-oriented, personalized, and context-sensitive networks of knowledge, thoughts, and beliefs" (Borg, 2003, p. 81).

Goertzel (2006) noted that belief systems are some of the most prominent structures in human minds. He defined a *belief system* as "a network of patterns embodying propositions about some aspect of the world or the mind," which are all interconnected as "autopoietic systems" (p. 275), which means belief systems can produce and maintain themselves by creating new beliefs. Goertzel further suggested that a belief system consists of not only beliefs about the world but also beliefs about other beliefs in the mind. When a belief about an aspect changes,

another will consequently be generated by the interaction of different beliefs in the system under the influence of many factors, such as the environment. On a different note, Goertzel (2006) argued that the self-generating and self-organizing nature of belief systems can complicate the process of humans adapting their minds to new beliefs, as the old information is part of an autopoietic network of beliefs that are dynamically regenerated by other patterns that co-adapt with this network. That may explain why some studies found no immediate effect of teacher education courses on teachers' beliefs (e.g., Pennington & Urmston, 1998).

Tudor (2003) stressed the importance of focusing on the dynamics of teaching-learning settings and the nonlinearity of "the pedagogical reality," which "arises from the interaction of participants with one another" (p. 9); this is reflected in "a kaleidoscope of detail which may often seem confusing, contradictory and, at times, rather trivial" (p. 10). This is also evident in the tensions and inconsistencies between teachers' practices and their professed beliefs, which are the reality of the uniqueness of language teaching situations. Thus, Tudor (2003) proposed the use of the concept of "localness" when exploring language teaching and learning situations that are "lived out 'locally'" and where "decision making, too, needs to be a local phenomenon" (p. 8).

Another key feature is the dynamic nature of teacher cognition. For example, Tudor (2003) stressed the importance of focusing on the dynamics of teaching-learning settings and the nonlinearity of "the pedagogical reality," which "arises dynamically from the interaction of participants with one another" (p. 9); this is reflected in "a kaleidoscope of detail which may often seem confusing, contradictory and, at times, rather trivial" (p. 10). This is also evident in the tensions and inconsistencies between teachers' practices and their professed beliefs, which are the reality of the uniqueness of language teaching situations. Thus, Tudor (2003) proposed

the use of the concept of “localness” when exploring language teaching and learning situations that are “lived out ‘locally’” and where “decision making, too, needs to be a local phenomenon” (p. 8).

Moreover, complex systems are sensitive to initial conditions. Feryok (2010), Finch (2010), and Zheng (2015) found that cognitive processes, such as learning and belief formation, were sensitive to initial conditions, including prior experiences or education that form the teacher- and learner-self. Finch (2010), for instance, conducted a study concerning the effect of sensitivity to initial conditions on the language learning process of graduate and undergraduate students. Finch asked the students to reflect on their learning over a semester by identifying critical events from their previous schooling. The study concluded that sensitivity to initial events (critical incidents) could be found in language learning; however, learners had to realize and notice those incidents for subsequent learning to occur.

By the same token, Kiss’s (2012) study is concerned with teachers’ learning, specifically, how student teachers construct meaning by drawing on their past experiences in an intensive L2 teacher education course. The researcher found that teachers’ learning was dynamic, chaotic, unpredictable, nonlinear, and highly sensitive to initial conditions (their past experiences). By analyzing five teachers’ reflective journals, Kiss mapped out the participants’ thinking and learning process. The mapping revealed that what the teachers had learned was highly unpredictable and nonlinear and did not match the objective of the course. Furthermore, the data support the conclusion that teachers associate a network of experiences, people, and places in their past, present, and future to construct an understanding and make meaning of the course content. Kiss then concluded that this understanding has profound implications for teacher education programs, encouraging teacher candidates to change the idea that learning is a linear

process of predictable learning patterns. Teacher cognition research is loaded with evidence that teachers' beliefs, knowledge, learning, and decision-making are complex and dynamic, forming nonlinear interactions with their practices and context.

In teacher cognition research, Burns et al. (2015) elucidated that teacher cognition studies (e.g., Finch, 2010; Zheng, 2015) have shown that “systems are prone to alteration by attractors, which give them emergent rather than stable properties” (p. 594). Although attractors indicate different degrees of stability, strong or weak attractors, they can be used as the focus of analysis to represent core or peripheral beliefs that regulate the individual's behavioral patterns for a certain period (Larsen-Freeman & Cameron 2008). Attractors need different forces from both inside and outside the system for them to change to another stable mode (attractors) or stay as they are. Consequently, the systems under study are better positioned in the context to understand the topic of concern more profoundly (Morrison, 2008).

Moreover, from a CT perspective, Larsen-Freeman and Cameron (2008) stated that units of analysis should identify “collective variables or those that characterize the interaction among multiple elements in a system, or among multiple systems” (p. 242). In other words, the goal of research adopting CT should be thought of as identifying the interactions of the interconnected variables to understand CASs. For instance, de Bot et al. (2007) noted that a learning person is a set of interacting variables characterized by “complete interconnectedness,” so a change in a variable would influence other variables and thus the system as a whole. The agents in educational research (e.g., individuals, institutions, and societies) should be integrated into the system so that the unit of analysis emerges (Morrison, 2008) as a web or ecosystem (Capra, 2013), deriving from and emphasizing a particular topic or focus (Morrison, 2008). In my study, I consider the interactions between the teachers' L2 writing assessment beliefs, practices, and

contexts as the main connected variables and units of analysis, considering them as complex systems with agents and emergent patterns.

2.8.1 Significance of Using CT in Teacher Cognition Research

Conceptualizing teachers' beliefs as CAS aligns with long-standing themes in cognition research. Previous studies, such as those by Borg (2006), Burns et al. (2015), and Freeman (2002), have consistently affirmed the complexity and contextual sensitivity of teacher cognition. Traditional methodologies, which often segment phenomena for isolated study, struggle to account for such complex interconnections (Larsen-Freeman, 2011). CT, as a theoretical lens, promises a more holistic integration of these themes.

Critiques within the field of teacher cognition research often highlight the oversimplification of nuanced issues, such as the relationship between teachers' beliefs and their practices (Borg & Alshumamiri, 2019), and question the research's overarching objectives (Borg, 2018). An acute awareness of the multifaceted factors contributing to the discrepancies between beliefs and practices, coupled with a clearly articulated research purpose, are indispensable (Borg, 2018). In this study, CT serves as a metaphorical channel to elucidate the interaction of the EFL teachers' L2 writing assessment beliefs, their practice, and contexts.

This inquiry into teachers' writing assessment beliefs as a CAS can unveil the heterogeneity of beliefs (i.e., beliefs, beliefs in practice, core, and peripheral beliefs) and their complex interactions. It further allows the examination of the ecological agents that form, shape, limit, and facilitate these beliefs and actual practices, as well as their interrelationships. While many studies have sought to delineate EFL teachers' cognition broadly, a targeted investigation of the writing assessment belief system has the potential to deepen our comprehension of teacher cognition, particularly within the underexplored context of PYPs in KSA.

Moreover, adopting CT as an analytical framework opens vistas for future research directions that could underpin the development of specialized courses for EFL teachers' writing assessments and foster contextualized educational reforms. Davis and Sumara (2005) have contended that "complexity science will not tell educators or educational researchers what to do in any prescriptive sense... But it can provide direct advice on how to focus efforts when preparing for teaching" (p. 318). Educational transformation is achievable through modifications at every conceivable problematic stratum of a system; CT can expose those levels (Larsen-Freeman, 2013).

Additionally, CT shares congruence with other prominent theoretical paradigms in applied linguistics research, such as SCT and ecological approaches, as noted by Larsen-Freeman and Cameron (2008). While SCT and ecological approaches underscore the significance of context in shaping teacher cognition, CT posits that the context itself is a complex dynamic system, both influencing and being influenced by teachers' belief systems (Zheng, 2015). CT acknowledges that interventions in one part of the system may propagate unpredictable consequences throughout the system (Larsen-Freeman, 2013). Feryok (2010) observed that language teaching encompasses a constellation of complex systems, including but not limited to teacher and student cognitions, language, and educational frameworks, all of which are interdependent and mutually influential. Focusing on the L2 writing assessment beliefs of teachers within a PYP at a Saudi University through the lens of CT allows for a nuanced exploration of the localness of their situation across various ecological levels (macro-exo-micro), investigating the interactive interplay of agents that shape writing assessment beliefs and practices, and thereby allows proposing localized solutions and reforms.

2.8.2 CT as a Metaphor

Adapting theories from natural sciences to social sciences can be challenging. CT, as an analytical framework in natural science, is usually characterized by mathematical constraints and computational modeling, which are difficult to apply to the study of language learning and teaching. Therefore, the applied linguistics field has been criticized for using CT concepts metaphorically, and for “just being metaphorical” in their use of CT concepts (Larsen-Freeman & Cameron, 2008, p. 11). Larsen-Freeman and Cameron (2008) disputed this view by asserting the importance of this new metaphor in promoting a new way of thinking about issues in the field of applied linguistics that could “push the field toward radical theoretical change” (p. 11).

Metaphors are essential in picturing how the human mind works. Researchers rely on metaphors when they face difficulty in making sense of the abstract. Metaphors help them comprehend one aspect in relation to another by “analogical structure mapping” between two knowledge domains (Cameron, 1999, p. 25). For example, the two metaphors “the brain is the body’s control center” and “the brain is a computer” both refer to the brain but offer different understandings: the brain is the source of data in the body, and it transmits the data throughout the body (Larsen-freeman & Cameron, 2008, pp. 11-12). Moreover, some conceptual metaphors became the norm in describing concepts in applied linguistics, such as *input* in Krashen’s (1981) input hypothesis and *output* in Swain’s (1993) output hypothesis.

Teacher cognition researchers have used CT in its metaphorical sense (Feryok, 2010; Zheng, 2015). Their findings have resulted in understanding teacher cognition profoundly through a new lens that could effectively contribute to improving teacher professional development. For example, Feryok (2010) intentionally emphasized that her investigation of teachers’ beliefs represents CT features metaphorically. The researcher also asserted that teacher

cognition may not be strictly categorized as complex, yet it can be seen as a complex system. In addition, using CT can fit the teacher cognition research properly and inspire new ways of perceiving the interactions of EFL teachers' beliefs, practices, and contexts. Zheng (2015) also used CT metaphorically and integrated an ecological view of the teachers' contexts. Larsen-freeman and Cameron (2008) recommended welcoming metaphors from different knowledge domains, such as metaphors from ecology, to reconceptualize contexts in the applied linguistics field. Consequently, in this study, I use CT metaphorically and as an "umbrella notion" (Davis & Sumara, 2006, p. 127) to explore EFL teachers' writing assessment beliefs and practices and support it by using an ecological metaphor of the teachers' contexts.

2.8.3 Ecological View of Context

Bronfenbrenner's (2005) *ecological systems model* offers a framework for analyzing relationships among people within communities and the wider society. The ecological systems model positions a developing person in interaction with five environmental systems: a microsystem, mesosystem, exosystem, macrosystem, and chronosystem. The *microsystem* represents the environmental structures at the level closest to the individual and with which they have direct contact, relationship, and experiences, such as the classroom, workplace, and colleagues. The *mesosystem* focuses on the interactions between two or more structures from the individual's microsystems, such as the relationship between the individual's colleagues and their workplace. The *exosystem* also depicts the layer of interaction between two or more structures; however, one or more structures have no direct contact with the individual, such as the relationship between their colleagues and the colleagues' families. The interaction is not at the individual's direct system-level but could negatively or positively affect them. The *macrosystem* represents the "societal blueprint for a particular culture, subculture, or other broader social

context," such as social beliefs about education. Finally, the *chronosystem* represents the sociohistorical dimension of a person's life, such as changing events that affect the development of a person. Those complex systems operate at different levels but are interrelated or "nested" within one another. They focus on a dynamic reciprocal interaction between the developing person and their environment (Bronfenbrenner, 2005, p. 96).

Researchers working with CT have adapted an ecological view of context (Larsen-Freeman & Cameron, 2008). The ecological view is considered complementary and contributes substantially to CT frameworks in applied linguistics (van Lier, 2004; Kramsch, 2003), as it emphasizes the interconnectedness of individuals, groups, communities, cultures, and societies and perceives them as systems within systems (van Lier, 2004). Larsen-Freeman and Cameron (2008) explained that an ecological approach in applied linguistics research uses a specific complex system, a context, as "a base metaphor" to explain how language teachers and learners interact with their environment and context on many levels (p.19). It helps to perceive individuals and their environment as "coupled" and that "because of the coupling, the context itself could change in a process of co-adaptation between the individual and the environment" (p. 7).

Researchers who adopt an ecological perspective in investigating language teachers do not employ all levels of Bronfenbrenner's ecological systems. To illustrate, Edwards (2021) conceptually used an ecological view to review literature on the impact of action research on language teachers' development. The researcher resorted to three ecological systems: micro (individual), meso (school), and macro (wider educational sector). Another example is Hofstadler et al.'s (2021) study of the professional subjective well-being of content and language-integrated learning teachers in Austria from an ecological perspective. Their objective

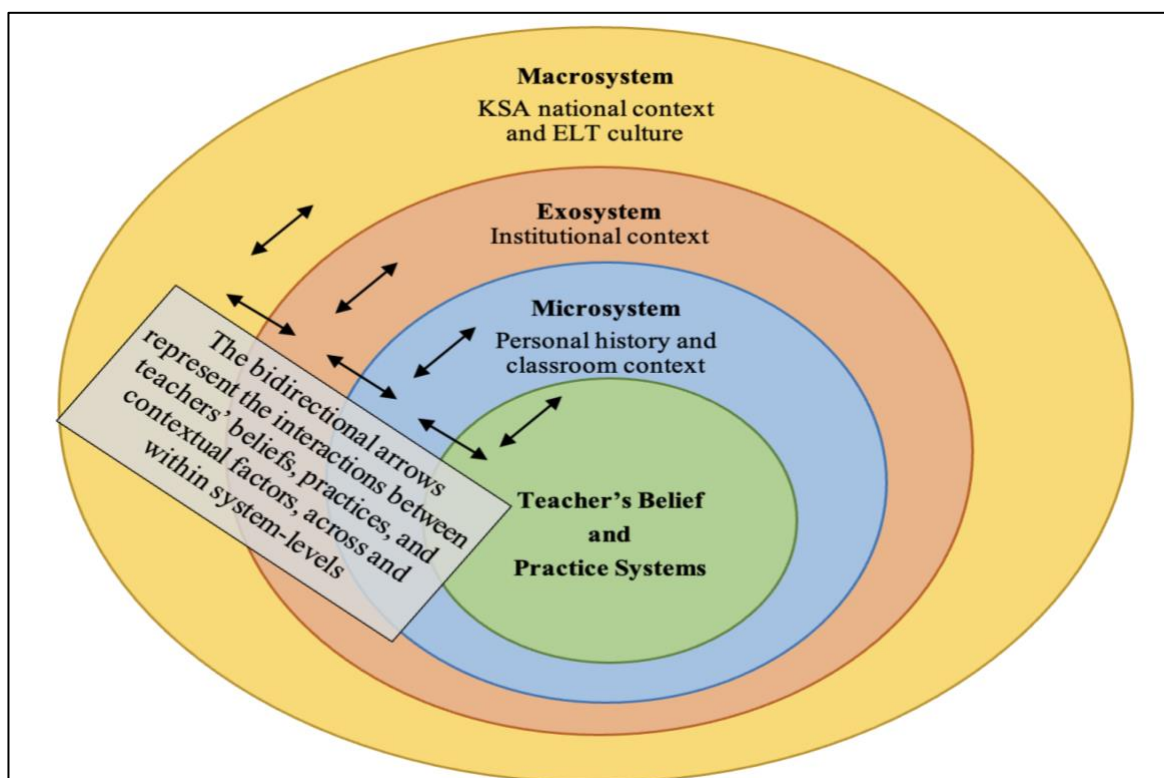
was to analyze the interconnected factors existing in the teachers' environmental subsystems, namely, micro (personal context), meso (class context), exo (school context), and macro (national context) that affect the teacher's subjective well-being.

The most widely used ecological systems in teacher cognition research are microsystem, exosystem, and macrosystem. For example, Zheng (2015) divided the contexts in which she investigated the beliefs of EFL teachers in China into micro (context of classrooms), exo (context of schools), and macro (context of society). The complex ecological perspective of contexts underscores that there is nothing stable about teachers' beliefs (Zheng, 2015), as context itself is not "a stable background variable outside the individual" (Larsen-Freeman & Cameron, 2008, p. 7), which affects teachers' beliefs and practices. Based on those understandings, the ecological view of teachers' context could form the needed boundaries and holistic understanding of the contextual factors influencing teachers' writing assessment beliefs and practices in their environment.

I adopted an ecological view of the teachers' context for the present study by dividing it into three subsystems, as illustrated in Figure 2.2. First, *the macrosystem* represents the national context, including social beliefs about education in KSA, educational regulations and policies, language policies, language culture, and societal beliefs about education and ELT. Second, *the exosystem* represents the institutional context, including aspects that influence the teachers but have no direct connection, such as the University and PYP contexts, encompassing curriculum, program administrators, University policies, testing regulations, and rules. Finally, the *microsystem* represents what the teachers have direct contact and relationship with. For example, it includes teachers' personal history, (i.e., teaching and learning experiences) and classroom context.

Figure 2.2

The Nested Ecological Systems and the Interactions Between Multiple Layers (Adapted from Bronfenbrenner, 2005)



CHAPTER 3: CONTEXT OF THE STUDY

This chapter provides general background information on the educational context surrounding ELT in KSA, with a particular focus on the pedagogy and assessment of L2 writing within Preparatory Year Program (PYP) at KSA universities. It also presents an in-depth description of the context of the study, a PYP at a Saudi University specializing in health sciences. Moreover, this chapter offers detailed information on the writing course under investigation, English II, elucidating its objectives, assessment methods, instruments, policies, and teachers' roles.

3.1 Education and ELT in KSA

According to the Ministry of Education (n.d.), there are 30 public universities, ten private universities, and 41 private colleges distributed across 13 different geographic regions in KSA. The Ministry of Education oversees various aspects of education, including general education, higher education, overseas education (such as scholarships and Saudi schools abroad), special and adult education, and training institutions. It receives the largest governmental budget among all sectors in KSA (Ministry of Finance, 2019). The Ministry of Education holds authority over all educational institutions, including schools, universities, colleges, and training centers, regardless of their public or private status (Ministry of Education, n.d.).

Additionally, as noted by Hilal (2013), KSA is unique in implementing a gender segregation policy throughout its education sector, with exceptions in certain medical colleges and universities, such as King Abdullah University of Science and Technology, as well as international schools governed by international embassies in KSA, like the British International School of Jeddah.

The official language of KSA is Arabic, with English being taught as a foreign language. Since 1970, the English language has gained significant prominence in professional and educational domains, becoming the lingua franca for everyday communication in KSA. Both Arabic and English are used in various contexts, such as stores, street signs, restaurant menus, and Saudi websites. English holds a crucial position in the core curriculum of primary, middle, and secondary education, as well as in higher education (Mahboob & Elyas, 2014). Most private and public university courses in KSA are conducted in English, except for disciplines related to Islamic and Arabic studies, which are taught in Arabic. As noted by Al-Seghayer (2005), “English is used as the medium of instruction in most university technical departments and in science, medicine, and engineering” (p. 126).

3.2 PYP in KSA

Whether private or public, all universities and colleges in KSA offer a PYP as a first-year foundation program. This program consists of two modules: introduction to a subject-track and English Language. The English language courses are standardized across faculties in each university, with variations primarily observed in the subject-track courses. For example, within the science module, students in science colleges are taught fundamental sciences, including introductory physics, chemistry, and statistics courses. The English module primarily focuses on developing the four basic language skills: reading, writing, listening, and speaking, with the occasional inclusion of grammar instruction. It is noteworthy that even students majoring in Islamic studies, who do not study any content courses in English, are still required to enroll in English language courses as part of the degree requirements during their PYP (Al-Seghayer, 2005).

Preparatory Year Program is the commonly used name for the program, although it may have different names in various universities. Despite the variations in names, the program's core objectives remain consistent, focusing on preparing students for English-medium instruction at the university level (Springsteen, 2014). Consequently, the PYP serves as a vital transitional phase bridging the gap between secondary school and university (Al-Shehri, 2017).

According to Al-Shehri (2017), PYPs in Saudi universities share similarities with foundation programs found in international universities, as they primarily aim to enhance students' English language proficiency and academic skills. The PYP is considered an integral part of students' educational program where, for instance, if an undergraduate program spans four years, the PYP constitutes the first year and contributes to students' cumulative grade point average. Hence, students are strongly encouraged to excel in all PYP courses, as their grades significantly impact their prospects of gaining admission to their preferred colleges and achieving subsequent success.

The English modules within most PYPs are typically managed by English language institutes affiliated with universities (Springsteen, 2014) or by English language departments responsible for delivering the preparatory English language courses. The PYPs' staff members primarily consist of Saudi teachers and expatriates from English-speaking countries such as Australia, the U.S., India, and the U.K. (Springsteen, 2014).

PYPs vary across universities in terms of their choice of textbooks, curricula, and assessment practices. Each university creates and develops its curriculum for the PYP it offers. The units responsible for curriculum development and assessments differ from one university to another (Springsteen, 2014). However, both private and public higher education institutions must adhere to the requirements set by the Education and Training Evaluation Commission (ETEC)

when designing curricula and establishing regulations for PYPs (Ministry of Education, n.d.). ETEC, an independent legal entity reporting directly to the prime minister, holds the highest authority in evaluating, assessing, and accrediting qualifications in education and training across public and private institutes in KSA. ETEC establishes the framework for educational regulations and quality standards that the Ministry of Education utilizes to develop curricular guidelines and manage educational institutions (ETEC, n.d.). This centralized approach has drawn criticism, as it limits institutional autonomy in hiring, academic policy development, budgeting, and regulatory establishment, thus hindering independence and competition among institutions (Alamri, 2011; Alkhazim, 2003; Smith & Abouammoh, 2013). Consequently, teachers find themselves at the lower end of KSA's educational administrative hierarchy, and centralized control remains a dominant characteristic of KSA's educational systems and subsystems (Hamdan, 2013).

3.2.1 Teaching and Assessing Writing in PYPs

English writing classes in KSA predominantly follow a traditional product-oriented approach, emphasizing the teaching of writing structures and practical skills, primarily focusing on improving learners' linguistic accuracy, including correct spelling and grammar. Feedback provided to students often centers solely on sentence-level errors (Al-Hazmi, 2006; Al-Seghayer, 2014, 2015; Alshakhi, 2019; Obeid, 2017).

Furthermore, the education system in KSA is characterized by an exam-oriented culture that permeates educational institutions, teachers, and students. As a result, the emphasis is primarily on summative assessments rather than on students' actual language abilities and learning progress (Al-Seghayer, 2005). This exam-oriented culture in Saudi education has a detrimental washback effect on writing classrooms. Washback refers to the impact of tests on

language teachers and learners, influencing their behaviors and decisions that can either support or impede language learning (Messick, 1996). The result is that both teachers and students in Saudi PYP classrooms tend to prioritize exam preparation over meaningful learning experiences (Al-Seghayer, 2014; Alshakhi, 2019; Obeid, 2017).

For instance, Alshakhi's (2019) study examining writing assessment in a Saudi PYP found that writing classes were primarily teacher-centered, featuring extensive mechanical drills, a focus on teaching writing patterns, and neglecting the communicative aspect of writing. Consequently, students tend to emphasize the structural elements of writing, such as grammar, punctuation, and spelling (Alharbi, 2017; Al-Seghayer, 2014, 2015). Alshakhi (2019) further highlighted that PYPs heavily rely on timed essay tests, often neglecting authentic writing tasks. Teachers commonly employ prescribed holistic rubrics and error codes for grading students' essays, but they may lack adequate professional training in effectively utilizing these rubrics and codes.

After examining multiple studies on EFL writing teaching and assessment in Saudi universities, Obeid (2017) identified several challenges teachers and learners face in writing classes. These challenges include the presence of restrictive and static instructions in teaching writing, a lack of engaging genres, an excessive reliance on summative assessments, a predominant focus on exam preparation rather than learning, and a lack of effective writing instruction and learning experiences in high schools. Furthermore, various negative factors were found to impact the teaching and assessment of writing in KSA schools. These factors encompass time constraints, large class sizes, students' weak writing abilities in their L1, ambiguous writing assessment criteria for learners, and a lack of experience and professional

development among L2 teachers in the field of writing instruction (Al-Seghayer, 2015; Al-Jarf, 2011; Ghalib & Al-Hattami, 2015).

In PYPs, while some teachers occasionally adopt recommended assessment practices, such as employing criterion-referenced assessments instead of norm-referenced assessments (Alshakhi, 2019; Obeid, 2017), writing assessment continues to be rooted in a test-oriented culture and fails to align with international standards of writing assessment (Ahmed, 2018). The evaluation and assessment procedures are still in their early stages, with teachers adhering to top-down educational policy and management approaches (Ahmed, 2018). A pre-determined and inflexible curriculum and assessment framework is enforced, placing significant pressure on teachers to conform to prescribed syllabi and assessment requirements within strict time constraints (Ahmed, 2018; Al-Hazmi, 2017).

Alshakhi (2019) highlighted that PYPs often overlook the contextual and interactive aspects of writing when designing writing materials and assessment tools. These materials are typically developed outside the writing classrooms without considering the roles of teachers and the needs of learners. Consequently, diagnostic assessments are rarely utilized, while traditional summative assessments take precedence, primarily focusing on assessments of learning (AoL) (Ahmed, 2018). In PYPs, writing classrooms commonly employ performance-based assessments as part of the achievement-based assessment, emphasizing pre-determined tasks that learners must complete to pass the writing course, such as writing assignments and mid-term and final writing exams (Alshakhi, 2019; Obeid, 2017). Furthermore, Alsamaani (2014) found that teachers in PYPs within Saudi universities need additional training in language test construction and to be made aware of various aspects of assessment. There was a lack of emphasis on

teachers' roles and contributions to the writing assessment processes, as well as insufficient encouragement for teachers' professional development in writing assessment literacy and skills.

3.3 The Context of the Study

This section offers comprehensive information about the study's context. Firstly, it provides an overview of the chosen Saudi University and its PYP, including relevant details about the University and its PYP's administration structure and policies. Secondly, it delves into the general role of teachers within the PYP. Thirdly, it provides a detailed description of the writing course, outlining the syllabus, assessment methods, and tools utilized.

3.3.1 PYP in a University for Health Sciences in KSA

According to the University's website, the institution is a public university dedicated to health sciences. It operates under the supervision of a University Council (UC), which serves as the highest decision-making body within the university. The council consists of various members, including the minister of education as the council's chief, the University's president and vice president, the assistant vice presidents for educational development, quality management affairs, and students' affairs, the general secretary of the council, as well as deans and other officials from multiple colleges.

Based on the academic policies and procedures documents retrieved from the University's website, the University follows a hierarchical structure for its curriculum committees. It consists of the Central University Curriculum Committee and individual curriculum committees at the college level, such as the College of Health Sciences and Professions Curriculum Committee. The Central University Curriculum Committee holds higher authority and is responsible for recommending significant changes to the UC, including the addition of new programs or modifications to existing academic programs and curricula. On the

other hand, the curriculum committees at the college level are responsible for proposing changes to the Central University Curriculum Committee. Their role involves evaluating course content to identify areas of deficiency or redundancy in the curriculum. The college-level curriculum committees typically consist of associate deans (one serving as the chairperson), deputy chairpersons, and course coordinators or representatives from various courses.

The University's website indicates that it has three branches located in three different cities in KSA. All branches operate under a single entity, the UC, and adhere to the same policies, curricula, and study plans. Each campus is divided into two sections, one for female and one for male students. Although the sections share the same overall campuses, they have separate classes, laboratories, departments, faculties, and administrations. Additionally, each campus consists of four colleges: Health Sciences and Health Professions, Medicine, Applied Medical Sciences, and Nursing.

The Health Sciences and Health Professions College is responsible for offering the PYP to first-year undergraduates. This college consists of three departments: the Basic Science Department, which provides courses in biology, health informatics, and anatomy; the Humanities Department, which offers courses on Arabic language and Islamic culture; and the English Language Department, which offers the English language module.

The University's website states that the PYP implements a unified study plan and curriculum that applies to both male and female students, irrespective of their chosen track or major, such as medicine or nursing. However, it is important to note that while the study plan is standardized for all students throughout their PYP, the program has two primary divisions. The first division, called *unified students' sections*, is for students competing for seats in the medical and applied sciences college seats. The second division, called *nursing students'*

section, is for students who are not engaged in a competitive process and will be directed to the nursing college. The assignment of students to these divisions is based on their preferences and grades as part of the admission process. After completing the PYP, students are then assigned to their respective tracks based on their division and GPA.

Based on information derived from the website of the University, the English Department provides an intensive three-semester English language module for all students. The primary objective of this module is to enhance students' English skills through regular and extensive practice in academic reading, vocabulary, oral communication, grammatical structures, and writing in English. According to the PYP unified trimesters' study plan, students must complete 20 credit hours of English language courses during the academic year, covering various components such as reading and vocabulary, grammar, writing, and listening.

3.3.2 English Teachers in the PYP

Based on the University's website and data obtained from initial interviews with participating teachers, the English Department consists of a head, administrators, and teachers who may also serve as course coordinators. This department, where the study was conducted, has approximately 40 teachers who teach various English courses. There are no specific criteria for assigning courses to teachers. Teachers express their course preferences to the department head before the trimester begins. However, the department head assigns courses based on vacancies, the number of sections, and teacher preferences. The number of teachers varies each trimester, depending on the weekly hours required for each course and the number of student sections. Female teachers teach exclusively in the female section, and male teachers teach exclusively in the male section.

The baseline minimum qualification for teachers joining the University is a bachelor's degree with no experience required, meaning that teacher qualifications range from bachelor's degrees to doctorates, with no additional certifications required. Teachers' experience levels vary from novices to those with many years of teaching experience. There are five job titles within the University: teaching assistant, lecturer, language instructor, assistant professor, and associate professor. Each course has two coordinators, one for the male section and one for the female section, and there are two heads of the English department. The teachers are mostly Saudis, whose L1 is Arabic, and expatriates from different English-speaking countries, such as India, the UK, and the U.S. The specific role of teachers in writing courses will be discussed in detail below.

3.3.3 PYP Writing Courses and the Role of Writing Teachers

The PYP study plan indicates that there are three mandatory courses that focus on developing English writing skills. To maintain the confidentiality of the University and teachers, I refer to the writing courses as English I, English II, and English III. English I is taught in the first trimester, English II in the second trimester, and English III in the third trimester. Each trimester spans a duration of 13 weeks. See Table 3.1 for a description of each of the three writing courses obtained from the course syllabuses of each course.

Table 3.1

Brief Description of the Three Writing Courses in the PYP

Course	Credits & Instructional hours/week	Brief Course Description
English I (First Trimester)	<ul style="list-style-type: none"> ● Three credit hours ● two instructional hrs/week for the oral component & three 	An introductory level course builds on students' previous knowledge to develop their academic reading and critical thinking skills while giving

English II (Second Trimester)	<p>instructional hrs/week for the writing component</p> <ul style="list-style-type: none"> ● Four credit hours ● Three instructional hrs/week for the reading component & three instructional hrs/week for the writing component 	<p>practice in the academic writing process. It emphasizes teaching the basic structure of the logical division of ideas, cause/effect, and process paragraphs. The course consists of two interconnected components—an academic writing component and an oral component.</p> <p>The reading component of this course is designed to help students improve their reading and critical thinking skills. It stresses vocabulary enhancement, extracting implied meaning, analyzing the author’s purpose, and drawing conclusions. The writing component of this course aims to develop the writing skills introduced in English I, to introduce new genres of writing paragraphs, to develop students’ abilities to paraphrase, and to introduce the structure of an essay.</p>
English III (Third Trimester)	<ul style="list-style-type: none"> ● Four credit hours ● Three instructional hrs/week for the reading component & three instructional hrs/week for the writing component 	<p>The reading component of this advanced course is designed to develop students’ academic reading skills for the university level. It also stresses chart analysis, extracting implied meaning, analyzing authors’ purpose, and providing responses to written material. The writing component aims to build on the writing skills learned in English I and II. Students will further develop their academic writing skills from the pre-writing planning stage to the final revision stage. Students will learn to plan and write a clear, detailed, and structured academic essay. Students will also learn how to synthesize information from 2-3 academic sources to support their point of view and cite their sources following the established conventions of the Vancouver Referencing system.</p>

I collected data for this study in the second trimester of the academic year 2022-2023. According to the PYP study plan, the writing course that was taught in that trimester was English II.

English II Objectives and Syllabus. The course syllabus indicates that English II is a general English module designed for first-year students, emphasizing reading and writing skills. The course is specifically designed for intermediate-level students and consists of two components: reading and writing. The reading component aimed to enhance students' reading fluency rate, critical thinking skills, and ability to analyze the author's purpose and draw conclusions. The writing component generally focuses on developing students' skills in recognizing different genres of paragraphs, paraphrasing, and structuring essays.

Based on data collected from initial interviews with participating teachers, it was revealed that although the course ostensibly integrates reading and writing skills, in practice, classes, assessments, and grading are distinct and conducted by different teachers. The focus of this study is specifically on the writing component of English II and the instructors responsible for teaching writing, aligning with the objectives of the study.

The specific objectives of the writing component, as stated in the syllabus, are to enable students to: (a) write well-structured comparison and contrast as well as definition paragraphs that include a topic sentence, supporting points, and a concluding sentence; (b) demonstrate knowledge of genre-specific structures; (c) generate ideas during the pre-writing stage using techniques such as brainstorming, mind mapping, WH-questions, and clustering; (d) exhibit awareness of the multi-step writing process; and (e) identify and analyze the main components of an academic essay.

The English II content includes three chapters from the textbook, *Longman Academic Writing Series 3: Paragraphs to Essays, 4th Edition* (Oshima & Hogue, 2013), covering comparison and contrast paragraphs, definition paragraphs, and essay organization. In addition to the textbook, teachers used pre-prepared PowerPoint slides for the course, provided by the course coordinator. However, the participating teachers indicated in the initial interviews that they did not know who developed these slides, but they were sure that the slides were not created by the course coordinator.

English II Assessment Methods. According to the course syllabus, English II includes both formative and summative assessments to evaluate students' writing skills. It is important to note that grades are assigned for both types of assessments. As detailed in the syllabus, grades are an integral part of the formative assessment process, not just the summative assessments. The distinction between these two assessment types seemed to be based on the assessment methods, such as tasks, quizzes, or exams, rather than their intended use for summative or formative purposes. Consequently, the terms 'formative assessment' and 'summative assessment' in this study reflect the labels used by the participating teachers to describe their assessments, not necessarily how these terms are used in the assessment literature.

Firstly, the formative assessment methods in the PYP include both indirect and direct assessments of students' writing skills. These methods are intended to evaluate students' writing abilities and comprehension of English II content throughout the trimester. Students are to be assessed based on their responses to Multiple-choice Questions (MCQ) and their ability to construct paragraphs demonstrating their understanding of writing processes, genres, organization, coherence, and other writing skills and grammatical rules.

The indirect assessments comprise an MCQ paraphrasing task, five online MCQ tasks, an MCQ quiz, and an evaluation of students' participation. The paraphrasing task comprises ten MCQs of a similar structure, each asking the students to select the best paraphrase for the given sentences. See Figure 3.1 for a copy of the MCQs in the paraphrasing task.

Figure 3.1

MCQs Example from the Paraphrasing Task

<p>Instructions:</p> <ul style="list-style-type: none"> • Please read the sentence/s and choose <i>the best</i> paraphrase. • Only <i>one</i> answer is possible. <p>1. A summary is a concise overview of the most important points from a communication, whether it's from a conversation, presentation, or document. Summarizing is a very important skill for an effective communicator.</p> <p>A. A summary of a conversation, presentation, or document is a copy of the main points. Therefore, summarizing is a very important skill in academic writing.</p> <p>B. Summarizing is a very important skill for an effective communicator because a summary is a concise overview of the most important points from a communication, whether it's from a conversation, presentation, or document.</p> <p>C. A summary is a short review of the key ideas in a communication, such as conversations, presentations, or papers. The ability to summarize is crucial for effective communication.</p> <p>D. To have a summary of any communication form like conversations, presentations, or papers, you should have a professional communicator who can summarize effectively.</p>

All online tasks follow a similar structure, consisting of a paragraph accompanied by five MCQs that pertain to identifying elements of the paragraph's structure, such as the topic sentence and main points. Refer to Figure 3.2 for an example of one of the five online tasks.

The quiz consists of two sections. The first section includes an essay, followed by nine MCQs testing the student's understanding of the essay's structure, sentence purposes, sentence types, and use of transitions within a paragraph. The second section presents a paragraph with blanks, asking students to fill in the blanks by choosing the appropriate answer from a provided table. For an example of the MCQs from the quiz, see Figure 3.3.

Figure 3.2

Copy of an Online MCQs Task

BEFORE READING:

Q. What is this paragraph about?

A. The economic strategies of Brazil and the United States.
 B. The differences in climate between Brazil and the United States.
 C. The similarities between Brazil and the United States.
 D. The history of colonization in Brazil and the United States.

DURING READING:

Q. Which sentence is the topic sentence of this paragraph?

A. Sentence 2
 B. Sentence 3
 C. Sentence 4
 D. Sentence 5

Q. Which sentences are the main points of this paragraph?

A. Sentence 4, 12, 14
 B. Sentence 4, 11, 16
 C. Sentence 4, 12, 15
 D. Sentence 3, 11, 16

AFTER READING:

Q. What is the organizational style of this paragraph?

A. Point-by-point organization
 B. Block organization

Q. What is the most appropriate concluding sentence for this paragraph?

A. In summary, Brazil and the United States share many cultural and political differences.
 B. In conclusion, Brazil and the United States stand out for their unique languages and traditions.
 C. In conclusion, Brazil and the United States share significant similarities in size, diversity, and values.
 D. Finally, Brazil and the United States both have coastlines that attract tourists from all over the world.

Figure 3.3

MCQ Examples from the Quiz

Section A: Read the previous essay and answer the following questions:

1. Which sentence is the thesis statement?

A. Sentence 3
 B. Sentence 4
 C. Sentence 5
 D. Sentence 6

2. Which is a fragment sentence?

A. Sentence 1
 B. Sentence 2
 C. Sentence 3
 D. Sentence 4

3. Which sentence contains an appositive?

A. Sentence 2
 B. Sentence 3
 C. Sentence 4
 D. Sentence 5

4. What is the purpose of sentence 7?

A. to show a similarity
 B. to give supporting details
 C. to give an example
 D. to give a main point

Section B: Read the following paragraph and fill in the gaps.

Communication Now and Then

Communication between people now and in the past has similarities and differences. (11) _____ modern and past forms of communication are similar; they differ in speed and tools. Forms of communication are similar to past methods as follows. In the past, letters and phone calls were both frequently used. (12) _____, both are used now. Email and text messaging are still popular written ways of communication. Oral types include the phone, mobile phone, and instant voice messages. The main difference in long-distance communication is speed. This is most obvious in writing. Letters in the past took days to arrive, [13] _____ e-mails arrive almost instantly and can be read in seconds today. In the past, a memo would be carried around the workplace to send a simple message, which took time. (14) _____, nowadays, text messages can be sent immediately. Communication strategies also differ. For example, in the 1960s, the telephone and letter were the main long-distance communication tools, (15) _____ today's communication methods are numerous. For example, these include phone, letter, email, text messaging, Snapchat, FaceTime, and Twitter. All in all, the way people communicate now and then are similar in some ways and different in other ways.

11.	12.	13.	14.	15.
A. And	A. Like	A. as	A. Also	A. also
B. But	B. Even though	B. just like	B. And	B. but
C. While	C. Similarly	C. like	C. However	C. in contrast to
D. On the other hand	D. Whereas	D. whereas	D. Similarly	D. however

The direct assessments consist of two writing tasks, which require students to compose paragraphs in two different genres, comparison and definition. Students have to compose 200–250-word paragraphs following a multi-step writing process of brainstorming, outlining, and writing a first draft and then a final draft.

The grades assigned to the mentioned formative assessments account for 30% of the final grade, with 8% allocated to the MCQ tasks (i.e., paraphrasing and online tasks), 4% to the MCQ quiz, 2% to students' participation, and 16% to the writing paragraph tasks. Therefore, teachers' formative assessments comprise 30% of the overall course grade; this portion is based on specific predetermined assessment methods. It is worth highlighting that the course also includes formative assessment methods for the reading component, which account for 15% of the overall course grade. This suggests a greater emphasis on writing in the course grade compared to reading. The ratio of the writing focus to reading focus in the formative assessment is 2:1, signifying that the weight given to writing is twice that for reading in the formative assessment. The reading assessments are conducted by a teacher other than the participating writing teacher.

Secondly, as stated in the course syllabus, the summative assessment methods in the PYP include one midterm and one final exam. The exams mainly include the following question forms: multiple choice questions, cloze exercises, fill-in-the-blank questions, and error identification questions. However, there is also a classroom-written final exam that takes place before the integrated final exam. This component does not apply to the midterm exam.

The exams integrate both reading and writing components. The distribution of writing assessment grades is 15% for the midterm, 25% for the final, and 15% for the final written exam. This means that the emphasis on writing in the summative assessment is also greater than for reading, 35% to 20% or 7:4 ratio. The ratio suggests that writing is prioritized more in terms of

assessment focus, with approximately seven units allocated to writing for every four units allocated to reading. Refer to Table 3.2 for information regarding the assessment methods of the writing component of the English II course, as obtained from the course syllabus.

Table 3.2

Summary of the Writing Assessment Methods in English II

Course	Semester Start and End Date/ Duration	Assessment Methods
English II	From 4 th December 2022 till 2 nd March 2023 (13 Weeks)	Formative Assessment Components <i>Indirect Assessment of Writing:</i> 1. Five online tasks (5%) 2. Paraphrasing task (3%) 3. One quiz (4%) 4. Participation (2%)
		<i>Direct Assessment of Writing:</i> 1. Task 1: Compare and contrast paragraph (8%) 2. Task 2: Definition paragraph (8%)
		Summative Assessment Components <i>Indirect Assessment of Writing:</i> 1. Midterm exam (15%) 2. Final exam (25%)
		<i>Direct Assessment of Writing:</i> 1. Final written exam (15%)

The rubric employed to assess the two writing-paragraph tasks and final written exam is analytical in nature and offers a comprehensive evaluation of a student's performance on multiple criteria, as shown in Table 3.3 below. The rubric breaks the writing assessment into five categories: process and format, organization, content, lexical resource, and mechanics and grammar. Each category consists of five levels of achievement: excellent (5), above average (4), average (3), below average (2), and inadequate (1). The total score of the rubric is 25.

The *process and format* criterion evaluates the completion of the writing process and the correct formatting and neatness of the paragraph. *Organization* assesses the effectiveness of the

topic sentence, the introduction of main points with relevant supporting details, and the strength of the concluding sentence. *Content* examines the unity, coherence, logic, and quality of ideas in writing, as well as the use of appropriate transition signals. *Lexical resource* considers the range of academic vocabulary, language style, sophistication, and the appropriate use of a formal register. *Mechanics and grammar* criteria focus on the accuracy of grammar, punctuation, spelling, and sentence structure. The rubric provides specific descriptions for each level of achievement within each category.

Teacher's role in English II. As mentioned previously, writing instructors do not have the authority to change the grade distribution or assessment methods and format. Each task and quiz have specific due dates listed in the syllabus, which instructors are required to follow. However, the participating teachers indicated that there is a collaborative effort in developing the formative writing assessment methods. Concerning the MCQ tasks, teachers suggest tasks to the course coordinator, who then organizes and returns the tasks to the teachers for classroom use. For paragraph writing tasks, each instructor proposes topics, and the coordinator selects 12 for use. Instructors then choose four topics for their students, who in turn select one for their paragraph writing task. Regarding the MCQ quizzes, teachers adhere to a standardized format for developing them. After completion, the assessments are reviewed by the course coordinator for approval prior to classroom implementation.

Table 3.3

Copy of the Writing Paragraph Tasks Rubric

Category	Excellent 5	Above Average 4	Average 3	Below Average 2	Inadequate 1
Process & Format (<i>brainstorm, outline, draft, & final</i>) /5	<ul style="list-style-type: none"> Fully completes and exceeds in the parts of the writing process (brainstorm, outline, draft, and/or final copy) Paragraph is correctly formatted and neat 	<ul style="list-style-type: none"> Completes the parts of the writing process with minimal procedural errors. Minor errors in formatting and/or neatness 	<ul style="list-style-type: none"> Partially completes parts of the writing process (1 part is incomplete or missing) Some errors in formatting and/or neatness 	<ul style="list-style-type: none"> Minimally completes parts of the writing process / parts are incomplete or missing) Several errors in formatting and/or neatness 	<ul style="list-style-type: none"> Fails to complete parts of the writing process are incomplete or missing) Several errors in formatting and/or neatness
Organization (<i>topic, supporting, & concluding sentences</i>) /5	<ul style="list-style-type: none"> Provides a strong, original topic sentence using the mechanisms according to the genre. Clearly introduces main points with developed, relevant supporting details. Strong concluding sentence 	<ul style="list-style-type: none"> Provides a good topic sentence using the mechanisms according to the genre. Clearly introduces main points with relevant supporting details. Good concluding sentence 	<ul style="list-style-type: none"> Provides an adequate topic sentence. Introduces the main points with some relevant supporting details. Adequate concluding sentence 	<ul style="list-style-type: none"> Topic sentence is missing the topic or controlling idea. Weak/missing 1 or more main points, supporting details. Weak concluding sentence 	<ul style="list-style-type: none"> Fails to provide 2 or more of the following: a clear topic sentence, main points, supporting details, and/or concluding sentence.
Content (<i>unity, coherence, logic, & quality of ideas</i>) /5	<ul style="list-style-type: none"> Maintains unity throughout Arrange and synthesize ideas logically and coherently with examples, facts, in-depth analysis, etc. Variety of transition signals Evidence of original thought and support for the topic content errors (0-1) 	<ul style="list-style-type: none"> Maintains unity throughout Arranges ideas and information logically and coherently with general examples, thoughts, etc. Little variety of transition signals content errors (2-3) 	<ul style="list-style-type: none"> Lacks unity: Off-topic 1 time in the paragraph. Adequately arranges ideas and information logically and coherently with general examples, thoughts, etc. Missing or repeated transition signals content errors (4-5) 	<ul style="list-style-type: none"> Lacks unity: Off-topic 2 times in the paragraph. Inadequately arranges ideas and information logically and coherently; lacks progression of ideas. Inappropriate transition signals content errors (0-1) 	<ul style="list-style-type: none"> Lacks unity: Off-topic (all or most of the ideas) <i>and/or</i> Ideas are disorganized, illogical, unoriginal, and underdeveloped. Little to no transition signals
Lexical Resource (<i>vocabulary, language, & sophistication</i>) /5	<ul style="list-style-type: none"> Exemplary and wide range of academic vocabulary, style, and sophistication Formal register Word choice errors (0-1) 	<ul style="list-style-type: none"> Sufficient range of academic vocabulary, style, and sophistication Formal register Word choice errors (2-3) 	<ul style="list-style-type: none"> Limited range of academic vocabulary, style, and sophistication Some formal register Word choice errors (4-5) 	<ul style="list-style-type: none"> Very limited range of academic vocabulary, style, and sophistication; uses basic/repetitive vocabulary. Informal register Word choice errors (6-7) 	<ul style="list-style-type: none"> Uses only basic/repetitive vocabulary; lacks style and sophistication. Informal register Word choice errors (8+) Word form errors (8+)
Mechanics & Grammar (<i>grammar, punctuation, spelling, etc.</i>) /5	<p>Error free sentences:</p> <ul style="list-style-type: none"> Free of distracting spelling, punctuation, capitalization, and other grammatical errors Free of fragments, comma splices, and run-ons Variety of sentence structures errors (0-2) 	<p>80% of sentences error-free:</p> <ul style="list-style-type: none"> Few spelling, punctuation, capitalization, and other grammatical errors allowing reader to follow ideas clearly. Minimal fragments, comma splices, and run-ons. Variety of sentence structures errors (3-5) 	<p>½ of the sentences error-free:</p> <ul style="list-style-type: none"> Some spelling, punctuation, capitalization, and other grammatical errors allowing reader to follow ideas adequately. Some fragments, comma splices, and run-ons. Limited range of sentence structures errors (6-8) 	<p>Less than ½ of the sentences error-free:</p> <ul style="list-style-type: none"> Several spelling, punctuation, capitalization, and other grammatical errors making reading sometimes difficult Several fragments, comma splices, and/or run-ons. Very limited range of sentence structures errors (8-10) 	<p>All/majority of sentences have errors:</p> <ul style="list-style-type: none"> Frequent spelling, punctuation, capitalization, and other grammatical errors create distraction, making reading difficult. Frequent fragments, comma splices, and run-ons. Incorrect sentence structures
Total Score / 25	Additional Comments:				

Besides the predetermined rubric, the participating teachers reported that they have to use a standardized textbook, syllabus, PowerPoint presentations, and writing task descriptions. They are not allowed to make any modifications to the provided teaching materials, although they are allowed to supplement them with additional materials. Furthermore, teachers, including course coordinators, are unaware of who was responsible for developing the standardized teaching materials. The participating teachers reported that it is possible that these materials were developed by the Central University Curriculum Committee.

Based on the participating teachers' reports of their practices, teachers contribute to the construction of the midterm and final exams by following a specific format and sending their contributions to the course coordinator. For instance, each teacher suggests and write three multiple-choice questions, three cloze exercises, three fill-in-the-blank questions, three error identification questions, and three paragraph prompts. The coordinator then selects, organizes, and combines the teachers' questions, sending the complete exam back to the teachers for review. Subsequently, the reading and writing coordinators schedule a meeting with the writing and reading teachers to discuss the exams, primarily focusing on identifying errors and making necessary modifications to the content.

Regarding the grading of midterm and final exams, there is a specific system in place. The MCQ parts of the exams are corrected automatically by an assessment unit. In the written part, the participating teachers reported that they do not grade the exam papers of their own students to maintain objectivity. Instead, the course coordinator assigns each teacher to rate and correct the answers of students they are not teaching. Once the ratings and corrections are completed, the papers are returned to the course coordinator. After all the teachers have submitted their corrections, the coordinator asks each teacher to review a set of papers from

students they are not teaching. In cases where there is a discrepancy in the assigned grades, the two teachers involved engage in a discussion to reach a consensus. Once the review is completed, the writing teachers return the reviewed papers to the coordinator, who then posts the students' grades on the student portal.

CHAPTER 4: METHODS

As described in Chapter 1, the main objective of this study was to examine the L2 writing assessment beliefs and practices of EFL teachers within a specific program (PYP) at a Saudi University while also exploring the interaction between these beliefs, practices, and contextual factors. A qualitative multiple-case study design involving five EFL writing teachers was employed to address the following research questions:

1. What are EFL writing teachers' beliefs regarding the assessment of L2 writing?
2. How do teachers' writing assessment beliefs relate to their actual practices?
3. What are the interrelationships between EFL teachers' writing assessment beliefs, actual practices, and contextual factors?

This chapter provides a comprehensive description of the study design, encompassing various levels of sampling, namely the Saudi university, the writing course, and the participants, as well as the recruitment and demographic backgrounds of the participants. Furthermore, it describes the qualitative data collection methods employed, including initial interviews, observations, think-aloud protocols (TAPs), follow-up interviews, and documents. Additionally, the chapter details the data analysis process, which employed an inductive approach to thematically analyze data from the five cases, incorporating both within-case and cross-case analyses.

4.1 Qualitative Research

Qualitative research encompasses data collection procedures that produce descriptive and “non-numerical data,” which is then analyzed using “non-statistical methods” (Dörnyei, 2007, p. 24). The qualitative approach's purpose is not to test theories but to describe and interpret social phenomena from individuals' perspectives within their natural setting (Cohen et al., 2007), such

as schools, hospitals, and workplaces (Holliday, 2015). Qualitative approaches have been used to investigate teacher cognition and teachers' beliefs under the umbrella of CT (Feryok, 2010; Gao & Zhou, 2021; Kiss, 2012; Zheng, 2015). Qualitative methods yield rich data and explain "the complexity, texture, and nuance involved in how individuals and groups experience themselves and their worlds" (Kamberelis & Dimitriadis, 2005, p. 17). It shares the CT purpose of focusing on "meaning and holistic concerns rather than discrete variables, statistics, and standardization" (Phakiti & Paltridge, 2015, p. 13). Given that this study aims to explore the complex interactions between teachers' writing assessment beliefs, practices, and ecological factors, it was deemed appropriate to collect and analyze data qualitatively to understand how these interactions occur within complex systems.

4.1.1 Multiple-Case Study Design

Creswell and Poth (2018) defined a case study design as follows:

A qualitative approach in which the investigator explores a real-life, contemporary bounded system (a case) or multiple bounded systems (cases) over time, through detailed, in-depth data collection involving multiple sources of information (e.g., observations, interviews, audiovisual material, and documents and reports), and reports a case description and case themes. The unit of analysis in the case study might be multiple cases (a multisite study) or a single case (a within-site study). (p. 97)

Mackey and Gass (2015) delineated that a case study design aims to provide a holistic picture and rich contextualized data of the complexities embedded within a specific setting. It allows for exploring the topic of concern in a real-life context (Yin, 2014) and in-depth (Blatter, 2008). A case study is a contextual study (van Lier, 2005, p. 205) that is primarily applied when the phenomena of concern cannot be decontextualized and when a deep understanding is influenced

by the interaction of multiple factors (Yin, 2014). Moreover, case study designs are suitable for CT because the nature of CASs calls for an investigation of interactions across the boundaries of their contexts (Anderson et al., 2005). Li (2020) emphasized that case studies have emerged as an insightful strategy to examine the complexity of teacher cognition and contextual factors.

Many recent studies in the field of teacher cognition have identified themselves as case studies (Li, 2012; Li & Walsh, 2011; Min, 2013; Kabouha & Elyas, 2015; Zheng, 2015). One of the reasons for the effectiveness of using case studies in teacher cognition research is that “beliefs cannot be directly observed or measured but must be inferred from what people say, intend, and do” (Pajares, 1992, p. 314). Therefore, Pajares suggested that research approaches such as “teachers’ *verbal expressions, predispositions to action, and teaching behaviors*” must be considered when investigating teachers’ beliefs to ensure the validity of the results and the significance of the study (p. 327).

According to Stake (2013), a case study design can be multiple or a collective of cases. The multiple-case study design examines phenomena within and across multiple cases to provide an in-depth understanding (Punch & Oancea, 2014) and offers a more convincing and comprehensive explanation of the research findings (Richards, 2011; Yin, 2014). One limitation of the case study design is its inability to generalize from a single case; however, Mackey and Gass (2015) demonstrated that adopting a multiple-case study design can lead to more conclusive results. The goal of this study was to understand and explore the complexity of interactions between teachers’ writing assessment beliefs, practices, and contextual factors in the targeted program. Readers in similar contexts, particularly other PYPs in KSA, can relate the findings of this study to their own experiences and situations (Merriam & Tisdell, 2016). Another limitation of the case study design is the potential bias of the researcher. However, as

Yin (2014) argued, researchers' subjectivity can be minimized by collecting data using different methods.

This study examined teachers' writing assessment beliefs using five case studies of EFL teachers who assessed their students' writing in a one-trimester writing course. The aim was to explore and describe the interactions among these five teachers' writing assessment beliefs, writing assessment practices, and contextual factors. By focusing on five cases within one program, a deeper understanding and more detailed study findings were achieved, allowing for the examination of a broad and diverse range of assessment beliefs and practices from different teachers within a single context. Furthermore, to address the limitations of the design, data were collected from the five teachers using multiple methods including: initial interviews, observations, TAPs, follow-up interviews, and document collection.

4.2 Sampling Procedures

Dörnyei (2007) explained that qualitative inquiry is not concerned with the representativeness of the respondent sample or the distribution of experiences within the population. Instead, the primary aim of sampling is to identify individuals who can provide insightful perspectives on the phenomenon under study, thereby enhancing our learning potential. Furthermore, Dörnyei suggested that the 'purposeful' or 'purposive' sampling method best suits this objective.

In this study, purposeful sampling was used to gather information from EFL writing teachers in a PYP at a Saudi University who can offer the most relevant insights into the research problem being studied. According to Creswell and Poth (2018), the purposeful sampling approach in qualitative research involves two key factors that can vary depending on the approach: the decision regarding participant or site selection for the study and the specific

sampling strategy employed. Creswell and Poth (2018) also suggested that researchers should consider different sampling levels in case studies. Researchers have the option to sample at the site level, the event or process level, and the participant level. A well-designed qualitative study may include one or multiple sampling levels, and it is essential to identify each level clearly.

In this study, the first level of purposeful sampling involved selecting the University based on several criteria. Firstly, the University stands out as one of the top-ranked public institutions in KSA, as indicated by the QS World University Rankings (2021). Secondly, public universities in KSA generally hold higher positions in national rankings than private institutions, making them more preferred among teachers and students (Jamal, 2021). Additionally, the chosen University offers a PYP, which is noteworthy because certain public universities offering graduate degrees do not provide PYPs; instead, they require students to submit IELTS or TOEFL scores for admission. Lastly, the University is unique as the only Saudi university operating multiple branches across KSA, boasting a significant student and faculty population and substantial government funding. Therefore, it was believed that conducting this study within the PYP at this University would yield valuable insights and carry significant implications for teacher education and policymakers in the higher education landscape of KSA. The data were collected from one branch of the University, as all branches adhere to the same policies, study plan, curriculum, and syllabus.

The second level of sampling involved the writing course. As mentioned above, the chosen program offers three writing courses, with one course offered in each trimester. The aim of this study was to investigate teachers' beliefs and practices in writing assessments within a single-trimester course. Therefore, the selection of the writing course was based on the timing of

the data collection. Data collection for this study began in the second trimester of the 2022-2023 academic year. The specific writing course taught during that trimester was English II.

The final level of sampling pertained to the teachers. The participating teachers for the study were selected using a purposive sampling strategy, following specific criteria (Dörnyei, 2007). The recruitment of teachers was based on the following criteria:

Firstly, teachers who taught the writing course in the second trimester were included, regardless of whether they were teaching in unified or nursing sections. Since both sections follow the same study plan and curriculum, any differences in students' language proficiency levels would be considered a contextual factor. Secondly, Saudi teachers with at least three years of teaching experience in the PYP were chosen. This criterion aimed to account for the predominant Saudi teaching staff and eliminate potential influences arising from differences in linguistic backgrounds. It also aimed to gather rich data on the impact of the PYP context on teachers' writing assessment beliefs and practices. Thirdly, female teachers from the female section were targeted. This decision was based on the practical challenges faced by a female researcher in obtaining permission to observe male participants in the male section. Finally, age, educational background, qualifications, and other teaching experiences were not considered during participant selection because they were deemed part of the teachers' personal context.

4.2.1 Recruitment of Participants

Prior to participant recruitment and gathering any data, ethics approval was obtained from York University's Ethics Review Board, as detailed in Appendix A. To recruit participants for this study, I emailed the English Department at the University to request permission to collect data. After the head of the department approved the study, I asked for the names of the Saudi teachers teaching writing courses in the second trimester. Once I received the names of nine

teachers and their WhatsApp numbers, I sent them WhatsApp messages introducing myself and providing a description of the study and the data collection procedures. Six teachers responded, expressing their willingness to participate. I sent the consent form to each teacher and informed them that their participation would be entirely voluntary and that they could withdraw from the study at any time. While six participants signed the informed consent letter (see Appendix B), only five completed all the study data collection methods. One participant withdrew after the initial interview. Consequently, I decided to exclude her from the study.

Pseudonyms were assigned to the participants to protect their privacy and confidentiality throughout the study. Table 4.1 presents the demographic data of the five participants. All participants are female Saudi teachers. Their ages ranged from 25 to 40 years, and their ELT experience varied from 3 to 8 years. At the time of the study, all participants had been teaching at the PYP for a minimum of 3 years. Each participant holds an MA degree, in teaching English to speakers of other languages (TESOL), except for Aya and Noor, who hold MAs in linguistics and education, respectively. None of the participants possessed any additional certifications beyond the MA. However, apart from Aya, all participants had undergone training experiences such as attending various language teaching and learning workshops. Leena, Batool, and Rana were assigned to teach different nursing sections, while Aya and Noor were assigned to teach unified sections.

Table 4.1*Participants' Demographic Data*

Pseudonym	Gender	Age	Years of teaching experience	Years of teaching experience in the PYP	Training Experiences	Qualifications	Writing Sections Taught
Leena	Female	30–35	8	6	Attended workshop on effective teaching & giving feedback in writing	MA in TESOL	Nursing
Aya	Female	35–40	3	3	N/A	MA in Linguistics	Unified
Batool	Female	25–30	6	5	Attended several Cambridge workshops (i.e., task-based language learning & teaching)	MA in TESOL	Nursing
Noor	Female	25–30	5	4	Attended workshop on writing assessment and learning (How to correct papers using error codes? And how to write essays?)	MA in Education	Unified
Rana	Female	30–35	7	4	Attended workshop on understanding students' level	MA in TESOL	Nursing

4.3 Data Collection Methods

I employed three primary methods to address the research questions: initial and follow-up interviews, observations, think-aloud protocols (TAPs), and documents. These data collection methods were chosen to investigate teachers' writing assessment beliefs, their writing assessment practices, as well as their interactions with contextual factors. The first method involved

conducting initial interviews with the participants in a semi-structured format. The second method consisted of observing two writing classes for most of the teachers, with all observations being audio-recorded and supported by field notes. The third method utilized audio-recorded concurrent TAPs, where teachers assessed students' written paragraphs while verbalizing their thought processes, followed by follow-up interviews that focused on both the TAPs and observations. Additionally, various documents were collected throughout all phases of the study, such as teaching and assessment materials, to supplement the other data collection methods.

Table 4.2 provides a summary of the data collection methods and the purpose of each method.

Table 4.2

Data Collection Overview

Data Source	Details	Elicited Information
Main Data Sources		
Initial Interviews (IIs)	<ol style="list-style-type: none"> 1. One interview for each teacher 2. 45 to 60 minutes 3. In English 4. Audio-recorded 5. Semi-structured in nature 6. Question guidelines were used 	Participants' PYP and University context information, background information, and professed writing assessment beliefs, beliefs in practice, and contextual factors
Unstructured Classroom Observations (OBSs)	<ol style="list-style-type: none"> 1. Two observations 2. 1 hour to 2 hours 3. In English 4. Audio-recorded 5. Observation guidelines were used, and field notes were taken 	Participants' actual writing class assessment practices
Concurrent Think-aloud Protocols (TAPs)	<ol style="list-style-type: none"> 1. Two TAP sessions 2. 25 to 35 minutes 3. In English 4. Audio-recorded 5. Participants thought aloud while performing a task—assessing six written texts by students. 6. TAP instructions and training were provided to the participants. 	Participants' actual rating beliefs and practices—providing written corrective feedback and grading.
Follow-up Interviews (FIs) regarding the TAPs and observed classes	<ol style="list-style-type: none"> 7. Two FIs 8. Conducted immediately after the TAP sessions. 9. In English 10. Audio-recorded 	The relationship between participants' writing assessment beliefs, practices, and contextual factors

	<ol style="list-style-type: none"> 11. Participants reflected on their TAPs and observed classrooms. 12. Semi-structured in nature 13. Question guidelines were used 	
Documents	20 writing samples of participant's feedback and rating practices on the first drafts of students' paragraphs	Understanding participants' actual rating practices—written feedback form, type, and focus, and their grading practices
Supplementary Data Sources		
Documents	<ol style="list-style-type: none"> 1. University Policies 2. PYP study plan and curricula 3. Writing course syllabus 4. Teaching materials 5. Observation field notes 	<ul style="list-style-type: none"> • Informing the description of the University context • Understanding the PYP context, teaching and assessment policies • Understanding the writing course objectives, scope, and assessment methods and tools • Triangulating the main data about writing assessment beliefs, practices, and contextual factors

The data were collected over the course of one trimester, spanning 13 weeks, starting from 4th December 2022 till 2nd March 2023. The writing component of the course consisted of three hours of instruction per week. Table 4.3 outlines the timeline for the data collection procedures of the initial interviews and observations in relation to the course weeks. The timelines for TAPs and follow-up interviews will be described later.

Table 4.3

The Data Collection Timeline of Initial Interviews and Observations

Pseudonym	Initial Interview Timeline	Number of Observation	1 st Observation Timeline	Length of 1 st Observation	2 nd Observation Timeline	Length of 2 nd Observation
Leena	Week 1	2	Week 4	One hour & 50 minutes	Week 8	One hour & 50 minutes
Aya	Before the beginning of the semester	2	Week 4	One hour	Week 7	One hour & 50 minutes

Batool	Before the beginning of the semester	2	Week 4	One hour & 50 minutes	Week 7	One hour
Noor	Before the beginning of the semester	1	Week 3	One hour & 50 minutes		
Rana	Before the beginning of the semester	1	Week 3	One hour		

As indicated in Table 4.3, the initial interviews were conducted over the course of one month before the observation. The initial interviews began on October 3rd, prior to the commencement of the second trimester on December 4th, except for Leena's interview. The interviews were conducted before the start of the trimester to ensure the completion of the interviews and allow time for listening to the interviews prior to the observation phase. This timing was essential for focusing on specific aspects of assessment during the observations. Commencing the interview phase earlier did not compromise its purpose, as all participants had prior experience teaching writing in the PYP and were teaching the English I course in the first trimester.

The classroom observations commenced on December 21st and spanned a duration of seven weeks. The timing of the observations was strategically planned based on the analysis of the course syllabus. Specifically, the observations were scheduled to take place during weeks 3 to 9 of the course, which concentrated on the instruction of paragraph writing and various writing tasks. This timeframe was selected due to its relevance to the study's focus on writing assessment beliefs and practices. Moreover, the initial two weeks of the course were primarily dedicated to introductory activities and organizing students' schedules, with limited assessment practices taking place during this period.

Due to the teachers' busy and overlapping schedules, as well as the limited time available for conducting both the observations and TAPs, I was able to observe each participant twice, except for Noor and Rana, who were observed only once. The duration of class observations varied, depending on the teacher's scheduled teaching hours for that day. Teachers typically taught writing for three hours per week, with most observations lasting between one and two hours.

4.3.1 Initial Interviews

I used semi-structured interviews in my study as they provide an appropriate method to understand teacher cognition and experiences in-depth (Borg, 2006). They also offered several advantages. First, semi-structured interviews are two-way conversations that are flexible enough to accommodate other emergent questions and themes during the conversation (Kvale & Brinkmann, 2009), such as adding further questions to the pre-planned questions to improve the natural flow of the conversation and elicit additional insights, clarifications, and elaboration of the participants' answers (Dörnyei, 2007). Second, this method allowed me to build rapport with the participants using the flexible nature of the semi-structured interviews and avoiding researcher-dominated procedures that might guide the participants' expressions of beliefs, which was crucial to the quality of the investigation (Borg, 2006).

Third, semi-structured interviews can provide insight into participants' mental lives, "perceptions, meanings, definitions of situations and constructions of reality" (Punch & Oancea, 2014, p. 182), and allow researchers to interpret participants' experiences from their point of view (Borg, 2006). Finally, I used semi-structured interviews to initiate the exploration of the first layer of teachers' professed beliefs, feelings, beliefs in practice, and influential contextual factors (Li, 2017). They also helped outline the complex and multi-layered understanding of the

interaction of language teachers' belief systems, practices, and ecological systems (e.g., Feryok, 2010; Li, 2013). For example, asking about the teachers' feelings regarding writing assessments revealed how some contextual factors shaped their feelings and, thus, their belief-practice relationship. Although interviews are criticized for providing only a summary of what the participants say rather than how they say it, the data analysis procedure could help minimize this issue by including delays, hesitations, feelings, and intonation (Li, 2017), which could reveal the intensity of a belief. Additionally, I used additional methods to support the interview data, including observations and TAPs.

To address the research questions, a set of pre-planned questions about writing assessment beliefs, practices, and contextual factors was utilized. Refer to Appendix C for the initial interview questions. The interview questions were divided into four main sections: a) teacher background information; b) teacher general writing assessment beliefs and practices; c) teacher specific writing assessment beliefs and practices; and d) teacher context. The first section aimed to gather participants' general personal information and their language learning and teaching backgrounds. This section encompassed inquiries regarding age, teaching experience (both in general and in the PYP at the University), courses taught, language learning experiences, professional training in language teaching and assessment, academic qualifications, as well as a general description of their writing course, students' levels, and job responsibilities.

The second and third sections included questions regarding writing assessment beliefs and beliefs in practice. In the second section, teachers were asked about their general beliefs surrounding writing assessment, such as its purposes, methods, and effectiveness. Additionally, their sentiments towards writing assessment were explored. This section also delved into teachers' overall writing assessment beliefs in practice, including the strategies, tools, and

approaches employed in their writing courses. The third section was specifically designed to uncover specific writing assessment beliefs and beliefs in practice that the previous section may not have addressed. For instance, it examined teachers' beliefs and practices concerning writing assessment processes and tools, such as corrective feedback, rating, and rubrics.

The final section investigated the contextual factors that influence teachers' writing assessment beliefs and practices. It also sought teachers' beliefs on what they think are an ideal context and optimal writing assessment practices. It is important to note that the divisions between these sections were not rigidly defined, as some questions about the context also touched upon teachers' writing assessment beliefs, such as inquiring about potential changes they would make to the rating tools and criteria used in English II and their reasons for doing so. Additionally, certain interview questions were adapted from previous studies, including Ben Hedia (2020).

Before conducting the initial interviews, a pilot test was conducted with an EFL writing teacher from another PYP in another Saudi university. The aim of the pilot was to ensure that the interview questions were clear, coherent, and free from ambiguity and to practice the overall flow of the interview procedure. Following the pilot interview, some adjustments were made to the interview questions. The initial version contained 61 questions, but they were reduced to 43 after the pilot. Redundant questions were eliminated, and some were combined for better efficiency. For instance, during the pilot interview, the participant discussed both rating and feedback practices when asked about general writing assessment beliefs in practice. As a result, the interview was reorganized, introducing the third section to provide additional support if participants did not elaborate on their writing assessment beliefs or beliefs in practice in

sufficient detail. The pilot participant also provided valuable suggestions for revising the wording and sequence of certain questions, leading to further improvements.

Initial interview procedures. Initially, participants who had consented to take part in the study were contacted to arrange a suitable date and time for the interview. Some participants requested additional details about the interview questions, so a list of the interview themes was provided to them. They were also given the freedom to choose their preferred communication medium. All interviews were conducted using FaceTime and audio-recorded via Voice Memos, except for the interviews with Leena and Aya, which were conducted in person within their respective offices. Throughout the interviews, English was the primary language of communication, although participants were given the option to respond in either English or Arabic. It is worth noting that despite the language choice, all participants opted to use English, occasionally incorporating Arabic expressions. While these expressions were not directly relevant to the focus of the study, they included words like ‘يعني’ /jeʕni/, meaning ‘so,’ and Islamic idioms like ‘إن شاء الله’ /ʔin fa:ʔa ʕlla:h/, meaning ‘if God wills it.’ The treatment of these expressions is discussed in the data analysis section below.

During the interviews, efforts were made to establish rapport with the teachers, encouraging natural conversation and providing guidance through questions only when necessary. The interview sessions with participants lasted between 45 and 60 minutes. Class materials such as the course syllabus were collected during this time to aid in understanding the course and guide the observations. Additional teaching materials, including class worksheets, were gathered during subsequent observation sessions.

4.3.2 Classroom Observations

Borg (2006) stated that a fundamental principle in teacher cognition research is that self-report methods of teacher cognition cannot be used as measures of actual practices. Verbal data are limited and needs classroom observation to complement participants' narratives of their beliefs, as people may claim to be doing one thing but are actually doing something else (Patton, 2002). Additionally, as Speer (2005) cautioned, "it is quite plausible that there are situations where teachers state beliefs that are (intentionally or unintentionally) inconsistent with what they carry out in their classrooms," possibly because of "incomplete or inaccurate understanding of terms and descriptions used by teachers and researchers" (p. 371). Consequently, classroom observation plays a crucial role in gaining insight into teachers' beliefs, allowing for a direct examination of their actual behaviors (Li, 2017). It also allows for examining teachers in real teaching moments and not relying on the possibility of teachers reporting an ideal teaching situation for them (Speer, 2005).

Borg (2019) acknowledged that recent studies in teacher cognition research often incorporate classroom observations as a valuable means to understand and examine the connection between teachers' thoughts and actions. Observation serves as a complement to verbal data such as questionnaires, TAPs, and interviews in investigating the relationship between teachers' beliefs and practices (Alzaanin, 2019; Burri et al., 2017; Liviero, 2017; Tajeddin & Aryaeian, 2017; Yu et al., 2020). Therefore, classroom observations were employed in this study alongside the data collected from the interviews and TAPs. The primary objective of this method was to explore teachers' writing assessment practices and their interplay with their beliefs, beliefs in practices, and contextual factors, as the observation of practices serves as "the manifestation of beliefs" (Hofman & Seidel, 2015, p. 108).

Classroom observation procedures. Before the trimester began, I reminded the teachers about the upcoming observation phase and informed them of my intention to observe classes during weeks 3 to 9, as shown above in Table 4.3. Before each observation, I contacted the respective teacher in the morning to confirm my attendance. Moreover, adhering to ethical guidelines, I distributed consent forms, detailed in Appendix D, to the students, which outlined the use of voice recording during the classes. Notably, no student declined to be recorded or to participate in the observed classes.

Borg (2006) listed several methodological and procedural dimensions of observations, which were considered in this phase of the study: (1) the observer role (participant/nonparticipant), (2) the authenticity of the observed setting (real/contrived), (3) the awareness of the purpose of observation by the observed (full/minimal), (4) the structure of the observation (structured/unstructured), and (5) the way the observation is recorded (manually/technologically).

Firstly, I adopted the role of a nonparticipant observer, positioning myself at the back of the classroom and taking notes without engaging in any interactions with the teachers or students. This approach aligns with Borg's (2006) assertion that nonparticipant observations are more appropriate for studying teachers' beliefs and practices, as they minimize the impact of the researcher's presence on the teachers' behaviors. To ensure clarity in my role, I explicitly informed the participating teachers that I would be assuming the position of a nonparticipant observer, thereby avoiding any direct interactions or exchanges (Borg, 2006). Throughout the observation period, neither the teachers nor the students directed any communication toward me.

Secondly, I conducted observations in authentic writing classes where teachers used materials and teaching methods from their regular writing course curriculum—observing these

genuine classes aimed to capture natural practices without introducing unfamiliar or artificial situations that could hinder the authentic manifestation of teaching practices (Borg, 2006).

Thirdly, prior to the observations, participating teachers were provided with a general overview of the study's aims and methods but were not provided with specific details regarding the focus of the observations. This was done to avoid biasing their writing assessment practices, ensuring that their behaviors and instructional choices during the observed classes remained unaffected by preconceived notions or expectations (Borg, 2006).

Fourthly, I employed qualitative unstructured observation, also known as open observation, for this study. Unstructured observation involves collecting detailed information about the events being studied through field notes and audio or video recordings, which are then used to generate narrative observation transcripts (Borg, 2006). Several factors influenced the selection of unstructured observation for this research. Firstly, structured or closed observations rely on predetermined categories, potentially leading to the oversight of events or behaviors that are not included in the structured observational checklist. Secondly, given the qualitative nature of my study, unstructured observations are well-suited for qualitative analysis (Borg, 2006). Thirdly, Clark and Leat (1998) argued that predicting behaviors can be challenging, making unstructured observation a more flexible approach to capture a wide range of writing assessment practices employed by teachers. Lastly, as the study utilized the framework of CT, it was essential to maintain an open approach to observe the entire process of writing classes, enabling exploration and consideration of emergent assessment practices and events (Larsen-Freeman & Cameron, 2008) without pre-identifying specific writing assessment categories.

Fifthly, I used Voice Memos to audio-record the observations. The use of audio recordings allowed me to document the activities and occurrences within the classroom,

providing the advantage of replaying them multiple times during the analysis phase (Cohen & Crabtree, 2006). I placed the recorder on a nearby table, close to the teacher, to ensure clear recording. Borg (2006) noted that while audio recordings have limitations in accurately capturing events, they can be supplemented effectively with field notes. Therefore, I maintained field notes as an integral part of the observational record. The field notes were used to document factual information such as dates, times, settings, and actions and capture moments that audio recordings alone could not fully capture. These field notes proved invaluable in aiding my recollection of those specific moments during the analysis process (Borg, 2006). For instance, in certain classes, periods of silence were observed as the teacher monitored students' progress in completing writing task outlines. Notably, this study did not use video recordings, as the participants consented solely to audio recordings.

To facilitate the organization and structure of field notes, I developed a set of guidelines to direct my observations and note-taking process, incorporating the “what, how, and who” questions (Cohen et al., 2007). These guidelines, adapted from Cohen et al. (2007), were divided into five distinct sections: the physical environment, teaching procedures, interactions, writing assessment focus, and additional notes. Refer to Appendix E for the field notes guidelines. The first section of the guidelines provided general information about the class, including details about the classroom layout and timing. The second section focused on capturing the materials, tools, and instructional strategies employed by the teacher during the lesson and tasks. The third section centered on documenting teacher-student interactions, encompassing various types of interactions such as greetings, information dissemination, and questioning. The fourth section specifically directed observations toward the teacher's writing assessment practices and areas of emphasis. Lastly, the final section served as a space for additional reflective notes, where I

recorded any questions or comments that could potentially guide further inquiries during the subsequent follow-up interviews. For an example of observation field notes, please refer to Appendix F.

4.3.3 Think-aloud Protocols

TAPs, which are an introspective data collection method borrowed from cognitive psychology, serve as a means to capture participants' thought processes (Ericsson & Simon, 1993). In the field of writing assessment, TAPs have been widely utilized in various studies (e.g., Cumming et al., 2002; Davison, 2004; DeRemer, 1998; Erdosy, 2004; Lumley, 2002; Smith, 2000; Weigle, 1999). These studies have employed TAPs to evaluate the decision-making processes of raters (Cumming et al., 2002; Lumley, 2002), explore the impact of raters' backgrounds and professional experience on scoring processes and criteria (Erdosy, 2004), develop models of the holistic scoring process (Sakyi, 2000), and investigate the effects of training on rating processes (Weigle, 1994).

Although the aim of this study differs from those mentioned above, the use of TAPs can provide a more comprehensive and nuanced understanding of teachers' writing assessment belief systems and their manifestation in practice. To enhance the trustworthiness of the study, both TAPs and follow-up interviews were employed to obtain a more comprehensive and accurate understanding of the cases, exploring diverse aspects of writing assessment beliefs and practices, as well as the complex relationship between these beliefs, practices, and ecological factors. It is important to note that rating and providing feedback play crucial roles in writing assessment practices, particularly within the Saudi PYP context, where a testing culture is prevalent. While observations were valuable, they were insufficient to fully comprehend these assessment practices, as many occur beyond the classroom confines. Therefore, the integration of TAPs and

follow-up interviews served as valuable tools for addressing the research questions and complementing data from the initial interviews and observations.

To achieve the objective of my study, I employed concurrent TAPs. Concurrent TAPs involve participants verbalizing their thoughts in real-time as they engage in a task (Ericsson & Simon, 1993; McKay, 2006). The rationale behind using concurrent TAPs is their ability to provide more detailed insights into participants' decision-making processes compared to retrospective protocols, which require participants to provide delayed explanations (Ericsson & Simon, 1993; Kuusela & Paul, 2000). Although concurrent TAPs may not capture every thought occurring in participants' minds, they offer valuable expressions or indications of their ongoing cognitive processes (Erdosy, 2004), as well as examples of their actual grading and feedback actions.

Barkaoui (2011) emphasized the importance of using TAPs cautiously and addressing the factors that can impact TAP data collection and analysis. In order to mitigate the limitations associated with TAPs, I implemented several methods. Firstly, I provided a verbalizing training to the teachers to address the issue of incompleteness, as thinking aloud is not a natural process (Dörnyei, 2007). To facilitate this, I developed a TAP training guide based on suggestions by McKay (2006). The TAP guidelines comprised two main sections: instructions and training. The instructions section outlined the definition of TAPs and clarified the roles of both the participant and the researcher. The training section involved the participants solving four mathematical equations while practicing verbalizing their thoughts. The TAP training guide can be found in Appendix G. I observed that the training was beneficial for the participants in grasping the concept of thinking aloud. For instance, Aya sought clarification on whether she was doing it correctly during the training. This presented an opportunity for me to explain that rather than

focusing on being right or wrong, she should concentrate on expressing her thoughts as if she were alone, without addressing her speech to me while assessing the students' papers.

Secondly, to enhance the validity of the TAP data, I conducted follow-up interviews that involved retrospective questioning. These interviews further validated and clarified the participants' thoughts and actions during the TAPs, as Erdosy (2004) recommended. The procedures for conducting these interviews will be discussed in detail in the subsequent section. Additionally, considering that TAPs take place within a social discourse and can be influenced by social and interactive factors such as audience awareness (Barkaoui, 2011; Sasaki, 2003), I minimized my interference during the TAPs to encourage participants to verbalize their thoughts as if they were alone. Only on certain occasions when a participant remained silent for an extended period did I provide gentle encouragement for them to continue thinking aloud, following a suggestion by Bohn-Gettler and Kendeou (2014).

TAPs procedures. The TAP procedures required teachers to evaluate students' responses to two paragraph writing tasks, task 1 and task 2. It is worth noting that other tasks mentioned in Table 3.2 were not considered in the TAPs, as they were indirect assessments of writing, receiving less emphasis in the assessment, and no feedback was provided for such tasks, such as the MCQ paraphrasing task (see Figure 3.1 and 3.2 for a copy of the MCQs in online and paraphrasing tasks).

The first task involved writing a compare and contrast paragraph, where students were required to compose a paragraph of 200–250 words following a multi-step writing process, which included brainstorming, outlining, drafting a first draft, and then a final draft. The second task, a definition paragraph of 200–250 words, followed the same process. For both tasks, students needed to construct a topic sentence, three main points, and a concluding sentence. The

first task took place during weeks 4 and 5, while the second was conducted in weeks 7, 8, and 9. For a copy of the descriptions and instructions of the two tasks provided to the teachers and students, see Appendix H. Additionally, Table 4.4 presents a concise description of the tasks assessed during the TAP sessions.

Table 4.4

A Brief Description of the TAP Sessions Tasks

TAP Session	Task Name	Task Weeks	Task Description
Session 1	Task 1: Compare and Contrast Paragraph (200-250 words)	Week 4: Brainstorming & Outline [50 min in class] Week 4\5: First Draft [at home] Week 5: Submission of Final Draft	Students are expected to produce a well-structured paragraph (comparing and contrasting two elements) making use of pre-writing and a multi-step writing process. Students should produce an appropriate topic sentence, well-developed supporting points, and a clear conclusion using a variety of grammatical structures, academic language, discourse markers, and contextual devices appropriate to the genre.
Session 2	Task 2: Definition Paragraph (200-250 words)	Week 7/8: Brainstorming & Outline [50 min in class] Week 7/8: First Draft [at home] Week 9: Submission of Final Draft	Students are expected to produce a well-structured definition paragraph, making use of pre-writing and a multi-step writing process. Students should produce an appropriate topic sentence, well-developed supporting points, and a clear conclusion using a variety of grammatical structures, academic language, discourse markers, and contextual devices appropriate to the genre.

In the first TAP session, participants were asked to assess three student papers responding to the compare and contrast paragraph task, using their standard rating criteria and feedback routines. In the second TAP session, participants assessed three papers responding to the definition paragraph task. Thus, each teacher evaluated a total of six papers across the two TAP sessions. Considering that TAPs were not the sole measure of teachers' assessment beliefs and practices; six student papers were deemed representative of each teacher's typical assessment methods and sufficient for data collection. Additionally, ten paragraphs rated by the teachers for

each task, including those in the TAPs, were collected, totaling 20 documents per participant. All participants agreed to complete two TAP sessions, with the exception of Rana, who participated in only one. Table 4.5 outlines the timeline and number of TAP sessions for each participant.

The participants reported using a standardized rubric for rating these writing tasks, as described in Chapter 3. Analysis of the teachers' initial interview responses indicated that students were required to submit two drafts for each task: a first draft and a final draft. According to the task instructions and policies, teachers were obligated to provide written feedback on the first drafts. Regarding the final drafts, the policy stated, "Teachers are not required to give additional feedback on the final draft. However, teachers can use their own discretion." Consequently, I chose to conduct the TAPs on the first drafts, as all teachers reported that they were mandated to provide both written feedback and a grade based on the rubric. This provided a richer opportunity to examine their writing assessment beliefs and practices.

Table 4.5

The Data Collection Timeline of TAPs and Follow-up Interviews

Pseudonym	Number of TAPs & FI	First TAP & FI	Number of Paragraphs Assessed	Number of Samples Collected	Second TAP & FI	Number of Paragraphs Assessed	Number of Samples Collected
Leena	2	Week 4	3	10 papers	Week 8	3	10 papers
Aya	2	Week 4	3	10 papers	Week 9	3	10 papers
Batool	2	Week 4	3	10 papers	Week 8	3	10 papers

Noor	2	Week 4	3	10 papers	Week 7	3	10 papers
Rana	1	Week 4	3	10 papers			10 papers

After each observation, I reminded the participants of the impending TAPs and follow-up interviews. I asked that they notify me when they planned to correct their students' first draft responses to schedule the TAP sessions accordingly. To preserve the natural context of the assessment, participants were allowed to choose both the location and timing for the TAPs. Furthermore, they were offered the option to verbalize their thoughts in either Arabic or English, to ensure a natural expression of their rating processes. All participants primarily used English for verbalization during the TAPs, with occasional Arabic phrases, consistent with their language use during the initial interviews.

I started the first TAP session with the training guidelines, while the second session commenced directly with the task. In each session, I asked the teacher to select three papers for assessment, as per their normal routine. All TAPs were conducted in the teachers' offices, except for those with Aya and Rana. Aya chose to meet in a coffee shop, while Rana opted for a video call on FaceTime. The sessions were audio-recorded using Voice Memos, with teachers' consent for audio-only recording. Each TAP ranged from 25 to 45 minutes in duration.

During the TAPs, most teachers used Microsoft Teams to review the students' first draft submissions, which were uploaded in .docx format. Feedback was provided using the comment features in Microsoft Word or Teams. While the TAPs were in progress, I took concise notes on

the teachers' decision-making processes and actions, which assisted in the follow-up interviews and subsequent data analysis.

4.3.4 Follow-up Interviews

This study employed follow-up interviews as a supplementary method to classroom observations and TAPs. Borg (2006) observed that it might be challenging for teachers to simultaneously teach and articulate their cognitive processes, making retrospective verbal accounts necessary to examine their interactive thinking. Follow-up interviews after observations are widely used in language teacher cognition studies, providing teachers the opportunity to reflect upon and elaborate on their observed practices (e.g., Borg, 1998; Phipps & Borg, 2009). Likewise, other studies have utilized follow-up interviews after TAPs to contextualize and deepen the understanding of the TAP findings (e.g., Bustamante & Yilmaz, 2020; Charters, 2003). Charters (2003) posited that retrospective questioning to clarify and expand upon TAP results can yield further insights into participants' cognitive processes. Charters also noted that conducting follow-up interviews enables participants to confirm or validate the researcher's interpretations of their verbalizations during the TAPs.

Therefore, the follow-up interviews included both post-observation and post-TAP questions. The primary objective of these interviews was to probe deeper into the participants' thoughts and obtain further details about their beliefs and practices related to writing assessment. While the initial interviews aimed to explore the participants' professed writing assessment beliefs and beliefs in practice, the questions during the follow-up interviews were designed to facilitate a more in-depth understanding of the complex relationship between these reported beliefs, actual practices, and contextual factors.

Follow-up interview procedures. I scheduled the follow-up interview with each teacher immediately after each TAP session. The interviews were semi-structured in nature and incorporated questions related to both the TAPs and the classroom observations. To ensure a systematic approach, guided questions were used, as outlined in the follow-up interview guidelines in Appendix I. The guidelines were divided into two sections: post-TAP session questions and post-observation questions.

While most of the questions were predetermined, additional questions related to the TAP sessions and observations were included based on each teacher's practices. I wrote down some post-observation questions during the observations to understand specific practices carried out in the classroom, for instance, asking the teacher about her reasons for discussing the rubric with her students in the class. Additionally, post-TAP session questions aimed to explore the teacher's thought processes for each paper she rated, such as asking the teacher why she thought a student sought help completing her paragraph writing task.

4.3.5 Documents

In case-study designs, documents can be used besides other methods for triangulation (Punch, 2009, p. 159). The present study incorporated document analysis as an adjunctive data source alongside interviews, observations, and TAPs to enrich the data collection methods. The collection of documents was an ongoing process throughout the data collection phase. I obtained the necessary permissions from the teachers and the Head of the Department to access and collect all the required documents.

Initially, I acquired the University policy, PYP study plan, and English II course syllabus. These documents detailed the specific study plan, course objectives, assessment methods, processes, tasks, weekly schedules, teaching topics, and assessment criteria, all of which were

vital for grasping the context prior to data collection. They also facilitated the composition of a detailed description of the study context (see Chapter 3). Furthermore, after conducting classroom observations, teaching materials were collected from the participating teachers. Subsequently, students' rated papers were gathered following each TAP session. These evaluated papers provided direct evidence of various aspects of teachers' rating and written feedback practices, such as feedback types and focus areas, complementing data from the TAPs. In total, I collected 20 examples of each teacher's assessments of students' writing tasks. The documents shared by the teachers varied in format. Some provided screenshots from Teams, while others downloaded the Word-formatted paragraphs and shared them with me by email. To enhance the readability of the screenshots and certain documents, they were converted to PDF format to consolidate all comments into a single view.

4.4 Data Analysis

The wide range of qualitative analysis approaches highlights that there is “no single right way to do qualitative data analysis;” the researcher needs to choose an appropriate analysis method based on the study's purpose (Punch & Oancea, 2014, p. 219). From a CT perspective, nonlinearity in interactions between agents causes fundamental uncertainty about how things will unfold (Larsen-Freeman, 2013). Dealing with CASs implies that numerous outcomes are possible and emergent depending on how the agents interact with and respond to contextual forces. Therefore, I implemented an exploratory open and inductive data analysis approach so themes could emerge from the data (Scott & Usher, 2010). The data that were analyzed qualitatively were initial interview transcripts, observation transcripts, TAP transcripts, follow-up interview transcripts, and documents. The process of thematic analysis consisted of reading each line of data, assigning codes and labels, identifying themes and sub-themes, and comparing

the data repeatedly within and across data sources (Nowell et al., 2017). According to Creswell and Poth (2018), in qualitative research, collecting and analyzing data and writing reports on data are not always clearly separate phases, as they frequently overlap and co-occur throughout the research process.

I began the data analysis process by conducting a preliminary document analysis before the first data collection method (initial interviews). As described in the data collection section, I reviewed the University policy statement and writing course syllabi to better understand the PYP context and its writing courses. Following each data collection phase, I transcribed the audio recordings promptly. For instance, I transcribed the initial interviews immediately after conducting them to ensure accurate recall of the teachers' responses and to clarify any uncertainties in their articulation. However, the main data analysis did not commence until all data collection was completed. Each case study was thoroughly reviewed and completed before moving on to the next one, except for transcription, which was an ongoing process.

I used data analysis software instead of manual analysis to ensure consistency and increase efficiency in handling the large dataset. According to King (2004), utilizing data analysis software is an efficient way of dealing with complex coding schemes and large volumes of text, which helps to facilitate in-depth and sophisticated analysis. Therefore, I used MAXQDA—a versatile software system designed for organizing, managing, and analyzing data.

4.4.1 Considerations for Analyzing Complex Adaptive Systems

Throughout the analysis process, I considered three substantial recommendations made by Larsen-Freeman and Cameron (2008, pp. 241-242) for analyzing data using CT. The first recommendation emphasizes the necessity of establishing boundaries for research using CT as its theoretical framework. Therefore, I established four main boundaries for analyzing and

discussing the case studies as follows. Firstly, I primarily focused on EFL teachers' beliefs and practices related only to L2 writing assessment. If other beliefs, such as those concerning teaching writing and students, were pertinent to writing assessment beliefs and emerged during the investigation, they were also analyzed and discussed. Secondly, I distinguished between teachers' writing assessment beliefs, beliefs in practice, and actual practices. *Beliefs in practice* refer to any belief that teachers reported about their actions or intended actions in writing assessments, and they are considered part of the teachers' belief systems (Zheng, 2015). This distinction was pivotal in clarifying the differences between reported beliefs (that is, beliefs and beliefs in practice) and actual practices. It afforded me the opportunity to discern between what teachers perceived they were doing and what they were actually implementing.

Thirdly, as indicated above, I adopted an ecological perspective to examine the contextual factors affecting teachers, which include the macrosystem (social and national factors), the exosystem (institutional and PYP factors), and the microsystem (direct personal and classroom factors). This approach enabled a clear delineation of these factors to facilitate discussion of their interplay. However, my primary focus was on ecological factors that interact with teachers' beliefs and practices in writing assessments. For instance, additional contextual factors that affect classroom practices or students, such as technological affordances or constraints, were discussed only when they directly impacted or interacted with the teachers' writing assessment beliefs and practices. Lastly, after engaging with the data, I established clear parameters that reflected the most pertinent areas for a comprehensive discussion of the participating teachers' writing assessment belief systems. I focused my analysis on five fundamental categories of beliefs that seemed to define the teachers' framework of writing assessment: (a) beliefs concerning the teaching and learning of writing; (b) beliefs regarding the purposes of writing assessment; (c)

beliefs about what to assess, encompassing evaluation criteria; (d) beliefs about how to assess, including assessment methods; and (e) beliefs regarding assessment processes and tools, including feedback, grading, and rubric use. These categories emerged from the data rather than being predetermined.

Larsen-Freeman and Cameron (2008) made a second recommendation that was also considered. They advised against reductionism by acknowledging the complexity of any case study, avoiding preconceived notions, and considering all possible factors. An open inductive approach to data analysis was employed, ensuring receptivity to unforeseen interactions and factors potentially affecting teachers' writing assessment beliefs and practices. For example, considering how teachers approached the interview questions differently was discussed when deemed enriching and influential in reporting a case's findings. The final recommendation is to think in terms of processes and changing relationships among variables. This was achieved by exploring teachers' writing assessment beliefs from a CT perspective, which involved examining and metaphorically using notions like *systems, attractors, heterogeneity, nonlinear interactions, openness, and adaptation*.

4.4.2 Transcription, TAPs Initial Data Analysis, and Translation

Despite the time-intensive process of creating detailed verbatim transcripts (Dörnyei, 2007), transcribing the interviews, observations, and TAPs was essential for a comprehensive understanding of the teachers' beliefs and practices. I produced naturalized transcripts by including as many details as needed, such as repeated words, fillers (e.g., 'like,' 'um,' 'uh'), and non-verbal behaviors (e.g., laughter, pauses, and sighs), to convey the teachers' expressions of their beliefs and beliefs in practice as authentically as possible. This approach aimed to provide deeper insight not only into the teachers' beliefs but also into how these beliefs were articulated

through their verbal reports. The same transcription procedure was applied to each participant. However, for the findings' chapters, some repeated words and fillers were omitted to enhance reader comprehension, retaining them only when they were crucial to the excerpt discussion.

The analysis of the audio-recorded initial and follow-up interviews began with transcribing the data. To transcribe the interviews, I initially used *Descript*, a Mac application that offers automated transcription, to convert the audio recordings into text. Subsequently, I carefully listened to the audio recordings while reviewing the corresponding transcripts. During this process, I revised the transcripts, translated any Arabic words, and incorporated transcription conventions. Some transcription conventions were adapted from Seedhouse (2004), while others were created specifically for this study. Refer to Table 4.6 for the transcription conventions.

The transcription process for the audio-recorded observations followed a similar approach to the interviews. I used Descript to transcribe the observation recordings and revised the transcriptions to include the necessary conventions. To aid in understanding the transcribed observations, I referred to the handwritten field notes as a guide.

Table 4.6

Transcription Conventions (Adapted from Seedhouse, 2004)

Transcription Convention	Description	Examples from the Study
P:	Participant in the interviews and TAP sessions	P: [Reads student paper] 'There are several significant'
T:	Teacher in class	T: You chose C, right?
L:	Learner in class	L: We can't add our opinion.
LL:	Learners in class	LL: Yes.
(Pause)	Long pause of more than 2 seconds	T: we are gonna explain the questions together. (Pause) What questions did you struggle with? (Pause)
?	An indication of a question including the use of falling intonation	P: I think it's like, what do you call it? Middle school?
...	An indication of intentional omission by the researcher to shorten the Excerpt when redundant sentences or meanings were	P: <u>Avoid</u> the traditional way. Because it will <u>kill</u> the students... I mean by traditional way, like, when

	present. Also, an indication of insertion of further clarification by the participant after further questioning.	the teacher only teaches the content, and the students <u>do nothing</u> .
(...)	An ellipsis between parentheses indicates a stretch of unclear or unintelligible speech.	L: A very special (...).
<u>word</u>	Underlined words indicate the speaker's emphasis by stretching out the vowel sound, speaking slower when saying the word, or pronouncing one syllable louder than the others.	P: It's not like <u>teaching</u> . Yeah, I mean I didn't have teaching responsibilities like grading the students.
“ “ Or ‘ ‘	An indication that the participant talks about what people told her or what she told them, for example, students. Also, an indication of the participants reading students' sentences in TAP sessions.	P: However, even when <u>I</u> told them that “you would <u>need</u> writing,” they told me, “Ms., the progress note is not academic writing”
/	An indication of an incomplete sentence or utterance	P: So, okay, so it's just like four/ if we want to talk about the writing specifically, it's just like three years.
[[]]	The researcher's interpretation or explanation of the situation or utterance. The interpreted segment is italicized and immediately followed by the interpretation between double square brackets.	P: Yeah. Like, <i>I try to give/ like, for these common mistakes</i> . [[giving supplementary teaching materials addressing students' common mistakes]]. I will not share with other teachers.
<i>Arabic word</i> ((translation))	Arabic words are italicized and followed by an English translation in double parentheses	P: I'm fine. الحمد لله ((thanks to Allah)). How are you?
{content deleted}	An indication of confidentiality	P: I taught in {university name} for one year?
{Laugh}	An indication of laughing	P: Red, red color. {laughs}
{Sigh}	An indication of sighing	P: I couldn't spot this one. {Sigh} Was it about pronoun reference or/ (Pause)?

TAPs initial data analysis. I also used Descript and the same procedure to transcribe the audio-recorded TAPs. Then I listened to the audio recordings and revised the transcripts. TAP transcripts are more fragmented and complex to understand than normal speech as they are not meant to be communicative to anyone but the thinker (Charters, 2003). Thus, as a further analysis step, I categorized the TAP transcripts into initial codes before starting to code all data sources. I adapted Cumming's (2002) framework of raters' decision-making behaviors to code the fragmented speech in the transcripts. The main purpose of those categories or codes was to

facilitate the understanding of the TAP transcripts during the analysis when differentiating between the teachers' various assessment beliefs, rating focus, and actual practices. Refer to Table 4.7 for the TAP initial codes.

Two main behaviors were identified: interpretation strategies (IS) and judgment strategies (JS). Each of these strategies could have a self-monitoring focus (SMF), rhetorical focus (RF), or language focus (LF). Under IS, SMF strategies relate to the participants' reading and rereading the texts. RF strategies involve discerning the rhetorical structure, scanning the whole composition, or observing the layout. LF strategies center on classifying errors into types, interpreting, or editing ambiguous or unclear phrases. Under JS, SMF strategies refer to the participants' consideration of their personal responses or biases, defining assessment criteria, articulating general impressions about the text, and revising rating or feedback decisions. RF strategies involve assessing reasoning or topic development, task completion, coherence, originality, text organization, style, or register. LF strategies refer to the participants' assessment of language, which include assessing the quantity of the whole text as well as considering lexis, syntax, morphology, spelling, or punctuation (Cumming, 2002). To illustrate the implementation of these codes, I included several examples from TAP transcripts in Table 4.7.

Table 4.7

TAPs Initial Coding (Adapted from Cumming, 2002, p. 88)

Codes	Description	Examples from the TAPs Initial Coding
Interpretation Strategies		
[IS-SMF]	Interpretation strategies with a self-monitoring focus, indicate: <ol style="list-style-type: none"> 1. reading or rereading composition. 	P: "Eating at home is not the same as eating at restaurants. <u>They</u> have three main differences" [IS-SMF].
[IS-RF]	Interpretation strategies with a rhetorical focus, indicate: <ol style="list-style-type: none"> 1. discerning the rhetorical structure 2. scanning the whole composition 3. observing the layout 	P: " <i>It is more suitable because doesn't need/ they can eat at any</i> " [IS-SMF] [[rereading and skipping sentences]]. Okay. I think this is the first main point [IS-RF].

[IS-LF]	Interpretation strategies with a language focus, indicate: <ol style="list-style-type: none"> 1. classifying errors into types 2. interpreting or editing ambiguous phrases 	P: ...[IS-SMF]. طيب ((OK)) ‘sometimes’ should come after the verb. So, um. What can I call this? word order? Yeah. Lexical. Uh, accuracy. word order [IS-LF]. P: ...[IS-SMF]. (Pause) she started using “people” and then “watching new movies. Just they can see them”. Okay, so I think she needs to use “people” again [IS-LF].
Judgmental Strategies		
[JS-SMF]	Judgmental strategies with a self-monitoring focus, indicate: <ol style="list-style-type: none"> 1. considering own personal response or biases 2. defining assessment criteria 3. articulating general impression 4. revising rating or feedback decision 	P: I don’t think that/ this is/ student/ I thought she would write better [JS-SMF]. P: Okay (Pause) <i>Fragment. comma rules. Good 1st main point</i> [JS-SMF] [[rereading her feedback comment from the feedback box on Microsoft Teams]]
[JS-RF]	Judgmental strategies with a rhetorical focus, indicate: <ol style="list-style-type: none"> 1. assessing reasoning or topic development 2. assessing task completion 3. assessing coherence 4. assessing originality 5. assessing text organization, style, or register 	P: (Pause) Okay, I think <u>good</u> first main point and related supporting details [JS-RF]. P: Okay, this also is about coherence, about pronoun, uh, reference [JS-RF].
[JS-LF]	Judgmental strategies with a language focus, indicate: <ol style="list-style-type: none"> 1. assessing the quantity of total written production 2. assessing lexis, syntax, morphology, spelling, or punctuation 	P: So here it’s a comma splice [JS-LF]. P: “Eating at a restaurant <u>differ</u> ” [IS-SMF]. Subject-verb agreement [JS-LF].

Translation. The participants in this study were EFL teachers whose L1 is Arabic and L2 is English. The data were primarily collected in English during all phases of the study, given that its purpose was to investigate the beliefs of EFL teachers regarding the assessment of writing in English as L2. However, since both the researcher and the students are Arabic speakers, some incidental use of Arabic occurred.

The main concern with data translation is the possibility of losing or changing the intended meanings and emphasis of the original text. For instance, metaphors may not carry identical meanings in different languages, and certain elements may lack clear equivalents that

can be easily translated into the target language (van Nes et al., 2010). Consequently, my primary objective during the translation process was to accurately convey the intended meaning of the participants throughout the data. To achieve this goal, two approaches to translation were employed: researcher translation and member checking (van Nes et al., 2010). Before selecting an approach for translation, I carefully examined the potential impact of Arabic use on the meaning of the participants' responses. Researcher translation was deemed appropriate when the use of Arabic was incidental and minimal and did not affect the interpretation or comprehension of the excerpt. As a native Arabic speaker from Saudi Arabia immersed in the culture and colloquial language, my judgment was sufficient for these cases. Examples of such incidental use of Arabic words include teachers calling their students saying, 'بنات' /bna:t/, which has the English equivalence translation of 'girls.' Some participants used Saudi-Arabic discourse markers such as 'طيب' /tʕjɪb/, which functions as a transition to a new idea, similar to the function of 'Okay' in English. Arabic-Islamic idioms also appeared in the data, such as 'والله' /wa:llh/, which is used to emphasize the truth of a statement by using the name of Allah (God).

On the other hand, all participants were bilingual English teachers. Therefore, when a participant used Arabic terms or longer phrases pertinent to the study's content, I translated them and consulted with the participant to confirm the intended meaning. For example, while discussing her grading beliefs, Rana switched to Arabic and said, "Listen, my students, زي مايقولون خط أحمر بالنسبة لي" which translates literally to 'as they say, a red line for me.' When translating such phrases, I verified the intended meaning with the teacher and then researched to determine if a similar expression existed in English. This ensured that my translation accurately reflected whether the phrase should be rendered literally or figuratively, according to its intended meaning.

In addition, the need for member checking varied among participants based on their use of code-switching. For instance, Leena did not use Arabic in her interviews or TAPs but employed Arabic phrases during class observations with her students, besides simultaneously translating them into English. Thus, checking the translation with Leena was not necessary. Seibert (2022) recommended transparency in translation and advocated for including participants' quotes in their original language when reporting. To comply with this recommendation, any excerpts used in the finding chapters include the original Arabic utterance followed by the English translation.

4.4.3 Within-Case Analysis

According to Nowell et al. (2017), researchers are the instruments of analysis in qualitative research. They must provide sufficient information about their documentation, systemization, and analysis techniques to allow readers to assess the trustworthiness of their data analysis. Therefore, this section provides a detailed account of the analysis methods employed in this study.

I followed the six-phase thematic analysis method of Braun and Clarke (2006) (as cited in Nowell et al., 2017). The six phases are: (a) familiarizing myself with the data, (b) generating initial codes, (c) searching for themes, (d) reviewing themes, (e) defining and naming themes, and (f) producing the report. Nowell et al. (2017) noted that it is more accurate to view the phases as “an iterative and reflective process” that requires repeatedly moving back and forth between the different phases (p. 4). Therefore, I used a reflexive journal starting from the first phase, as suggested by Lincoln and Guba (1985). The primary objective of reflexive journaling was to document my thought processes during the analysis phases and deliberate on my coding decisions as I progressed. Additionally, it enabled me to reinforce clear boundaries pertaining to

my theoretical framework, research questions, and coding strategies. It facilitated the meticulous tracking of all inclusions and exclusions across cases to ensure consistency.

Firstly, to familiarize myself with the data, I reread the transcripts of each case before uploading them to MAXQDA for two purposes: to ensure that the documents were correctly named and to become immersed in the data. As I read, I highlighted segments and annotated potential patterns that could inform the coding process (Braun & Clarke, 2006, as cited in Nowell et al., 2017). Subsequently, I uploaded the transcripts to MAXQDA and organized them by creating a dedicated folder for each case, labeled with the teachers' pseudonyms. Each case folder typically contained the following sequence of transcripts: initial interview, first observation, first TAP, first follow-up interview, second observation, second TAP, and second follow-up interview. The 20 samples of teachers' responses to students' written paragraphs, which were indicative of their rating practices, were primarily coded manually since many were in the form of screenshots.

Secondly, the phase of generating initial codes involved deriving codes directly from the data, an activity Nowell et al. (2017) described as "theorizing," where the researcher formulates concepts and insights about the phenomena observed in the data (p. 5). I began coding the documents in each case folder sequentially, starting with the initial interview transcript and concluding with the second follow-up interview transcript, followed by the 20 paragraphs. Initially, developing the codes presented challenges; however, after referring to Saldaña's (2013) coding manual, I gained a clear understanding of the various coding cycles, methods, and the rationale for selecting specific codes.

The inductive approach was the essence of my data analysis. However, Nowell et al. (2017) noted that to start coding the data, it is important to use a consistent and disciplined

approach. Saldaña (2013) also recommended that researchers should consider and explore various coding methods in advance to gain a fresh perspective on the data during each coding cycle. Therefore, generating the initial codes focused mainly, but not exclusively, on two coding methods, values and descriptive coding. Saldaña (2013) described that “*values coding*” focuses on selecting parts of the text that reveal the thoughts, emotions, and opinions of the participants (pp. 110-111). This is particularly helpful when conducting research that delves into beliefs, interpersonal experiences, and actions or when seeking to understand the human experience. “*Descriptive coding*” involves coding qualitative data excerpts with a word or short phrase, typically a noun, to summarize the main topic of the passage (Saldaña, 2013, pp. 87-88). Descriptive coding is helpful for studies with a wide variety of data forms and appropriate for coding social contexts. Hence, values coding was beneficial in coding the interviews and TAPs, while descriptive coding helped code the classroom observations and participants’ contexts and backgrounds. The utilization of these coding techniques facilitated the process of generating initial codes. However, they did not limit or predict the types of codes used (Saldaña, 2013).

The second cycle of coding focused on creating further levels of codes and categories using causation and structural coding methods. Saldaña (2013) explained that the *structural coding* method involves assigning a content-based or conceptual phrase to a data segment about a specific research question. It is appropriate for organizing information by creating lists of key topics, categories, and themes. For example, during this cycle, I kept in mind the focus of my research questions, which were writing assessment beliefs, practices, and ecological factors. In addition, *causation coding* involves the extraction of participants’ attributions or beliefs about the causal factors behind specific outcomes, not limited to the means through which they occurred but also encompassing the reasons for their occurrence. Saldaña (2013) noted that

causation coding is suitable for “discerning motives, belief systems, worldviews, processes, recent histories, interrelationships, and the complexity of influences and effects on human actions and phenomena” (p. 261). Utilizing this method of coding was important for my theoretical framework, CT, as it allowed me to understand the complex interactions between belief, practice, and ecological systems. I also used the MAXQDA feature of vivo coding, using the participant’s own words as a code in places where teachers expressed feelings or used interesting descriptive words.

MAXQDA facilitated the creation of numerous codes and subcodes as required. I utilized different colors for these codes and subcodes to align with specific areas of investigation, which simplified the categorization process. During coding, I also used the ‘memo’ and ‘comment’ features to augment my reflexive journal, documenting descriptions for some codes and subcodes and noting reflections on my thought processes and analytical approach to ensure consistency in the analysis. After completing each coding cycle, I reviewed the codes using MAXQDA’s feature to retrieve coded segments, ensuring no interchangeability or redundancy (Nowell et al., 2017). For a sample of the MAXQDA coding, see Appendix J.

Thirdly, the search for themes commenced once all the data had been coded and organized into initial categories, along with a comprehensive list of the various codes developed throughout the dataset. In this phase, the relevant coded data extracts were meticulously sorted and organized into themes (Braun & Clarke, 2006, as cited in Nowell et al., 2017). A theme is defined as a description of the underlying meaning within the data, which can either be explicitly evident or subtly embedded within the phenomenon it represents. The significance of a theme in relation to the research questions does not rely on quantifiable aspects but rather on its ability to capture something noteworthy (Saldaña, 2013).

The themes identified were data-driven and derived inductively, emerging directly from the analysis, thus maintaining a close connection to the data and the case itself. The primary set of themes pertained to the research questions, focusing on EFL teachers' writing assessment beliefs, practices, ecological factors, and their interaction. First, the data revealed that teachers frequently discussed beliefs and beliefs in practice regarding writing instruction and learning, purposes of writing assessment and its impact on student learning, assessment methods, evaluation criteria, and the processes and tools for writing assessment. Second, the themes related to actual practices centered primarily on classroom writing assessment and rating practices (i.e., feedback and grading). Lastly, themes capturing the interactions between teachers' beliefs, practices, and contextual factors included the influence of factors from the different ecological systems on shaping beliefs and the beliefs-practices relationship, encompassing both tensions and harmonies.

During the fourth and fifth phases of reviewing and defining the themes, I meticulously analyzed each theme in relation to its associated concepts and categories. This process involved naming the themes accurately and refining their wording and descriptions. Moreover, I delved into the unique definition and characterization of each theme within the individual cases to establish a comprehensive theme for each case to understand the teachers' core and peripheral beliefs. For example, this included examining the meaning of writing assessment purposes from the perspective of each participating teacher.

The final phase involved producing the reports. Merriam (1998) noted that "there is no standard format" for reporting case studies data (p. 220). Creswell and Poth (2018) explained that case studies can fulfill various purposes, from generating theory to merely describing cases. Some are more analytical, including cross-case comparisons. The intended purpose of a case

study design significantly shapes the organization and presentation of the final report. My reports were both descriptive and analytical in nature and organized based on my research questions and theoretical framework. Each case report provided an overview of the participant, detailed the elements of her belief system and practices, itemized ecological factors, and highlighted the interactions and relationships between these components. From these individual reports, cases were distinctly characterized by a central theme unique to each participant's core beliefs.

4.4.4 Document Analysis

Various types of documents can assist researchers in revealing meaning, fostering comprehension, and uncovering insights pertaining to their research questions (Bowen, 2009). Documents can also provide a static representation of contents at a given time (Rapley, 2018). Due to the aim of this study, document analysis was necessary for understanding the PYP context, the English II course framework, teachers' writing assessment practices, and the interaction of those practices with their beliefs and the PYP context. Prior (2008) noted that understanding the document contents alone is insufficient to understand how agents establish and implement those documents across complex systems. Consequently, I focused on the documents' functions by examining their positioning and utilization within the context of PYP. The primary approach employed involved exploring and understanding teachers' perspectives and their usage of the documents. For example, I explored how the teachers perceived and used the pre-prepared rubric and teaching materials.

The analysis of documents, including the University policy, PYP study plan, English II syllabus, teaching materials, rating criteria, and teachers' responses to the 20 written paragraphs, served mainly three purposes: (a) to understand the contexts of the program and where the teachers work before collecting data; (b) to explore how the documents as agents of the contexts

influence and interact with teachers' writing assessment beliefs and practices, and (c) to investigate how documents, such as student-rated paragraphs, reflected teachers' writing assessment practices and their interaction with teachers' assessment beliefs.

Document analysis is like other qualitative analytical approaches and necessitates the thorough examination and interpretation of data to attain a precise understanding and extract meaningful insights (Rapley, 2018). Bowen (2009) noted that the analytical process of documents “involves a careful, more focused re-reading and review of the data. The reviewer takes a closer look at the selected data and performs coding and category construction, based on the data's characteristics, to uncover themes pertinent to a phenomenon” (p. 32). The analysis of the documents in this study involved iterative readings of the raw data to extract and interpret the underlying meanings in relation to the dataset of each specific case.

4.4.5 Cross-Case Analysis

The qualitative datasets comprised five cases, each drawing from various sources that required integration. Group averages or norms are not insightful without scrutinizing individual cases, as suggested by Dörnyei (2014). Consequently, I analyzed each case in isolation and wrote a two to four-page case summary addressing my research questions. These summaries encapsulated elements from the teachers' writing assessment belief systems (core and peripheral beliefs), their writing assessment practices, and ecological factors from the macrosystem, exosystem, and microsystem. They also included an overview of the interplay between beliefs, practices, and contextual factors (influences, tensions, and harmonies).

Creating these summaries afforded me the opportunity to refine themes and reassess the findings for each case. Additionally, I developed a visual representation for each case, following the recommendations of Martin et al. (2018), to illustrate the interactions among the teachers'

writing assessment belief systems, their actual practices, and ecological systems. These visuals served as an efficient reference to the case findings and aided in the comparative analysis of the five cases. Writing report summaries and integrating visuals was instrumental in extracting overarching insights from the cases, which is in line with the core objectives of case study research as outlined by Creswell & Poth (2018). The visuals of each case were added to the finding chapters of each case (Chapters 5 to 9).

4.5 Trustworthiness

Richards (2009) discussed that evaluating the effectiveness of any research design involves examining specific criteria. In the case of quantitative research, this evaluation relies on criteria such as validity and reliability. However, when it comes to qualitative research, the use of these criteria to validate such research has been a topic of debate. This debate stems from the potential influence of researcher bias on both data collection and analysis procedures (Creswell & Poth, 2018; Richards, 2009).

As an alternative, *trustworthiness* was introduced by Lincoln and Guba (1985) and used to evaluate qualitative research instead of the quantitative criteria of validity and reliability. It constitutes a set of criteria through which researchers can convince both themselves and their readers that their research results deserve consideration (Lincoln & Guba, 1985). Nowell et al. (2017) integrated Braun and Clarke's (2006) six-phase data thematic analysis method with Lincoln and Guba's (1985) criteria for trustworthiness, which are credibility, transferability, dependability, and confirmability. They also highlighted how researchers might address criteria for trustworthiness during each phase of thematic analysis. Those criteria were considered to ensure the trustworthiness of this study.

Credibility. *Credibility* is “the methodological procedures and sources used to establish a high level of harmony between the participants’ expressions and the researcher’s interpretations of them” (Jensen, 2008, p. 138). This definition underscores the intricate nature of credibility within research. To ensure a satisfactory level of credibility in this study, the primary method employed was the triangulation of data sources. “Typically, this process involves corroborating evidence from different sources to shed light on a theme or perspective” (Creswell, 2013, p. 251). In the current study, triangulation was achieved through the utilization of multiple data collection methods, including interviews, observations, TAPs, and document analysis, which complemented one another. Firstly, interviews delved into participants’ professed beliefs and beliefs in practice, while observations provided an in-depth exploration of actual classroom assessment practices that could not be solely confirmed through verbal reports. Secondly, TAPs expanded the investigation to encompass teachers’ beliefs and practices in assessment areas outside the classroom, such as rating, grading, and providing feedback on students’ writing tasks. Thirdly, follow-up interviews augmented the researcher’s interpretation of observed classroom actions and verbalizations from the TAPs. Finally, document analysis corroborated the researcher’s understanding of the context and offered insight into actual teacher evaluation practices. Furthermore, another form of triangulation was implemented by seeking participants’ input in verifying and validating the English translation of specific Arabic phrases in their transcripts.

Dependability. Nowell et al. (2017) explained that “to achieve dependability, researchers can ensure the research process is logical, traceable, and clearly documented” (p. 3). According to Richards (2009), the analysis must comprehensively elucidate the interrelationship between the chosen methodology and the study’s overarching purpose. Such elucidation should include

detailing the data collection methods, clarifying how these methods were applied to generate the dataset, and articulating the subsequent data analysis procedures. In pursuit of dependability in the present study, I have thoroughly explained the pertinent research procedures. This approach may facilitate the potential utility of the current research in diverse contexts by enabling future researchers to draw upon and replicate the study's methods effectively.

Confirmability. Confirmability pertains to the assurance that the researcher's interpretations and findings stem directly from the data rather than from their personal assumptions. This necessitates clearly demonstrating the steps taken to arrive at conclusions and interpretations. Guba and Lincoln (1989) asserted that confirmability is attained when credibility, transferability, and dependability are achieved (as cited in Nowell et al., 2017). Additionally, Koch (1994) advocated for the inclusion of markers elucidating the rationale behind theoretical, methodological, and analytical choices throughout the entirety of the study, enabling others to comprehend the decision-making process (as cited in Nowell et al., 2017). In this study, meticulous justifications for each stage of data collection and data analysis were provided. Furthermore, to ensure consistency in findings and comprehensive analysis, a peer review or debriefing process was implemented, characterized as "an external assessment of the research process" (Creswell, 2013, p. 251). This peer review process primarily involved my advisor, who reviewed drafts of the data analysis procedures and offered detailed feedback on the interpretation and analysis of each case report. Additionally, the dissertation committee examined and approved the data collection methods before the commencement of the study.

Transferability. This concept pertains to the extent to which the findings of a study can be applied to diverse contexts beyond the specific research setting (Richards, 2009). The term commonly used to describe this is *generalizability*, although it remains a topic of considerable

debate in qualitative research. In this respect, Creswell and Poth (2018) emphasized that qualitative researchers typically exhibit caution when it comes to generalizing findings from one case to another due to the differing contexts of each case. They highlighted that the primary purpose of case studies is to comprehend the complexity of individual cases rather than making broad generalizations (p. 99-100). However, Lincoln and Guba (1985) argued that while researchers may not definitively determine the transferability of results to other situations or contexts, they can provide sufficient information and thick descriptions for readers to assess the applicability of the findings to their context. As a result, this study aimed to maintain transparency by reporting comprehensive details regarding the study methods, analytical processes, and findings.

While the present study specifically examined EFL teachers' writing assessment beliefs, practices, and contextual factors within a single program at a Saudi University, its implications extend to a broader context encompassing other Saudi EFL teachers within different PYPs across various Saudi universities. This transferability is predicated on the assumption that Saudi teachers share common cultural, educational, and linguistic backgrounds, which collectively influence the development of their writing assessment beliefs and practices within comparable educational settings. Furthermore, it was anticipated that the insights gathered from this study would be of considerable relevance to policymakers and academic institutions within the Saudi higher education landscape. Consequently, it is my contention that the findings and implications of the current study hold the potential for a degree of applicability beyond its specific setting.

4.6 Ethical Considerations

The previous section provides an in-depth exploration of the diverse methods employed to ensure the research's high quality. Another crucial aspect that greatly influences the credibility

of any research study pertains to the conscientious consideration and subsequent resolution of ethical concerns. This research project adhered rigorously to the ethical guidelines set forth by York University and the specified Saudi University throughout its entirety. Several ethical considerations were considered to ensure the proper conduct of the study. Before gathering any data, consent forms were distributed among the participating teachers. The participants carefully reviewed and signed the forms upon their agreement to participate in the study. During the interviews, observations, and TAPs, I reassured the participants about the confidentiality of their information and emphasized their right to withdraw from the study at any point. Moreover, permission was obtained from the participants to record the interviews, observations, and TAPs in audio format. It was clearly communicated to the participants that their information would be used solely for research purposes and discussed exclusively with my advisor.

In adherence to the ethical standards established for this research, I ensured the confidentiality of participant data by utilizing pseudonyms in all stages of the study. Documents pertinent to data analysis, such as printed observation field notes, were stored under these assumed names. Additionally, the identities of participants and their students referenced in electronic feedback comments, as well as in Word documents or Teams screenshots, were anonymized, either by blurring or redacting the information where feasible. Electronic data was safeguarded on my password-protected Mac laptop, which prevented access by unauthorized individuals. Following the successful defense of my dissertation, all data will be permanently erased, and any physical documents will be securely destroyed.

To safeguard participants' confidentiality, any information that could potentially compromise participants' identity was removed. For example, I redacted the names of the institutions where the teachers were employed, as well as any specific institutional references.

Moreover, in line with the ethical standards governing research observations, consent forms were distributed to all students to secure their consent for potential voice recordings during classroom observations.

CHAPTER 5: THE CASE OF LEENA:

Skill-based Approach and Assessment of Grammar Accuracy

This chapter, along with the following four chapters, Chapters 5 to 9 present the findings from the five participants: Leena, Aya, Batool, Noor, and Rana. Each chapter is titled after the core beliefs of the respective teachers and highlights the teacher's beliefs in five areas of writing assessment: teaching and learning of writing, purposes of assessment, assessment methods, evaluation criteria, feedback, and grading. The discussion also encompasses the contextual factors from the macrosystem, exosystem, and microsystem that shaped the teacher's beliefs and caused tensions between these beliefs and actual practices. Additionally, it underscores the harmony between the teacher's beliefs and practices. At the end of each chapter, a visual representation and a summary of the specific case findings are provided.

5.1 Background for Leena

Leena, an English lecturer in her early thirties, has worked at the University for six years at the time of data collection. She holds a bachelor's degree in English Language from a Saudi university and a master's degree in TESOL from the US. With almost eight years of English teaching experience, her career spans such roles as a two-year teaching assistant at another university and a lecturer for six years at the University's PYP. As a teaching assistant, Leena initially supported other instructors, assisting with courses like reading, writing, and sociolinguistics without directly teaching or assessing students.

As a lecturer at the University, she has taught grammar and reading and focused on writing skills within the last three years at the PYP. During the data collection, she led a writing class with 22 students from the nursing section, meeting them for three hours weekly to deliver the writing portion of English II.

Leena's English learning began with basic education in Saudi public schools, where English was not prioritized, starting only in middle school. Unlike her peers, she did not engage with English outside of school through movies or music. Her structured English learning commenced with her undergraduate studies, where English and writing skills were integral to her major. She recalls the rigorous academic demands of writing assignments, research papers, and essays that shaped her language proficiency, explaining,

Excerpt 5.1:

As you know, like, in all of the four years or the five years, we had to, like, submit writing assignments. We had to write in English, and we had to use perfect sentence structures. I remember we had to write research papers. Sometimes, article reviews. When we were asked/ we were required to write essays. (Leena, II)

5.2 Teaching and Learning Writing: A Skill-based Approach

Leena's core belief in teaching and learning writing appeared to be anchored in a skill-based approach, viewing writing as a technical and instrumental act. She conceptualized writing literacy as a collection of discrete skills, techniques, or strategies that students could acquire and utilize. For Leena, writing was a tool or instrument for achieving specific objectives, such as fulfilling academic or professional requirements, rather than for its intrinsic qualities. The practical utility of writing was seen as its primary value.

This technical perspective of writing was evident in Leena's interview responses, as she emphasized the importance of using written models and exploiting resources such as phrase banks to build vocabulary. Leena expressed this belief clearly,

Excerpt 5.2:

With writing, you can teach them, like, specific strategies, techniques. And I always tell my students this thing, and I always believe in this, "writing is just like techniques, chunks. If you learn them, *you can go* [[you can write]]." Especially academic writing, it's not like creative writing. If you/ for example, for this course, for the topic sentence, you can follow the models in the textbook. And even if they don't have يعني ((like)) perfect English language/ if they are smart enough, and they can, like, read and analyze the models, follow the models, they can write. And for me, this is something that I myself use as a technique. I like the phrase banks. I think it's the Manchester phrase bank. And there's another university. I like them. They help. And also, my students, يعني ((like)) it helps my students, it's not only about me. So, teaching writing is about techniques. Teaching them techniques or strategies. (Leena, II)

Leena seemed to believe that academic writing stands apart from creative writing, highlighting that academic writing requires specific techniques. Leena suggested that students, even those without sufficient English language skills, could write effectively if they were able to analyze and follow the models in their course books. Moreover, Leena believed that students could incorporate vocabulary from the reading course into their writings. She also advocated for a change in the current writing course textbook to introduce a book that provides comprehensive lists of academic words for students to use. She explained,

Excerpt 5.3:

We need to use another book that, like, has some/ مثلاً الـ ((for example, the)) academic word list. It is so important in academic writing. I feel that the textbook that we should give the students/ for example, at the end of each chapter, we should include some academic words. So, they [[students]] can use these kinds of words. (Leena, II)

Moreover, the skill-based approach typically involves explicit instruction, guided practice, and opportunities for students to apply and refine their writing skills through various writing tasks and exercises. Leena believed writing proficiency could be developed through deliberate practice and targeted instruction. She underscored the role of practice in learning to write, stating, “[the] key factor in any writing course is practice. If the students don’t practice, they will not be able to write” (Leena, II). Leena believed in emphasizing the importance of practice to her students, reinforcing the idea that regular practice is crucial for improving writing skills. She explained,

Excerpt 5.4:

I try to, like, emphasize that “even if you are nurses, you will need to write a lot. Like even in your last university years, you will be asked to write a research paper, so you need to understand and practice writing an academic paragraph or an academic essay.” (Leena, II)

In Excerpt 5.4, Leena highlighted the importance of writing in various professional contexts, such as nursing. She stressed the necessity for her students to practice constructing academic paragraphs and essays to prepare them for the writing proficiency required in their later university years and professional lives.

5.3 Writing Assessment Purposes: Direct Assessment Methods of Writing

Leena's response to the question about the purposes of writing assessment veered into discussing the most effective methods to assess a student's performance in the English II course within the PYP framework. She indicated that formative assessments aimed to give students opportunities to practice their writing. This belief seemed to be an extension of her belief in a skill-based approach to writing, where students must "practice writing." She then elaborated on the efficacy of involving students in direct writing tasks, explaining,

Excerpt 5.5:

I believe that, in the writing course, students need to focus mainly on writing assignments. But you know, because this course is not only about writing, it's reading and writing. So, we have, like, other assessments...for example, analyzing a paragraph, finding the topic sentence, all are reading skills.... But for me, for writing, I believe that focusing on writing assignments is the most important thing to assess the student's performance. Other than that, like, the multiple-choice questions are not good/ effective. (Leena, II)

In Excerpt 5.5, Leena's preference appeared to lean towards assessments, where students actively write paragraphs or essays rather than selecting answers from multiple-choice questions in final or midterm exams. She criticized the PYP curriculum implementing MCQ assessment in writing (see Figure 3.3. for a copy of MCQs from the quiz), suggesting that they do not provide insight into a student's understanding or reasoning behind choosing an answer: "I don't think that they are effective. Because students might choose this one, but, like, they don't understand why or the reason for choosing this specific one" (Leena, II).

Even when acknowledging the course's integrated nature, including reading and writing components, Leena maintained that assessments should not consist of multiple-choice questions. She argued for the inclusion of written components in quizzes and final exams, emphasizing that "it should actually include some writing" (Leena, II). Her rationale was that the inclusion of writing tasks in assessments, even within an integrated course, was essential to ensure that the assessment accurately reflects the student's writing ability and understanding.

Her belief in direct assessments extended to graded paraphrasing tasks, where she believed students should be assessed based on their actual paraphrasing skills rather than through indirect assessment methods like multiple-choice questions. She expressed concern about the lack of direct assessment of paraphrasing, emphasizing the need for such evaluation and stating that “students need to paraphrase, but all of the questions are multiple-choice questions in writing” (Leena, II).

In her responses to the question about assessment purposes, Leena articulated her preference for specific assessment methods. Thus, it seemed that, in her view, the methods employed in writing assessment are inherently linked to her conceptualization of the purposes of evaluation and the effectiveness of assessment. However, when asked about peer and self-assessment, Leena provided brief responses, indicating a lack of utilization of these methods. Nevertheless, she directed the focus of her discussion toward the significance of feedback and her pedagogical strategies, which will be discussed later.

5.4 Criteria for Evaluating Writing: Assessing Grammar Accuracy

Leena’s beliefs about writing evaluation criteria appeared to place substantial emphasis on grammatical accuracy. Her view of writing seemed primarily as a medium for teaching and reinforcing grammar rules, which she considered just as necessary as content. This stance is evident in Excerpt 5.6, where Leena highlighted the importance of topic and concluding sentences and the relevance and relationship between sentences. Yet, she ultimately steered the focus toward grammar, punctuating her point by stating:

Excerpt 5.6:

The most important thing is the topic sentence, the concluding sentence, how all of the sentences are related, are relevant, and the grammar. I think it’s more than punctuation. I focus more on the grammar... Because the content and the grammar are both important. It’s not only about the content. (Leena, II)

Leena's prioritization of grammar over other aspects of writing was further reinforced in her response to a post-TAP question about the most severe errors in student writing. She identified sentence structure as a critical area where students often falter, indicating her belief that grammatical accuracy is a cornerstone of effective writing. She explained, "Grammar. They [[students]] should focus on sentence structure. I think their worst mistakes are in grammar" (Leena, FI 2).

Leena's beliefs in practice are exemplified in Excerpt 5.7, where she described her ideal writing class. Instead of adhering to multiple-choice assessments, Leena advocated for continuous assessments with a vital grammar component. She expressed a desire to address grammatical elements such as comma rules and article usage, which were often neglected in the English II syllabus but were believed to be crucial in academic writing. She elaborated,

Excerpt 5.7:

I will not ask the student about the organization, about the topic sentence. No. I will ask about the grammar that I want them to focus on, like the comma rules. I'm gonna give them more/ like. For example, one thing that students always face difficulties with is the articles, but we have never tested them or explained the articles or the uses of articles. So, I'm gonna focus actually on grammar and give them an assessment of the grammar that should be used in academic writing... asking them about these rules. (Leena, II)

Leena elaborated on why she preferred to focus on grammar, believing that students "need them because, like, there are a lot of issues in their writing" (Leena, II). Leena seemed to connect good writing with compositions that are free from grammatical errors. She also highlighted the importance of fostering improvement, focusing on areas like the use of transition words and comma rules when providing feedback and explaining that her focus would be "use the transition words perfectly" or "the comma rules," for example (Leena, II).

Synthesizing these Excerpts, it was apparent that Leena perceived writing not as an end but as a means to an end—the end being the mastery of grammar. This view of writing as instrumental in teaching grammar was pivotal to understanding her writing assessment belief

system and her beliefs in practice for evaluating student writing that seemed rooted in her core belief of the skill-based approach.

5.5 Feedback: An Indirect Approach to Feedback

Leena shared her feedback beliefs using explicit categorizations, such as “detailed vs. general,” “verbal vs. written,” and “positive vs. negative.” Her definitions suggested that both detailed and general feedback fall within the scope of indirect feedback. According to Leena, detailed feedback involves pinpointing students’ errors without providing the correct forms, while general feedback gives a broader critique without pinpointing the exact error locations:

Excerpt 5.8:

إيعني ((like)) like to give them detailed feedback, specifying where they make mistakes, but I don’t give them the correction because I know that, like, some kind of feedback to the students/ The teacher, for example, would comment on this sentence and provide an alternative sentence. (Leena, II)

Leena’s beliefs in practice on feedback emphasized the importance of detailed comments to enhance writing skills. As she stated, “I try to give them detailed feedback because I want them to improve their final drafts... they need this kind of feedback” (Leena, II). She posited that the enhancement of students’ writing through feedback relies on avoiding direct corrections, encouraging students to take ownership of their error correction: “...otherwise, it seems like I wrote the paragraph, not the student. I leave the error correction to the students. They need to be able to analyze, to edit, to revise their own work” (Leena, II).

Leena also preferred written feedback, considering it more memorable for students than verbal feedback:

Excerpt 5.9:

I think that *they* [[written feedback comments]] are the most important thing in the feedback. Like, because even if I give the students verbal feedback, they will not remember it, especially because they are freshmen. Maybe they don’t know how to take notes yet, so they need something in front of them to follow. (Leena, II)

Nevertheless, Leena acknowledged that verbal feedback is often more detailed than its written counterpart, but she reported she reserved it for individual sessions in her office. She did

not endorse the practice of giving verbal feedback to the entire class, preferring to provide such feedback when students seek clarification after reviewing her written comments. Leena explained, “I only provide verbal feedback when students come to my office after they see my feedback. I try to explain to them verbally and give them more detailed feedback” (Leena, II).

Thus, there appeared to be a complex interplay within Leena’s feedback belief system. To elaborate, she held the practice belief that she ‘always’ provides her students with detailed feedback. However, while she acknowledged that verbal feedback tends to be more detailed than written feedback, her prevailing belief was that verbal feedback can be forgotten. This led her to limit detailed verbal feedback to instances when students specifically request it. However, external factors like time constraints might also influence this complexity.

Leena also elaborated on her belief in providing students with positive and negative feedback. She categorized any form of positive reinforcement and encouragement as positive feedback, while any evaluative feedback addressing errors as negative. However, Leena acknowledged that she often forgets to give students positive feedback beyond commenting on the paragraph structure. She justified this by emphasizing her belief in highlighting areas for improvement rather than solely focusing on positive aspects.

Excerpt 5.10:

I always forget to give them positive feedback. And even if there is يعني ((like)) any positive feedback, it is only about the topic sentence. So, I would write a sentence, like, “excellent topic sentence” or “excellent concluding sentence,” but, like, there are other things in the paragraph that I forget to mention. Here, I’m talking about positive *things* [[Feedback]]. Like, for example, when they use transition words. Some of them use the transition words perfectly or the comma rules, for example. But I always forget because I focus on giving them the comments that they need to improve. (Leena, II)

However, Leena’s belief in the significance of positive feedback and her acknowledgment of occasionally forgetting to provide it seemed to heighten her awareness of the need to offer positive feedback. During her TAPs, Leena verbalized a couple of times that she wanted to

include positive written feedback, such as “Okay. Let’s write some positive comments” (Leena, TAP 1) or “I am going to write good 1st main point” (Leena, TAP 2).

Additionally, in Leena’s beliefs in practice, error code usage emerged as a pragmatic strategy to streamline the feedback process, aligning with the materials used in the English II course. According to her, she was consistent with the error codes presented in the textbook, using abbreviations like “TS” for topic sentences and “PUN” for punctuation, thereby establishing a standard language of feedback that her students can understand and reference easily. As she explained, this practice helps maintain a clear and direct form of communication regarding specific writing issues. Her rationale was to emphasize simplicity and adherence to established guidelines: “I try my best to follow the same thing that is attached in the textbook... It makes it easier for the students and for me” (Leena, II).

However, her elaboration revealed a nuanced layer to her belief; while she employed standardized error codes, she exercised discretion and opted for elaboration in instances where a code might be insufficient, such as preferring to write out “sentence structure” overusing an abbreviated code (Leena, TAP 1). This selective application suggested that Leena valued clarity and comprehension over rigid conformity to the error code system.

5.6 Grading and Rubric Use: The Significance of Objective and Fair Assessment

Leena’s beliefs in assessment emphasized the essential role of rubrics in enabling an equitable and objective evaluation of student writing. She maintained that grading criteria are crucial in reducing the “subjectivity of grading writing assignments.” She asserted, “I think using grading criteria is so important” because it “makes us [[teachers]] more objective” (Leena, II). She recognized that subjectivity is unavoidable, even with detailed criteria, but argued that rubrics could steer educators towards more even-handed evaluations. She emphasized the

importance of applying the rubric uniformly to all students to ensure fairness by lessening the chances of bias that may come from an individual student’s classroom engagement or personal attributes. Leena elucidated this by stating:

Excerpt 5.11:

With the rubric, you will have fair grading...Because I think, as I mentioned, grading is subjective. You know, for example, this student is so excellent, and she always participates in class. You know her level, but she has, like, some issues in writing. You might ignore these issues, which will not be fair to other students. So, by following grading criteria, you will ensure fairness. (Leena, II)

Leena’s commitment to fair assessment was also reflected in her satisfaction with the rubric used within the PYP context (see Table 3.3 for a copy of the rubric), as she confidently stated, “I’m satisfied with the rubric that we use” (Leena, II), underscoring her belief in the effectiveness of using rubrics for writing assessment.

5.7 Ecological Factors Influencing Leena’s Writing Assessment Beliefs and Practices

Leena’s belief in a skill-based approach to teaching writing was not just theoretical; it was deeply rooted in the practical utility of techniques that have served her well in her academic journey. Additionally, participating in a feedback workshop seemed to have somewhat informed her beliefs regarding feedback approaches and types. Ecological factors, such as curricular constraints and the teacher’s limited authority, appeared to impede her beliefs about using indirect assessment methods to assess students. However, some factors, like the use of a rubric, aligned with her beliefs and facilitated their manifestation in her practice.

5.7.1 The Influence of Leena’s Learning Experience on her Writing Assessment Beliefs and Practices

Leena’s core pedagogical beliefs in teaching writing following a skill-based approach appeared to be influenced by her personal learning strategies. In Excerpt 5.2, Leena highlighted her preference for practical techniques, such as using phrase banks or word lists—a method she personally finds effective while writing. This implied that her teaching philosophy was rooted in

applying concrete tools that have proven helpful in her own academic journey. Consequently, in her actual teaching practices, she encouraged the students to use vocabulary from the reading course to enhance their use of advanced vocabulary in their writing:

Excerpt 5.12:

T: Excellent. But from where do you usually get the advanced vocab?

L: Maybe from books.

T: Excellent from the reading. One way/ So all of the vocab that you have in your reading course, use them in your writing. Okay? This is one way because you already know the meaning, the synonyms, the part of speech of these words. So, use them in your writing. (Leena, OBS 1)

Furthermore, Leena's participation in workshops on providing corrective feedback appeared to inform her feedback beliefs and some of her practices. Excerpt 5.13 details the insights she gained from attending such a workshop at a conference. She particularly recalled the emphasis on avoiding overly 'detailed feedback' that pinpoints errors and the importance of delivering a balanced mix of positive and negative feedback to students. Leena stated,

Excerpt 5.13:

I remember I attended this [[the workshop about corrective feedback]] at one of the conferences. I don't remember a lot. But the one thing that I took with me till now. It was about not giving the students detailed feedback. And not, for example, giving them only negative feedback. You should, like, mention both positive and negative feedback. These are the two things that I still remember from that workshop...I completely agree with. (Leena, II)

Leena's feedback beliefs, as she articulated in Excerpts 5.8 and 5.10, advocated for including positive feedback, especially on topic sentences, and providing detailed feedback to enhance students' first drafts.

The workshop's impact seemed to have informed her actual feedback practices of providing positive comments in the classroom, where she frequently employed affirming phrases like "excellent" and "bravo" to encourage students when they responded correctly. Yet, the analysis of her written feedback across 20 samples revealed that positive comments on topic sentences and conclusions were limited, with just two occurrences noted: "Excellent topic sentence and concluding sentence" and the other, "Good topic sentence and conclusion." This discrepancy between her expressed beliefs and tangible feedback practices might be considered

minor. Leena herself acknowledged this gap in Excerpt 5.10, admitting, “I always forget to give them positive feedback. And even if there is يعني ((like)) any positive feedback, it is only about the topic sentence” (Leena, II, Excerpt 5.10). This self-awareness suggested a reflective practice, yet it could also indicate the weak status of this belief within her assessment belief system.

On the other hand, regarding providing written, detailed feedback by specifying the error location, Leena’s belief showed a degree of complex interaction. She reported providing students with detailed feedback because “they need this kind of feedback,” yet also acknowledged learning from a workshop the importance of not giving detailed feedback. Despite this, Leena did not specify any errors across the 20 samples in practice. She provided only a list of error types found in the student’s paragraph. Refer to Figure 5.1 for her typical feedback comments approach.

Figure 5.1

Example of Leena’s Written Feedback in the First Task

Compare and contrast between eating home and eating in restaurants

We love the restaurant food but we don't now how they made it. Here we have many different types between home made and restaurants. First, is it healthy? Is not healthy if you eat every day in restaurants, you don't now is it clean or no, if they add something you should not eating. Moreover, if someone eat in restaurants maybe will be sick or fat. Second, save your money. These day we have many dreams that we want to achieve, but how you can achieve it if you eat in restaurants and you don't want to eat in home. When you try to save money you should eat in home p. keep your money for something to benefit from. In addition, someone eat every day in restaurants they will be waster. Finally, what is the testy food? The home made is tasty than restaurants. our mother feel tired when they make food every day. They work hard for our health so that we can stop going to the restaurants. Our mothers made tasty food and we love it. All in all, you should be carefully, we don't know the restaurants food. A lot of people tired because the food restaurants.

Take action in student view

Feedback

- sentence structure
- rewrite your main points using different structure
- cap
- comma splice
- rewrite your concluding sentence

Points

No points ✓

Return ▾

This interaction within her feedback belief system may indicate that her belief in specifying the error location resulted from external pressure due to students' levels and needs in this specific context. However, her belief in not specifying error locations, which stemmed from the feedback workshop, appeared to hold a more robust position within her writing assessment belief system. This was reflected more in her actual written feedback practices. This tendency was also evident in her in-class verbal feedback approach. For instance, Leena discussed examples of errors with the students after she corrected their first drafts, but she refrained from indicating the error type or providing corrections. Instead, she repeated the sentence containing the error to draw the student's attention without giving the correction, as shown in Excerpt 5.14.

Excerpt 5.14:

T: You usually make this mistake. You write, for example, "Learning new skills are important." What is the mistake here?

LL: Skills.

T: No. It's not "skills." "Learning new skills are important. New skills are important."

LL: Are.

T: Are. Why it's a mistake?

LL: Singular.

T: Yeah, because "learning" is singular, so you should, yeah, write "is." (Leena, OBS 2)

5.7.2 Tensions Between Leena's Writing Assessment Beliefs and Practices

Beyond the non-linear interplay between Leena's feedback beliefs and her practices, which were influenced by participation in a feedback workshop, it appeared that the imposed English II curriculum, inclusive of standardized assessment methods and textbooks, along with constrained teacher authority, were the principal factors of conflict between Leena's writing pedagogical beliefs and practices.

Leena rigorously discussed several ecological factors that had limited her instruction and assessment of writing. Several of these constraints were rooted in the curriculum. For instance, she shared her concerns regarding integrating reading and writing into a single course. She believed this combination limited the time needed to evaluate students' writing skills through

direct writing tasks exclusively. Leena believed alternative assessment methods, like MCQs, focus more on reading skills and do not effectively measure students' writing abilities or contribute to their writing competence (Leena, II, Excerpt 5.5).

Another curricular factor, specifically the standardized exam and task formats, had hindered the implementation of her belief in focusing solely on direct assessment methods and had profoundly shaped her teaching approach, steering it toward test-focused instruction. Consequently, a significant portion of Leena's classroom time was spent on activities that mimic the questions anticipated in exams, quizzes, or MCQ tasks. During the observed classes, she allocated considerable time to preparing students for exams, which included practice worksheets and targeted drilling exercises, while also giving explicit attention to the format and content of the final and midterm exam questions.

For instance, during her second observed class, Leena stressed the importance of understanding the use of connectors, as there would be indirect evaluations of writing in the midterm and final exams. These evaluations would specifically target linguistic devices' meaning, positioning, and punctuation that create connections and transitions between ideas in writing paragraphs. She explicitly instructed her students about the focus of the exam as follows:

Excerpt 5.15:

T: So, it's important to notice the position and how we use this connector in a sentence. Why? Because you are gonna have a question in the midterm exam and also in the final exam. The question/ Do you remember the fill-in-the-gaps paragraph? (see Figure 3.3 for a copy of this question)

LL: yes

T: So, you are gonna have all of the connectors in *this section*. [[referring to the fill-in-the-gap questions by choosing the best answer]] Okay? You'll be asked about all the transition words, the connectors, the conjunctions, the subordinations, and so on. Okay? (Pause) Why do we use 'therefore' before the last sentence? Why don't we use, for example, 'however'? (Pause)

L: It is a result.

T: Excellent. When you study for the connectors. What do you need to focus on? The meaning, the punctuation, and the positions. It is so important to understand these three points so you can answer any questions related to the transition words. (Leena, OBS 2)

In another case, during the first observed class, Leena handed out a supplementary paraphrasing worksheet that mirrored the MCQs from the students' graded paraphrasing task.

While working through the worksheet together, she pointed out the similarities and differences between this drill and the graded task, offering advice and strategies on how to approach and successfully answer such questions, advising:

Excerpt 5.16:

T: You will have 40 minutes. Use them for the graded task. Don't choose the answer immediately. Read the original text more than one time. Try to understand it. Okay? Next step, you choose an answer. Try to bring up a reason for not choosing this one. For example, if you chose B, think about why you didn't choose A?" (Leena, OBS 1).

Secondly, according to Leena, the textbooks used in the courses impacted her ability to employ materials reflecting authentic and original language usage, or as she put it, "the language that real writers use." She pointed out an instance of this lack of authenticity, highlighting a contradiction in the books used for different courses in the PYP regarding grammatical rules, such as the use of commas with 'while.' She explained, "In *this* [[English II] textbook, it states that there is an exception with 'while,' even if you use it in the middle, you should add a comma. So, it doesn't reflect real-life language." In contrast, the grammar course textbook instructed students not to use a comma when 'while' appears in the middle of a complex sentence.

Consequently, she found it frustrating to have to explain these conflicting rules regarding commas with 'while' to the students and assess them based on the differing requirements of each course. She believed this confused the students and failed to "provide them with real-life language" (Leena, II). Although the 'while' rule may not necessarily be contradictory in the course textbooks, it could pertain to the varied uses of 'while' in complex or compound sentences. Nevertheless, Leena considered the textbook to be "so old." As noted in Excerpt 5.3, she argued in favor of adopting a new textbook that incorporates an "academic word list" at the end of each chapter to better serve the students' learning needs of using academic words in their writings.

Restricted teacher authority emerged as a key factor that Leena repeatedly identified as influencing her beliefs about teaching and assessment of writing and, consequently, her overall practices. She asserted, “It’s all about teachers’ authority” (Leena, II). While Leena did not elaborate on how limited teacher authority affected her practical assessment methods, she did explain why teachers need autonomy within their specific context.

Leena highlighted the critical role of teacher authority in effective teaching, particularly in writing. She argued that a teacher’s ability to identify and address specific student needs was hindered by restricted authority, impacting the teaching materials' quality and relevance. She explained,

Excerpt 5.17:

Teachers’ authority is so important, like in any course, but even in writing, specifically. Because, for example, you notice that this group of students struggles with specific issues. So, you need as a teacher to focus on those issues. You don’t need to give them, like, other things that do not help them. So, like, based on these issues, you can design your own teaching materials and give these materials to the students. But we don’t have this kind of authority. (Leena, II)

According to Leena, this limitation was not only a barrier to effective teaching but also to fair and appropriate assessment practices. Leena argued that a uniform approach to assessment cannot accommodate the students’ varying proficiency levels. Leena specifically mentioned the unfairness in treating all nursing students as if they were on the same level of English proficiency as the unified students. She suggested that equitable education necessitates tailored assessments and resources and thus advocated for differentiated assessments to accommodate these differences:

Excerpt 5.18:

The point is that students’ levels are different. So, they need different assessments. Some students need more teaching materials, more activities, need to practice more, others don’t. So, for example, it’s always unfair to deal with nursing students like unified students because they have two different English levels. (Leena, II)

Leena’s criticism extended to the restrictions on creating resources like PowerPoint slides, explaining, “We are not even allowed to give *them* [[the students]] PowerPoint slides other than

the ones that are shared by the coordinator” (Leena, II). This restriction symbolized a broader issue of limited teacher autonomy, which she believed negatively impacted the organization and delivery of course content. Leena argued that such restrictions prevent teachers from optimizing the learning experience to suit their students’ specific requirements.

5.7.3 Harmony between Leena’s Writing Assessment Beliefs and Practices

Most alignments between Leena’s beliefs and her practices indicated a positive interaction between her belief in practice and actual practice. As might be anticipated, the distinction between beliefs in practice and actual practices can often be subtle. Furthermore, there was evidence of a consistent alignment between her professed beliefs and her actual practices.

Leena believed in the skill-based approach and viewed writing literacy as a technical discipline where discrete skills and techniques can be methodically learned and applied. This core belief was reflected in her teaching and assessment practices in the writing classroom at multiple levels. In her classes, Leena stressed the importance of mastering strategies, modeling, and consistent writing practice. Initially, as indicated, she encouraged students to utilize vocabulary lists from their reading course (Leena, OBS 1, Excerpt 5.12). Moreover, she underscored the need to understand how to emulate written models by providing additional essay and paragraph examples for student guidance. She instructed:

Excerpt 5.19:

T: When you analyze an essay or when you study إن شاء الله ((if God wills it)) for the midterm exam or the final exam, please read all of the models in the chapters. It’s not only about reading the essay or reading the paragraph. It’s about noticing how the words, how the transition words, how the sentence structures are written or used in the paragraph. In that way, you will learn, and you will avoid making mistakes in your writing. OK? (Leena, OBS 2)

In line with this, Leena’s classroom assessment approach involved closely monitoring and supporting the students’ comprehension of such skills before they began drafting their paragraph outlines. She actively assisted students in deconstructing written models to deepen their

understanding of paragraph elements such as topic sentences, supporting details, and concluding statements. As shown in Excerpt 5.19, Leena projected a textbook model onto the blackboard during class and dedicated time for students to dissect and question its structure:

Excerpt 5.20:

T: What I want you to do is to identify the thesis statement, the three main points, the concluding sentence. When you read the conclusion, I want you to tell me the strategies that the writer uses.... While reading, please, I want you also to tell me, like, where is the hook, or where is the general background of the introductory paragraph? What is the thesis statement? Okay? So, try your best to analyze the essay, not only read it. (Leena, OBS 2)

Leena then guided the students' understanding by posing numerous questions. She would ask a question, pause, and patiently await the students' responses, refraining from providing the answers herself.

Lastly, Leena's belief in the critical role of writing practice was reflected in her classroom activities. She facilitated a practice writing exercise before assigning the first graded writing task, encouraging students to compose a well-structured compare-and-contrast paragraph. In the first follow-up interview, she also emphasized the significance of providing supplementary practice materials and additional writing models to guide students. She stated,

Excerpt 5.21:

I feel that the students need to look at other examples. Because they couldn't figure *it* out [[the structure of compare/contrast paragraph]]. So, I brought other examples, other worksheets with other explanations, other ways of explanation. (Leena, FI 2)

This deliberate focus on repeated practice reinforced Leena's belief that regular and systematic writing exercises were vital for students to enhance their writing proficiency.

Moreover, further evidence of alignment was found between Leena's beliefs regarding evaluation criteria and her actual assessment practices, especially in her prioritization of grammatical accuracy in writing. Leena consistently underscored the significance of producing writing devoid of grammatical mistakes, and she identified grammar as the area where students most frequently erred (Leena, II, Excerpts 5.6 & 5.7). This emphasis on precision shaped her

feedback and grading, with her 20 samples indicating that most comments addressed form-related issues. As depicted in Figure 5.1, Leena’s checklist included sentence structure, comma splices, capitalization, and articulating the main point through varied structures. An additional example of this attention to grammar is presented in Figure 5.2, where Leena noted multiple grammatical topics, including comma rules, sentence fragments, verb tenses, prepositions, and comma splices, areas that the student struggled with in her paragraph construction. Leena also implemented error codes in her feedback, a method she reported using in her feedback (Leena, II).

Figure 5.2

Example of Leena’s Written Feedback that Focuses on Grammar

The screenshot displays a writing assignment titled "Eating at Home and Eating in a Restaurant". The main text area contains the following paragraph:

Some people like eating at home more than eating in a restaurant, but some people like eating in a restaurant more. And there are many differences between eating at home and eating in a restaurant. First, eating at home is healthier than eating in a restaurant. Home foods include many important nutrients such as vitamins, minerals, and proteins. It keep the body healthy, and it has less fat that makes the body fit and slim. Home food has all these benefits because they are doing it by themselves (handmade). Second, eating at home will save money because many people put their money into the right and important place like school fees, so because of some reasons people do not like to waste their money in a restaurant. Third, everyone will not waste the food that mothers did. For example, when they found food in the home they will eat it and that leads to saving food and not wasting it. Next, on the other hand, eating in a restaurant for others is more delicious than eating at home, because it has many flavors. In addition, eating in a

The right sidebar shows a feedback panel with the following content:

- Take action in student view
- Feedback
 - never start the sentence with conjunctions
 - comma rules
 - vt
 - prepositions
 - comma rules
 - comma splice
- Points
 - No points ✓
- Return

Additionally, during in-class activities and while discussing the essay model, Leena paused to pose questions on English sentence structure, such as, “What is the subject of this sentence? What is the verb?” She emphasized the frequency of errors in these elements, stating, “Because this is a kind of sentence that you usually use in your paragraphs, but you usually make mistakes

in it” (Leena, OBS 2). While Leena’s emphasis on grammar might stem from her belief that good writing equates to a polished paragraph, the interactions appeared much more intricate. This focus might also arise from the students’ limited English proficiency and their inadequate grasp of grammatical principles. In her interview, particularly in Excerpts 5.17 and 5.18, Leena expressed the necessity of teacher authority to concentrate on the predominant issues students faced. Additionally, she conveyed her concerns regarding the absence of English proficiency testing before the University admission, underscoring its negative effect on the students’ English learning journey, as she shared:

Excerpt 5.22:

You know, *here* [[in the University]], students are accepted without proof of their level in English. That is not helping them. *Nursing* [[students]] struggle now with basic English. I don’t know what they will do in the third or fourth year. (Leena, II)

To elaborate, Leena’s attention to grammatical accuracy aligned with her classroom instruction and feedback methods. However, this could also be shaped by external influences such as the students’ proficiency and the curriculum emphasis, which in turn molded her beliefs within the PYP context and, consequently, her actual practices.

While Leena’s feedback primarily concentrated on grammar and mechanics as per the rubric, she also gave minimal attention to additional criteria like organization and lexical resources. She directed students to ensure the inclusion of a topic sentence, three main points, and a concluding sentence with prompts like, “Where is your topic sentence?” and “Where is your 3rd main point?” She also concentrated on the use of formal language, advising against the use of first and second-person pronouns with guidance such as, “Do not use ‘I’ and ‘you.’”

Furthermore, Leena advocated for the use of a rubric as a tool for equitable assessment, a belief reflected in her classroom practices (Leena, II Excerpt 5.11). She demonstrated the value she placed on the rubric by instructing her students to employ it as a ‘checklist’ for self-

evaluation, emphasizing its importance not just for the teacher but for the students as well. She directed them:

Excerpt 5.23:

T: The rubric is so important for you. It's not for me only but for you. Why? Because you are gonna use the rubric as a checklist. After you finish writing, you are gonna go back to the rubric and you make sure that you follow the criteria. You do what it's asked/ what it's required from you. Okay? (Leena, OBS 1)

5.8 Summary of Leena's Case Findings

Figure 5.3 summarizes the key findings of Leena's case, illustrating the elements of and interplay among Leena's writing assessment belief system and actual practices, alongside the influencing ecological systems on their interrelationships.

The left side of Figure 5.3 illustrates Leena's belief system, encompassing a range of assessment beliefs and beliefs in practice. Although the connection between her core and peripheral beliefs was not always apparent, Leena predominantly believed in a skill-based approach in her teaching of writing, valuing technical mastery as a crucial component of writing as an instrumental skill. As indicated by the lower bidirectional purple arrow, this core belief was significantly influenced by her personal learning strategies (microsystem), suggesting her belief in the effectiveness of practical techniques such as phrase banks and model analysis. This was also in harmony with her actual practices, where she emphasized these strategies to her students and actively monitored their understanding and application. This core belief extended to one of her peripheral beliefs—that engaging in writing practices is critical to writing development.

One of Leena's peripheral beliefs was geared towards feedback strategies, like giving positive reinforcement and encouraging self-editing skills by offering indirect written feedback—identifying error locations but not providing the corrections. This belief seemed to be shaped by her experiences in professional development as a teacher (microsystem), including attending a workshop focused on feedback. Despite her emphasis on positive reinforcement,

particularly for topic sentences, her written feedback seldom reflected this. Although Leena recognized the importance of specifying error locations, she tended to provide only general categories of errors without pinpointing them, revealing a tension between her intended feedback beliefs in practice and her actual methods. This suggested that these feedback beliefs were not as deeply ingrained within her assessment belief system, as there appeared to be no influential ecological factor causing this belief-practice misalignment.

Examining the exosystem (orange box) revealed curricular and institutional constraints. The integration of reading and writing in one course, the use of standardized exam formats, and restricted teacher autonomy significantly steered Leena's practices toward test-oriented instruction. These external factors necessitated a shift from her belief in the importance of direct assessment methods, which emphasize extensive writing practice, to a focus on equipping students with test-taking strategies for answering MCQs in exams.

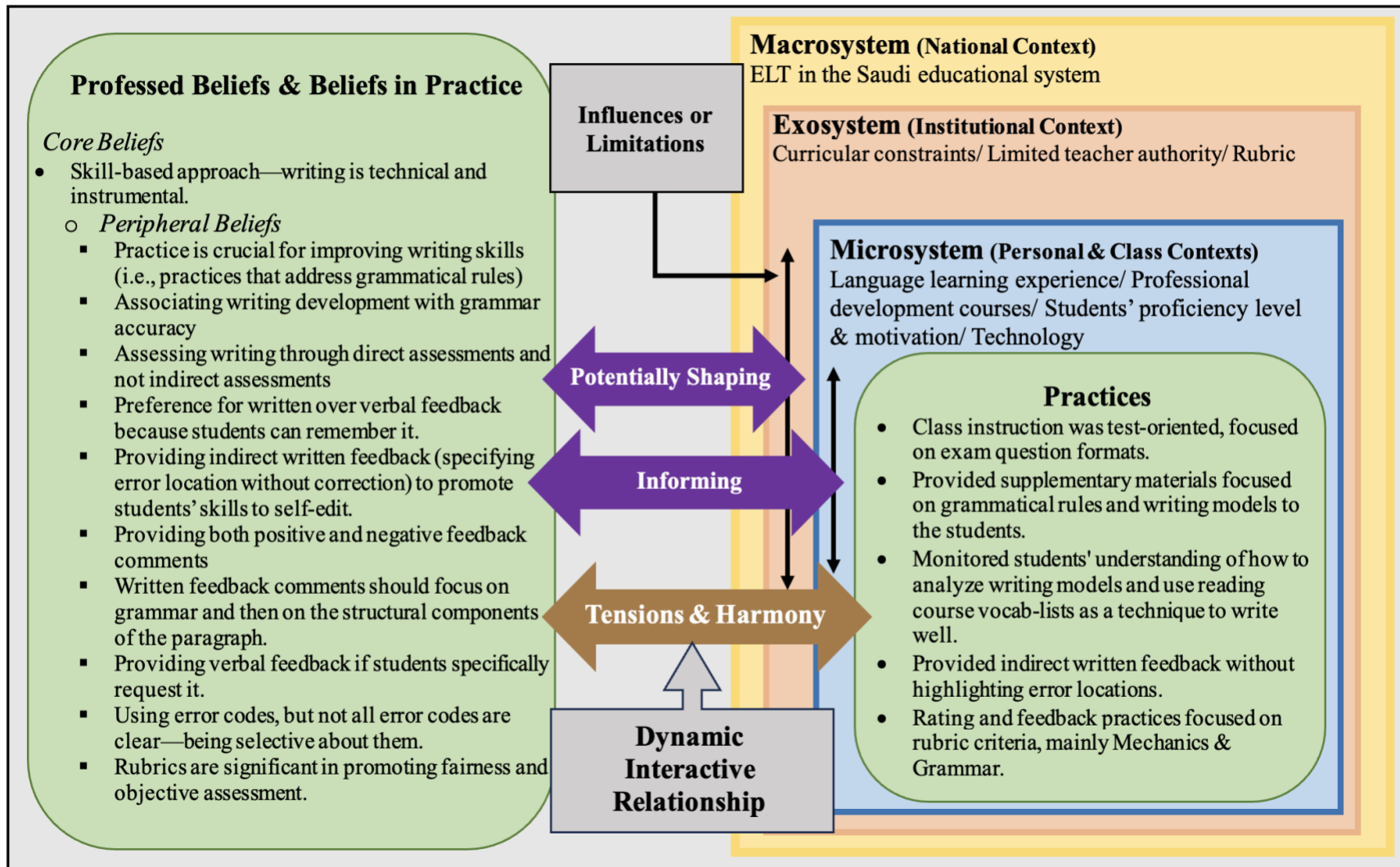
Furthermore, Leena associated good writing with texts devoid of grammatical errors. Her beliefs concerning evaluation criteria prioritized grammatical accuracy, focusing on elements like comma usage and sentence structure. This belief appeared to have a dual origin: it might be a fundamental part of her belief system or possibly a response to students' low grammar proficiency. As suggested by the upper bidirectional purple arrow, this belief could be contextually shaped by Leena's extensive experience in the PYP, where she has observed students' limited understanding of basic grammar. In practice, Leena's actual feedback and assessment methods were consistent with this belief, with her written feedback and classroom instruction primarily targeting grammatical correctness.

A further alignment was observed between Leena's assessment beliefs and her practices regarding the use of rubrics. She believed in the rubric efficacy for fair evaluation, and its

presence in the PYP context facilitated the translation of this belief into her practice. Leena emphasized the importance of the rubric by guiding her students to use it as a 'checklist' for self-assessment. The factors within the macrosystem appeared to exert no discernible influence on Leena's writing assessment beliefs or her practice.

Figure 5.3

Summary of Leena's Case Findings



CHAPTER 6: THE CASE OF AYA:

Assessing Individual Progress

6.1 Background for Aya

Aya, an English lecturer between 35 and 40, has been employed at the University for three years, where she has been teaching foundational English skills to undergraduate medical students in their PYP. She holds both a bachelor's and a master's degree in linguistics, acquired from Canadian institutions. Her professional journey, as an EFL teacher, commenced at this University, her sole teaching experience. During her higher education in Canada, Aya did not pursue many jobs; instead, she concentrated on her dual roles as “a mother and student” (Aya, II). After the completion of her academic pursuits, Aya returned to Saudi Arabia and secured her current lecturer appointment at the University.

As she narrated, Aya's English learning journey underscored a gradual and continuous acquisition of the English language. Her initial engagement with English began in a Saudi middle school, which laid the foundational groundwork for her English language development. The turning point in her language proficiency, as she denoted, was during her master's studies: “I think my master's was the turning point in my English” (Aya, II). This phase presumably provided an immersive environment that catalyzed her English language skills beyond the basic levels attained during earlier schooling.

She noted that the transformative experience in her English writing competence was attributed to the academic requirements encountered during her undergraduate studies in Canada, where the necessity to articulate complex thoughts in research papers and essays facilitated a deeper understanding and application of the language: “I learned how to write in English...when

I was doing my bachelor's because I had to do some research papers and I had to write some essays" (Aya, II).

During data collection, Aya was responsible for teaching a writing class comprising 23 students from the unified sections. She engaged with her students for three hours per week, focusing on the writing component of the English II course.

6.2 Writing Assessment Purposes: Assessment of Learning

Aya's core belief about the purpose of writing assessments appeared to be grounded in the 'assessment of learning' (AoL) paradigm. Her explanations during the interviews revealed a strong inclination towards employing assessment as an instrument for evaluating the individual writing progression of students and the extent to which they have assimilated the instructed material of English II.

When Aya was asked about the purpose of the assessment, she stated, "to assess their [[students']] progress in class" (Aya, II). Although she did not explicitly mention the term *assessment of learning*, her viewpoint suggested that she perceived assessment as a means to track and measure students' individual growth over time within the context of the writing course rather than as an evaluation of achievement. She explained,

Excerpt 6.1:

The purpose? I think to assess their progress in class. Yeah, that would be the main purpose. Not assessing the task, the specific task. I would check the progress. For example, if a student just starts writing a sentence يعني ((like)), she's different from someone who already knows how to write/ how to write a paragraph. (Pause) I would, yes, consider, as I said, the progress of each student. I wanna compare, for example, the beginning of the semester, how this student did, and how she did at the end of the semester. (Aya, II)

Aya seemed to define students' progress in terms of how much they have learned relative to their starting point within the context of what has been taught. She appeared to advocate for an individual evaluation for each student rather than comparing them to one another.

A further illustration of Aya's core belief in considering individual progress was evident during the second TAP session. While reviewing a student's paper, Aya observed that the student needed to sufficiently elaborate on the first main point of her definition paragraph. She thought, "It is not a big mistake. Yeah, especially for this student" (Aya, TAP 2), indicating that she considered the student's current proficiency level when evaluating their papers.

However, during the TAP sessions, Aya also seemed to consistently reflect on and give precedence to the content taught in class when rating students' papers rather than solely considering students' levels. She questioned, "Okay, what kind of transition she can use? Okay, what transitions did they learn?" (Aya, TAP 1) and "I have taught them that. She should know that" (Aya, TAP 2). In the follow-up interview, she clarified, "If I look at the paper, I consider what I'm teaching, I don't only consider their level, like I have to consider what I'm teaching, what I told them" (Aya, FI 2). Aya's assessment beliefs and beliefs in practice appeared to aim to determine what students had assimilated from the instruction provided.

In discussing the impact of assessment on students' learning, Aya did not broadly address the effect of writing assessment on learning as a whole. Instead, she focused on the specific implications within the PYP context, noting the pervasive testing culture there. She expressed concern over the emphasis on grades over actual learning. She observed that assessments influence students' learning "in that they become, or they care only about the grades. So sometimes, yes, they care about their grades only. And they forget to learn writing for writing" (Aya, II).

6.3 Teaching and Learning Writing: Emphasizing Idea Generation and Content Quality

When Aya reflected on her learning experience of English writing, she delved into the challenges she faced. She underscored the importance of idea generation and the quality of

content construction, saying, “I always believe the hardest part of writing is how to get the idea articulated” (Aya, II). While Aya acknowledged this as a challenge, her further explanation seemed to underscore her belief in the significance of crafting quality ideas, which she views as crucial to learning how to write. She elaborated,

Excerpt 6.2:

So, the first part is the idea, for example, the topic sentence that you need to start with. This is the hardest part. It's just the idea. It's not about writing. It's just how to get the idea clearly written in the introduction. (Aya, II)

Additionally, her focus on content and ideas extends to her beliefs about the selection of writing topics for the paragraph tasks. When discussing these writing tasks in the PYP context, Aya expressed a desire to make certain changes, particularly regarding topic selection. Her response appeared to reflect her beliefs in promoting student autonomy and teaching writing for specific purposes. She emphasized the importance of allowing students to choose their own topics, stating that students “would feel more comfortable and more confident writing about what they like or what they are interested in” (Aya, II).

Aya also highlighted the relevance of incorporating health-related subjects, considering the students' fields of study in the medical and nursing sectors, moving away from “repetitive” topics such as “talk about your holiday, about the popular places in your city” (Aya, II), and creating opportunities for personal interest exploration. Her suggestions seemed to demonstrate her belief in enhancing students' engagement and aligning writing tasks with students' fields of study or personal preferences for a more meaningful and effective writing learning experience, stating that “Maybe if we give them, as I said, the freedom to talk about what they like, or we make it more about health-related topics, they will learn better” (Aya, II).

Further evidence from the data concerning Aya's belief that idea generation and argument development are critical to writing will be addressed in subsequent discussions. This belief was

also highlighted in a dedicated brief section, as it was found to be a core belief, and additional data suggested a correlation between her emphasis on content quality and her actual practices in writing evaluation.

6.3.1 Criteria for Evaluating Writing: Prioritizing Clarity of Ideas

Aya's emphasis on content in teaching and learning writing appeared to align with her evaluative criteria beliefs. She stated, "I would focus on whether the student clearly expressed the idea they intended to convey" (Aya, II). This indicated that Aya's criteria for assessing writing were intertwined with her belief in the crucial role of content mastery in L2 writing development.

In her response to a question about the focal points when evaluating students' compositions, Aya discussed both her beliefs in practice within the PYP context and her personal stance. She acknowledged the PYP requirement to "grade them on their spelling" yet expressed her disagreement with this practice. Preferring a more lenient approach to spelling, she indicated her intent to be "generous," considering the prevalence of digital spellcheck tools. She elucidated,

Excerpt 6.3:

So as long as the student gives me/ I would focus يعني ((like)), did the student clearly express the idea about the thing that she wanted to talk about? I would be generous with spelling because nowadays, we have spelling checks. So, because the students have *this* in their in-class writing practices [[the focus on spelling in the PYP context]], so we have to grade them on their spelling. I don't think they should be graded on spelling. I don't think spelling is something important. As I said, in the future, when they wanna go outside, they will use their computers everywhere. So, they don't really need to be perfect spellers. (Aya, II)

Aya criticized the PYP's focus on spelling in assessments, contending that impeccable spelling became less critical with technology. Thus, her approach to writing assessment appeared to favor the substance and clarity of the students' ideas over strict spelling accuracy.

In a TAP instance, Aya evaluated a student's entire definition paragraph after correction, focusing on content aspects, stating,

Excerpt 6.4:

P: She didn't explain the main point separately, and she jumped right away.

P: The details/ well/ Conclusion is/ it could have been better.

P: It's okay. They're still learning it. It's not/ it's not bad. It's not bad paragraph. (Aya, TAP 2)

Excerpt 6.4 demonstrates that Aya's thoughts and beliefs during the TAPs revolved around evaluating the content, including the main point, details, and conclusion. She explicitly articulated, "It's not bad content." The discussion of the assessment criteria was also highlighted in Aya's second follow-up interview, particularly after this TAP session, where she was asked about her priorities in rating students' papers. She emphasized content-related aspects like the topic sentence, conclusion, and main points.

Excerpt 6.5:

I focused on the topic sentence. Did the students write a good topic sentence, a good conclusion, have three main points? Spelling errors. Not taking the mark from the first time. Let's say the student had two spelling mistakes. Maybe. Yeah, I need to forgive her for those mistakes. (Aya, FI 2).

Consequently, Aya expressed a lenient stance on spelling, choosing not to deduct marks for minor errors, which seemed to reflect her focus on content over form. However, her reflection did not extend to other writing aspects like grammar or vocabulary.

6.4 Writing Assessment Methods

Aya maintained that exams were unnecessary, arguing that students should be assessed on their writing abilities rather than their knowledge of grammatical rules. She emphatically stated, "I would طبعاً ((for sure)) cancel the exams. They're pointless" (Aya, II). Upon inquiry about her stance on exams, she elaborated, "Why? Because we're teaching them writing. Why would they need to know the name of the rules? They should be able to apply those rules in their writing. This is the purpose of writing" (Aya, II). Emphasizing a preference for practical application, she added, "I would prefer to teach them two tasks about writing very well, better than giving them four tasks that some of them are multiple-choice/ about grammar" (Aya, II).

When Aya was asked to discuss her assessment methods beliefs, she touched upon self-assessment and peer assessment and expanded on her beliefs regarding writing assessment processes and instruments, detailing her perspective on feedback, grading, and using rubrics in practical application.

Aya expressed that she sees peer assessment as advantageous for her students. She conveyed that leveraging the varying skill levels in PYP classes can be beneficial, noting, “It would definitely help, especially if you put different levels of students with each other. They can help each other for sure” (Aya, II). She also mentioned the value of peer assessment, as “students are always interested in helping each other” (Aya, II). Despite this, she noted that she seldom incorporates peer assessment in her writing instruction, attributing this to time limitations—an ecological factor that will be explored later.

Additionally, Aya conveyed her belief that students naturally partake in self-assessment, noting “they’re doing that anyway” (Aya, II). She reasoned that activities such as “proofreading” and “checking” their work before submission are forms of self-assessment. She suggested that students particularly undertake this process after receiving feedback on their first drafts of English II tasks and advised that “they should consider all this feedback when they start writing or before they submit their final draft” (Aya, II). Aya seemed to endorse the belief that students should be accountable for their writing and leverage feedback to improve their skills independently, without the need for her to direct or highlight self-assessment explicitly.

6.5 Feedback: Preference of In-class Individual Verbal Feedback

Aya emphasized her belief that the most beneficial form of feedback is verbal, especially when she observes students actively engaged in the writing process. She stated that “the major feedback, or the feedback that really helps the students, is when I see them writing in class”

(Aya, II). Her response suggested a preference for providing targeted and immediate verbal feedback in the writing class.

Aya further described her beliefs in the practice of providing verbal feedback, indicating that she would analyze the students' sentences and offer specific guidance, such as highlighting when "the topic is unclear" and "the structure of the sentence is not good" (Aya, II). She appeared to value real-time observations in providing targeted comments to improve the content's clarity and quality.

Aya articulated the reasons for preferring verbal feedback, highlighting its interactive nature, the immediate assistance it provides, and the option to use the student's native language when necessary. She explained,

Excerpt 6.6:

I feel the verbal really helps. Why? Because I can tell them, "This is not very clear," "You can do it like that," "You can add something," "The structure of the sentence is not good," and she can correct it at the exact moment. And she would ask me again, "Miss, did I do it well this time?" Maybe I can correct it. Maybe I can give her another example. So, the verbal feedback/ I feel, really helps. (Pause). I also have a feedback session, so I would go around, and I would talk to every single student. Or if the students have questions they would ask, raise their hands. Yeah, I feel the verbal really helps. Because it is about talking to them, explaining to them, sometimes the student can explain what she wanna say in the sentence so I can help her. I even can sometimes use Arabic, especially with beginner students, even with native-like students. That's why maybe I prefer verbal feedback because usually, I can use the L1. But with the written, I cannot. (Aya, II)

To clarify, Aya emphasized the interactive and dynamic nature of verbal feedback in the classroom, which she believed fosters immediate learning opportunities. She could promptly address unclear areas, allowing students to make necessary adjustments right away. The importance of dialogue in verbal feedback allows her to scaffold students' compositions by providing tailored support and addressing individual questions and errors. Moreover, Aya appeared to find that verbal interactions facilitate better understanding and allow her to adapt her language to meet the student's needs, such as using their first language, an approach not possible with written feedback.

Furthermore, Aya's core belief was assessing students' individual progress, which seemed to inform her preference for personalized verbal feedback in class. This is evident in Excerpt 6.6, where she stated, "I would talk to every single student... the student can explain what she wanna say in the sentence so I can help her". In a different part of her initial interview, while discussing her beliefs in practice about writing paragraph tasks, she mentioned her method of providing in-class verbal feedback: "I go around, and I check on every student. And every student is different from the others" (Aya, II). These remarks underscored Aya's commitment to individually assessing students and offering customized guidance to meet their learning needs through in-class verbal feedback.

However, when asked directly about her beliefs on written feedback, Aya reported that she was obliged to provide students with written comments, particularly for the first drafts of their paragraph tasks, saying, "I mean, for sure, the first draft, I have to give them written feedback. Sometimes the course coordinator says we have to" (Aya, II). Though she acknowledged the requirement to give written feedback, Aya held that verbal feedback is "enough" and had adapted by allocating "a session about *feedback* [[verbal]] and the whole session about feedback" (Aya, II). It seemed Aya's firm belief in the effectiveness and adequacy of verbal feedback had led her to dedicate class sessions to it, regardless of the coordinator's influence or assessment regulations. Additionally, Aya reflected once more on the time constraints factor, which she believed limited her ability to provide comprehensible written feedback. This issue of time will be addressed later.

Furthermore, when discussing the focus of her written feedback, Aya emphasized that her "feedback focuses on clarity, always" (Aya, II), which corresponded with her core belief in the importance of idea generation and content in producing quality writing. However, as a belief in

practice, Aya also reported providing direct feedback on spelling, punctuation, and formatting, explaining, “because I feel, you know, they’re very minor mistakes, and the students can avoid them” (Aya, II). This discrepancy between her stated beliefs and beliefs in practice might stem from assessment policies, including the rubric, which mandated her to assess linguistic forms and not just content and rhetoric.

When Aya shared examples of her verbal feedback comments in Excerpt 6.6, it appeared that she believed in employing a combination of indirect and direct feedback methods. However, it seemed that she leaned more toward a direct approach. For example, she mentioned phrases she would use, including “This is not very clear,” “You can do it like this,” “You can add something,” and “The structure of the sentence is not good.” Remarks such as “this is not clear” and “not good” demonstrated her method of pointing out errors without explicitly correcting them, but she provided clear guidance. In contrast, phrases like “you can do it like this” and “Maybe I can correct it. Maybe I can give her another example” implied that she also believed in offering direct corrections at times.

Additionally, when I asked about her direct written feedback beliefs, mainly her reasons for correcting specific errors while grading students’ papers in TAPs, Aya justified her method by saying, “I want them to learn. I correct mistakes that are easy to fix, like spelling, tense. I don’t think they need to lose grades because of this” (Aya, FI 1). This suggested that Aya believed in following a direct written feedback approach when correcting mechanical and grammatical issues over content-related errors such as ideas and main points.

6.6 Grading and Rubric Use: The Importance of Fairness

Aya reported that she favored an alphabetical grading system: “I can give the grade this way, like with A+, A, B+, B” (Aya, II) rather than numerical scores as used in the PYP context.

Aya also believed in her ability to assign these grades without the aid of a rubric, asserting, “I believe I don’t need the rubric if I can see the paper is an A+.” Aya seemed to lean towards an impressionistic assessment, drawing on her expertise to judge students’ writings.

Crucially, Aya’s grading beliefs aligned with her core belief of prioritizing content-related criteria and evaluating individual student progress. Firstly, she explained that she believed in centering her grades mainly on whether the student’s “paragraph is well-structured. The idea is clearly stated... has some arguments... the conclusion is good and well structured” (Aya, II). Her response reemphasized the strong position of her belief that good writing hinges on clear content and the writer’s ability to convey ideas compellingly and cohesively.

Additionally, her belief about the best way to grade students was informed by her belief in the importance of considering each student as an individual learner, emphasizing their progress and the knowledge they have gained in class. Aya articulated this grading belief as follows:

Excerpt 6.7:

What would really help with the grading is if we consider every student individually. يعني ((like)) sometimes the rubric is not/ is unfair. Cause some students have more English knowledge than others. So, I have to consider, for example, their starting point and what I taught them in class. And then if they did what I said in the class, not based on what they already knew. (Aya, II)

Excerpt 6.7 underscores Aya’s beliefs in a tailored approach to grading. She contended that a standardized rubric could be “unfair,” acknowledging the diverse levels of English background knowledge among students. Her grading beliefs aimed to reflect individual learning trajectories within the English II framework, preferring a progress evaluation over a rigid, one-size-fits-all rubric.

However, Aya’s elaboration on her grading beliefs brought to light a teacher’s dilemma: how to balance assessing students using the PYP rubric—a measure of student achievement—and her own belief in evaluating student progress. She articulated,

Excerpt 6.8:

I always wanna assess the student based on where she started. So, the progress. Yeah, the progress again... (Pause). And I ask myself the question many times. If this student, I know her level, and I know how she started and what she's doing now is perfect. So, she deserves an A for her progress. She started from a very low level, and she wrote this at this point. So, I would give her an A, and I wanna give her an A for her progress. But if you take her paper and compare it to already a good student in the class, she doesn't deserve an A. (Aya, II)

As indicated, 'progress' for her signified the extent of a student's improvement from their initial position, not how much they have learned relative to what the PYP sets as achievement expectations. Aya also appeared to perceive that applying specific criteria implies norm-referenced standards, which necessitate comparing students against each other, whereas, in reality, she was only required to assess the students according to the rubric criteria (see Table 3.3 for a copy of the rubric). Thus, the rubric posed a challenge to her.

Nevertheless, Aya acknowledged that rubric and numeric grades were important in the PYP context because they promote consistency among teachers who teach writing. She explained the reasons behind this belief as follows:

Excerpt 6.9:

Not using *it* [[the rubric]] would affect the consistency, of course. Because I'm not teaching the whole batch, right? We have different classes... So, for the sake of consistency, I think, yeah, maybe we should stick to the rubric, especially in unified classes because those students are competing for seats in medical school. So, yeah, we try to do our best to be fair and consistent. (Aya, II)

Excerpt 6.9 indicates that Aya believed that an evaluation process that promotes fairness and consistency was suitable for the PYP context because students were competing for seats in medical school. Furthermore, during her first follow-up interview, when asked about the practice of counting the student's errors before assigning the final grade in the TAP session, she explained, "I try to be fair. Let's say the students don't have a competition, and we're teaching English just to teach them English, not for the grades. Maybe I don't need to do that" (Aya, FI 1).

Aya's reflections in this section suggested a complex interaction between various grading beliefs that seemed contradictory but contextual. She preferred assessing progress over

achievements, envisioning a personalized evaluation over relying on the rubric. Yet, she acknowledged the rubric's role in ensuring consistency and fairness within the PYP. Aya recognized that not teaching all the students' sections could undermine this consistency. She thus saw the value in a standardized rubric for equitable, impartial student assessments, ensuring a fair playing field for all.

6.7 Ecological Factors Influencing Aya's Writing Assessment Beliefs and Practices

Aya's experience as an L2 English learner shaped her approach and beliefs about student challenges, fostering a sense of generosity in her grading. Ecological factors, such as assessment policies (e.g., rubrics), curricular constraints, and time limitations, appeared to impede the full realization of her evaluation and grading beliefs into actual practices. However, there was a notable harmony between her beliefs in feedback practices and her actual feedback practices.

6.7.1 The Influence of Aya's Learning Experience on her Writing Assessment Beliefs and Practices

Aya's perception of student performance seemed to be influenced by her own experiences as an L2 English learner, shaping her belief that English learning is continuous. Consequently, she appeared to believe that writing should be evaluated on individual progress rather than achievement.

Reflecting on her English learning journey, Aya stated, "I'm still learning. I'm a second language learner. So, I speak English as a second language learner clearly" (Aya, II). Her identification with being "a second language learner" underscored her view of learning as an enduring process, one that involves constant engagement with the language rather than a fixed level of knowledge. Her TAP thoughts also reflected her view of the students as developing learners, stating, "It's okay. They're still learning" (Aya, TAP 2).

Aya elaborated on this belief when discussing how she would assess her students or what she would prioritize if she were to create her own writing class. Her response in Excerpt 6.10 below indicates that she recognized the importance of considering her students' second-language learner status. She emphasized the need to avoid overly stringent assessments that expect native-like proficiency from non-native speakers, stating that in the PYP, the assessment was "strict" and "harsh."

Excerpt 6.10:

I would consider that the students are second language learners. Okay. So sometimes we don't really consider that. They're not asking the students to be native, but we're very strict and harsh on them. We have to consider that those students, for example, are exposed to those words or to those questions for the first time. So, the ideal assessment will consider that this is a class for second-language learners, not for native speakers. (Aya, II)

In Excerpt 6.10, Aya highlighted the significance of acknowledging that the students were encountering certain words or questions for the first time and stressed that her ideal assessment would consider that this is a writing class for second-language learners. Aya seemed to perceive that the PYP assessment of achievement following numerical numbers and focusing on an analytical rubric was an overly stringent assessment.

When specifically asked about the impact of her learning experience on her writing assessment beliefs and practices, Aya clarified that her background as an L2 learner has given her insight into the difficulties students encounter while learning and improving their language skills. Consequently, she seemed to adopt a more lenient and understanding assessment approach with her students. She elucidated,

Excerpt 6.11:

I consider myself a second-language learner. So, let me say, I understand the struggle of learning English. I've been through this. And I think this made me more generous and considerate of my students' mistakes. Yeah, I consider myself generous, but/ big 'but' here {laughs}. I'm very generous, but I'm fair with my students. (Aya, II)

Nevertheless, fairness in assessment also seemed to be a significant belief in Aya's assessment belief system, which, as highlighted in Excerpt 6.9, might stem from the competitive environment among students within the PYP context.

6.7.2 Tensions Between Aya's Writing Assessment Beliefs and Practices

As indicated, the PYP assessment policies led to a conflict between Aya's preference for subjectively evaluating students' individual progress and the mandated use of an analytical rubric for measuring achievement. The PYP's numerical grading system also limited her ability to use broader categories like A or B-plus, which she felt more accurately represented each student's personal growth. Aya was conscious of these tensions and the impact of ecological factors on her practices.

Aya strongly believed in assessing students based on continuous progress and individual development, valuing the idea of "learning writing for writing" (Aya, II), rather than for grades. Yet, she recognized that her actual practices had to focus on grades to meet the demands of the students and the PYP context. She emphasized the need for consistency in her assessment and grading, adhering to the rubric and providing specific grades due to the testing culture and the PYP's emphasis on grades. Aya noted, "I think you have to be clear with them [[the students]]. You have to be very clear and consistent with what you gonna grade them on because this is what the students and teachers really care about" (Aya, II). Ecological factors from the PYP context, such as curriculum design, assessment policies, and the expectations of students, teachers, and coordinators, created a dissonance between Aya's beliefs and her practical actions.

Moreover, in an environment influenced by the competitive nature of college admissions, Aya was committed to fair and consistent grading practices, as illustrated by her systematic approach to writing assessments that involved error counting and reviewing each paragraph three

times. She implemented a structured system for grading deductions during TAP sessions: a half grade was deducted for content-related errors, such as an incorrect topic sentence or insufficient details, and one grade was deducted for every four mechanical or grammatical errors. Her verbalizations, “No, no grammatical mistakes. OK, it’s two mistakes, part of speech and/but won’t take marks,” and “one, two. OK, two punctuations” (Aya, TAP 1 & TAP 2), along with her clarification, “Four mistakes for each mark, but I’ll take half a mark for topic sentence or conclusion mistakes” (Aya, FI 1), reflect this meticulous approach. This grading method is exemplified in Figure 6.1, where a student with eight mechanical and grammatical errors received a two-point deduction overall.

Figure 6.1

Example of Aya’s Written Feedback in The First Task

The screenshot displays a digital writing environment. On the left, a document titled "True Friendship" contains the following text:

True friendship is about a connection between two strangers. It refers to give support, be honest, and care for each other. The first thing in true friendship is support and respect. True friends must be there for each other during both good and bad times. For example, they support each other successes and offer help during problems. You can bring a gift for your friend in their birthday. This support makes the connection between them stronger. Secondly, true friends have to be honest with each other. They have to tell the truth even if it is going to make your friend sad. Lying is not acceptable in true friendship. They have to trust each other words and hearts. For example, if my friend is in bad relationship with another friend I have to tell her. Finally, true friendship refers to taking care of each other. If my friend is in trouble, I have to be there for her. I have to see what she needs and give her what she needs at the time. For example, when my friend has a problem with her mother, I have to tell her what to do to make the problem less. In conclusion, true friendship is all about supporting, telling the truth and caring for each other. It is important to know who your true friends are because sometimes you think that this is your true friend, but it is not.

On the right side of the interface, a list of feedback comments is visible, each with a "mention or reply" button:

- giving (July 25, 2023 at 6:11 PM)
- being
- caring
- other's
- other's
- add a

At the bottom right, a "Feedback" section shows a score of 23/25 and a "Points" section showing "No points".

In addition to the grading practices previously discussed, Aya implemented a multi-read approach to ensure a thorough evaluation of each paper. She would first read and correct a

paragraph, then re-read it to grade and count errors. During her TAP sessions, Aya meticulously double-checked corrections before grading, paying special attention to the topic sentence and conclusion, as indicated by her saying, “I’ll take another look at the topic sentence” or “Let’s take a look at the conclusion” (Aya, TAP 1). She referred to this as “the second round,” which she clarified during her follow-up interview: before finalizing a grade, she conducted three rounds of review. The first round was to identify student errors, the second to catch any of her own oversights to maintain fairness, and the third to assess consistency and consider class-wide performance levels, adjusting her strictness if necessary (Aya, FI 1).

Moreover, while Aya placed a higher value on idea generation and content (Aya, II, Excerpt 6.3), she acknowledged the necessity of adhering to the PYP rubric, which required attention to spelling, grammar, and formatting. Despite her belief that spelling is a “minor” error, she diligently corrected every spelling error to comply with the rubric’s explicit criteria. Her comprehensive review extended to all aspects of writing as delineated by the rubric, including part of speech, subject-verb agreement, and article use, often issuing specific corrections such as “Add a and delete plural s,” and “add the.”

Additionally, curricular and time constraints emerged as discordant factors between Aya’s pedagogical beliefs in assessment and her enacted practices. Aya lamented that the syllabus was laden with what she considered unnecessary assessment methods, which curtailed the time available for what she valued more, such as direct writing practices. Reflecting on the PYP’s requirements, Aya emphasized, “We should focus on writing, only writing” (Aya, II).

Aya acknowledged the value of peer assessment but noted the impracticality of its implementation due to time restrictions, saying, “We don’t really have time for peer assessment.

Usually, they do it, but we don't have time. And the students are also busy. They have a lot to do in this course and other courses, and it's really hard to do the peer assessment" (Aya, II).

Time constraints also affected Aya's ability to provide written feedback. She indicated that the multitude of tasks required of both teachers and students often meant there was insufficient time to review every piece of student writing, particularly the first drafts. Aya described this challenge, stating:

Excerpt 6.12:

So, for the written feedback, it's a little bit (.) I'm gonna say, complicated. It depends on the time. Cause they're asked to do a lot of things, and we are asked to do also a lot of things, so sometimes, we don't really have time to look at every practice before the real or the graded task... So I give written feedback. But as I said, it depends on the time. (Aya, II)

When inquired about what "things" competed for her time, Aya clarified that she would prefer to prioritize direct writing tasks and allocate more time to them rather than dispersing efforts across various indirect tasks and multiple-choice questions about grammar. She further elaborated:

Excerpt 6.13:

I feel what takes time/ Cause we have a number/ a lot of/ maybe if we reduce the number of the tasks and exams, we would have enough time to/ I had to focus/ let's say, I would prefer to teach them two tasks about writing very well, better than giving them four tasks that some of them are multiple-choice/ about grammar. And then, we have to finish everything on time, and we don't have enough time. I'm not gonna say it's a headache, but it requires energy and focus. (Aya, II)

Furthermore, there were instances where tensions coexisted with harmony, such as in her beliefs and practices regarding individual approach to grading and feedback. Aya was able to align her beliefs about providing individual verbal feedback with her practices; however, with grading, she was not able to do so. She needed to follow what was imposed on her by the PYP policies to be consistent in grading with other teachers. The harmony between feedback beliefs and practices will be discussed in the following section.

6.7.3 Harmony between Aya's Writing Assessment Beliefs and Practices

Aya appeared consistent and clear when reporting her writing assessment beliefs and beliefs in practice. Her beliefs regarding the purpose of writing assessments, evaluation criteria, assessment methods, feedback, and grading, all seemed in alignment with each other, demonstrating a core belief in assessing students' individual progress and providing individual evaluation and guidance of content-related aspects. In terms of her actual writing assessment practices, some were aligned with her beliefs and beliefs in practice, specifically in the area of feedback.

During the interview, Aya discussed her preference for providing in-class verbal feedback, focusing on clarity of ideas, and her actual practices aligned with what she reported (Aya, II, Excerpt 6.6). In Aya's two observed classes, she circulated around the class to provide students with immediate individual verbal feedback. During the 50-minute outlining sessions of the paragraph writing tasks, Aya gave the students almost 20 minutes to write, and then she started circulating around the class to monitor their progress and provide verbal feedback. Aya focused on each student, asking them to show her their topic sentences and inquiring, "Where is your topic sentence?" (Aya, OBS 1). She then offered direct verbal feedback on their sentences, pinpointing the specific errors, and providing suggestions on how to rephrase their topic sentences. For example, she commented on a student's definition paragraph outline, "You can use the phrase 'refers to being' and write your sentence about true friendship. What does true friendship mean to you?" (Aya, OBS 2) and on another student's compare and contrast paragraph, "You can say 'There are similarities and differences between eating at home and eating in a restaurant'" (Aya, OBS 1). Aya appeared to place particular emphasis on ensuring that the students wrote a correct topic sentence in several instances while interacting with them.

Additionally, she indeed used the students' L1 to enhance their understanding of her verbal feedback as she explained in Excerpt 6.6, using phrases such as “اكتبها بالطريقة” ((write it this way))” (Aya, OBS 1) and “كذا غلط غيري الجملة” ((this is wrong, change the sentence))” (Aya, OBS 2). Furthermore, if a student wrote the topic sentence correctly, she would encourage them, stating, “Excellent topic sentence,” “This is very good” (Aya, OBS 1), and “Start writing your main points” (Aya, OBS 2). During the outlining sessions, she focused on individual students only, not providing group verbal feedback, and made sure to review the outline of each student. She did not use a pen to provide feedback; it was purely verbal, with suggestions and corrections provided orally. Her verbal feedback practices aligned with her core belief in individualized approaches to assessment.

Aya's written feedback practices were in line with her belief in the importance of content in compositions, ensuring that responses accurately addressed the prompts and that the main ideas were clear. This was evident during the first TAP session, where she evaluated a student's compare and contrast paragraph, showcasing her focus on content assessment and the student's adherence to the prompts:

Excerpt 6.14:

P: [[Reads student paper]] There are several significant similarities and distinctions between meals eaten at fast food restaurants and those cooked at home.

P: طيب ((OK)) It's a good sentence بس ليش خلتها ((but why she made it about)) fast food. She could have said, “eating in a restaurant” زي الـ ((like the)) question. (Aya, TAP1)

In Excerpt 6.14, Aya questioned the student's deviation from the prompt. The student's topic sentence contrasted ‘meals at fast food restaurants’ with ‘meals cooked at home,’ missing the actual topic of ‘eating in a restaurant versus at home.’ Aya wrote a feedback comment for the student, stating, “The title or the topic was about eating in a restaurant.”

Moreover, as indicated earlier, Aya's actual written feedback, while attentive to linguistic accuracy, such as spelling and grammar, also concentrated on the substance and originality of

students' work. For instance, as shown in Figure 6.1, she addressed content in two remarks. Initially, she prompted the student to elaborate on her topic sentence on 'leadership' by adding three main points. Specifically, for the third main point, she sought a deeper exploration, advising the student to "add more details to [her] example," thus emphasizing the need for specificity and elaboration.

Figure 6.2

Example of Aya's Written Feedback in the Second Task

The screenshot displays a writing task interface. On the left, a student's paragraph titled "Leadership" is shown. The text is as follows: "Leadership refers to guiding and motivating a group of people to do a specific task or job." "A group leader is a person who give direction and instructions to other people. Firstly, leadership refers to good communication skills. A good leader listens to the ideas and problems of team members. For example, a team leader in a group project must listen to everyone's ideas and put a plan to do these ideas. Then tell every member what to do. Secondly, leadership means to motivate the group to do good work. A leader leads and encourages team members to do their best. For example, a good coach of a sport team encourages the players to train hard and play together as a team in a game. Finally, leadership is making big decisions when the team is lost. For instance, if your team members do not know what to do, a good leader must tell them what to do." "In conclusion, leadership is about good communication, encouraging the group, and making important decisions. It is important to learn what is leadership to find a good job and to succeed in your education and life." On the right side of the interface, there are four feedback comments from Aya, each with a "reply" button:

- Comment 1: "add all your main ideas to your topic sentence"
- Comment 2: "gives"
- Comment 3: "motivating"
- Comment 4: "add more details to this example"

Analyzing Aya's other feedback comments on the student paragraphs also revealed her emphasis on content and structure. Her comments, such as "You can explain the difference separately and then give the supporting details" and "You can add more details for each point here," indicated her emphasis on a detailed and methodical exposition of ideas. She sought clarity in the students' explanations of differences, a critical aspect of comparing and contrasting paragraphs. Furthermore, Aya's feedback, "Is this the topic sentence for the paragraph? If so,

where is the definition part?” signaled her insistence on a clear thesis that guides the reader through the paragraph. The suggestion “you could’ve used a transition to explain the difference, e.g., on the other hand,” demonstrated her attention to the flow of the essay, guiding the student toward the smooth and logical progression of ideas. Lastly, comments on the concluding sentence, such as “you can mention the differences briefly,” implied her focus on the quality of the conclusion.

In her interviews, she disclosed that her written feedback was predominantly direct, especially concerning formal errors like mechanics and grammar. Her indirect feedback, on the other hand, tended to address content, pushing students toward the appropriate corrections. An example of her direct method is shown in Figure 6.2, where she rectified a subject-verb agreement error by replacing ‘give’ with the correct form, ‘gives.’

6.8 Summary of Aya’s Case Findings

Figure 6.3 encapsulates the complex interplay between Aya’s writing assessment beliefs, practices, and the ecological systems that shape them. The graphic delineates a belief system rooted in the importance of individual progress and clarity of content over mere linguistic accuracy, both at the core and peripheral levels. It also highlights Aya’s personal journey as an L2 English learner and how this microsystem factor informed her belief in the continuous nature of learning, steering clear from the “strict” and “harsh” assessment culture of the PYP. The interplay between these personal beliefs and her pedagogical practice revealed a thoughtful process and empathetic approach to considering students’ backgrounds as L2 learners when assessing writing.

Aya’s core belief in evaluating students’ content based on individual progress rather than aggregate achievement was both a guiding principle and a practice. This belief extended to her

emphasis on content over spelling and grammar and her preference for in-class individual verbal feedback. Her peripheral beliefs, such as the value of peer assessment and the unnecessary nature of exams, further underscored her focus on the writing process as a developmental journey rather than a test of proficiency.

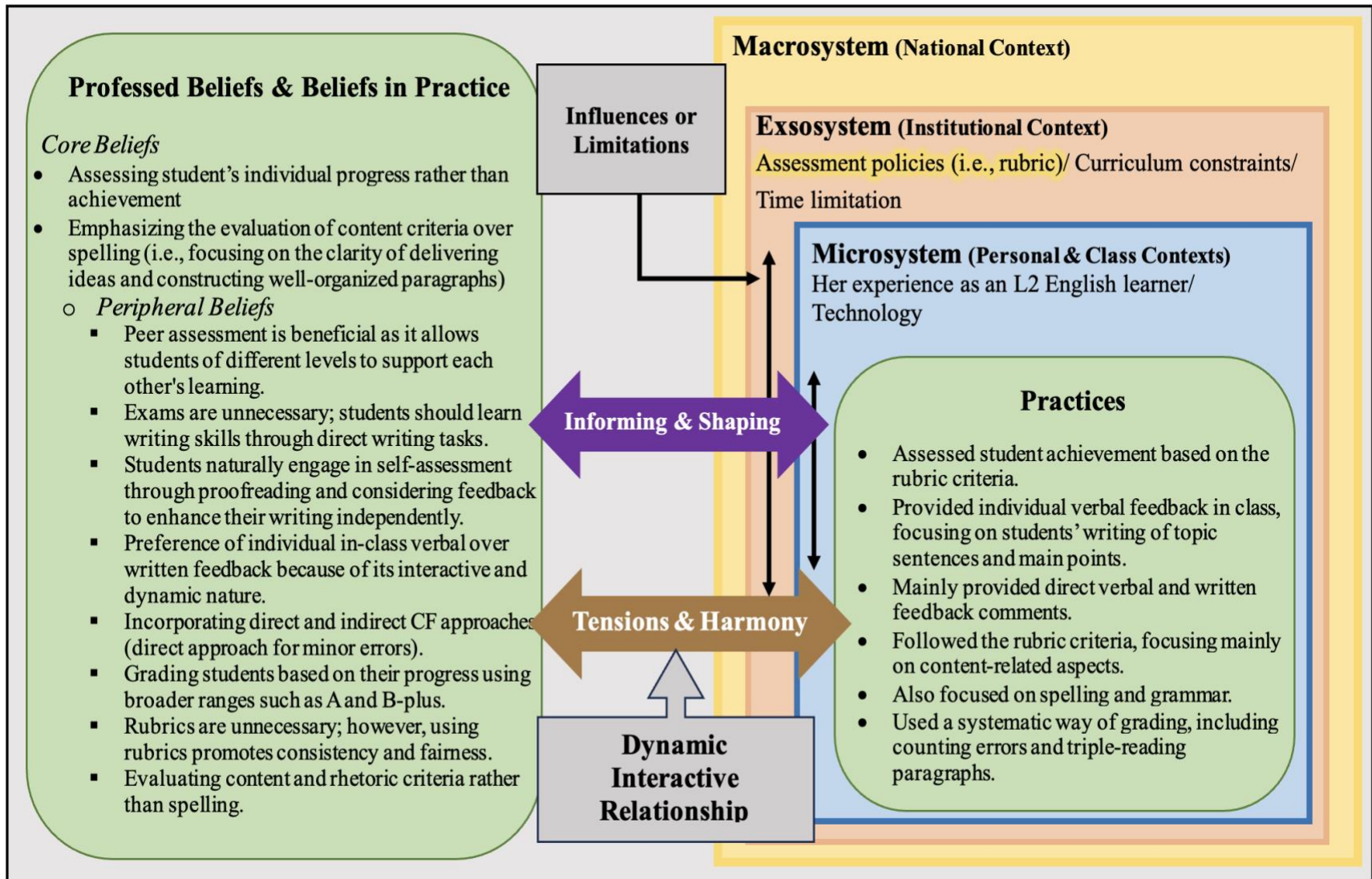
However, the visual also highlights tensions arising from the exosystem (orange box), with institutional constraints like rubric-based assessment policies, curricular demands, and time limitations challenging the application of her beliefs. Aya adhered to the rubric criteria, including spelling, and followed a systematic approach to grading, such as counting errors and thoroughly reading the paragraphs multiple times to ensure fairness, as the PYP context, including students and teachers, focused on grades. Time limitations hindered Aya from implementing peer assessment or allocating time to focus on direct assessment methods such as writing tasks and providing comprehensive written feedback.

Despite these tensions, Aya's approach had a discernible harmony; her practices closely aligned with her beliefs, particularly in her preference for providing in-class verbal feedback. Aya circulated the class and provided verbal feedback on students constructing topic sentences and main ideas. Her verbal and written feedback concentrated on the clarity of ideas, a practice she consistently carried out despite systemic pressures.

In the context of Figure 6.3, the complexities of Aya's beliefs and practices were not in a vacuum. Still, they were actively shaped and sometimes constrained by the various layers of the ecological system surrounding her professional environment. The diagram thus serves as a map, charting the territory where Aya's educational ideology and the reality of her classroom intersect, revealing a fluid and responsive belief system that adapts to the challenges and opportunities within her educational setting.

Figure 6.3

Summary of Aya's Case Findings



CHAPTER 7: THE CASE OF BATOOL:

The Value of Formative Methods and Self-Assessment

7.1 Background for Batool

Batool, an English lecturer between 25 and 30, has been teaching at the University for almost five years at the time of data collection. She has a bachelor's degree in English Language from KSA and a master's degree in TESOL from the US, with six years of teaching experience. Batool commenced her teaching career at a PYP in another public university, serving as a teaching assistant for approximately a year and a half. During that time, she taught students of various language proficiency levels basic English skills, including reading, writing, listening, speaking, and grammar. She explained,

Excerpt 7.1:

I used to teach foundation year students. So, it's basically general English, all the skills. It integrated English reading, writing, oral, and grammar, all in one book.... I have taught most levels, to be honest. They used to have four levels: 1, 2, 3, and 4, which is the advanced one. (Batool, II)

After completing her master's, she started teaching English at the University. She is responsible for teaching basic English skills to undergraduate medical students during their PYP. Throughout these five years, Batool explained that since she joined the University, she taught grammar and writing courses to diverse students' levels, in both the unified and nursing sections.

Batool recounted her English language learning experience, noting her middle school education as her first encounter with the language, as detailed in Excerpt 7.2 below. She never attended a private or international school, yet she was immediately captivated by English despite it not being widely spoken in her surroundings. She observed that Saudi public schools at the time merely covered the fundamentals of English, focusing on basic communication skills for everyday situations.

Excerpt 7.2:

Well, honestly, I was/ I've been studying in a public school. I've never been like in any private or international school. But, you know, in public schools, we used/ The first time I got introduced to the English

language was in middle school. So, once we took this very new subject, it attracted me... So, you know, in the public schools, they didn't really focus on the English language back then. Because, like, all they cared about was the basics, ABCD and, like, small conversations, if you want to go to a restaurant, doctors, and very easy things. (Batool, II)

In the initial interview, Batool indicated that she learned to write in English individually and outside of an academic setting. Later, her English writing skills were honed into formal academic skills during her bachelor's studies. Her initial engagement with English writing was driven by social motivations, as she found it "very cool to use the English language with [her] friends while texting." This informal practice with abbreviations like "LOL" piqued her interest in writing more in English despite acknowledging that it was "full of grammatical mistakes." This phase of her learning was characterized by a desire to conform to peer norms and be considered "cool," especially in texting using a "Blackberry." The transition to academic writing in her bachelor's years marked a significant shift. During this period, Batool was first introduced to "the basics of academic writing," highlighting a structured approach that necessitates adherence to "special rules" for formal writing. Her journey of learning to write underscored the interplay between informal, socially motivated language learning and formal, institutionalized education (Batool, II).

At the time of the study, Batool was responsible for teaching a writing class comprising 24 students from the nursing section. She engaged with her students for three hours per week, focusing on the writing component of the integrated reading and writing course, English II.

7.2 Writing Assessment Purposes: Formative Assessment

Batool demonstrated a core belief regarding the purpose of writing assessment that was primarily formative, that is, informing, directing, and improving students' learning. Batool appeared to establish a connection between the purpose of writing assessments and the

importance of her role in providing proper feedback. She believed that feedback enables students to learn from their errors and mistakes, thereby helping them learn and progress. She explained,

Excerpt 7.3:

because the students would not/ they are still learning. They won't know what mistakes they've made. The things that they need to improve, unless you give them the proper feedback, provide them with good, reasonable feedback that they would understand, and also learn from their mistakes. Then they will apply it and of course, like, they will improve, and they won't repeat the mistakes the other time they have a task. (Batool, II)

In Excerpt 7.3, Batool underscored that students in their learning process, are not inherently aware of their errors or areas needing improvement. To address this, she emphasized the importance of the teacher's role in assessment, such as providing constructive feedback that is comprehensible and valuable to the students. By doing so, Batool believed that students can learn from their errors and integrate these lessons into their writing. She seemed to imply that this iterative process of assessment, feedback, and improvement leads to enhanced writing skills and a reduced likelihood of recurring errors.

Additionally, when asked about the influence of assessment on students' learning, Batool addressed the question differently and discussed her approach to making students aware of their progress throughout the trimester. It appeared that she interpreted the question as relating to how to make students assess their learning and progress or how she determines if students have grasped the material she taught. She replied,

Excerpt 7.4:

Okay, so this is what I tell the students. In the first class, the introduction class, I tell them, "You know, in my class, I want to make sure that you can see and sense your progress." I tell them, "You have to have a reference for your writing." I ask them to write a paragraph, whatever they want to write. I take this piece of writing that they've written, and at the very end of the semester, I ask them to compare their writing with the paragraph they have written on the first day... Yeah, I want them to, like, see themselves how they've improved and how their writing improved and got to, like, this academic level. So, this is what matters to me in assessment. The things that I'm teaching them explicitly. I want them to be able to/ I want to see it, to see it in their writing. (Batool, II)

In Excerpt 7.4, Batool started by conveying a clear intention to help students perceive their progress throughout the course. By emphasizing the importance of a writing reference, Batool

seemed to believe in fostering a sense of continuity and development in students' work through her teaching and assessment. Although Batool did not discuss the influence of assessment on students' learning, her discussion of her belief in practice of having students write a piece initially and then comparing it to their writing at the trimester's end is a tangible way to make students aware of their own progress.

7.3 Writing Assessment Methods: Formative Methods and Self-Assessment

Batool appeared to believe that formative assessment is more effective for students' learning and measuring their performance than summative assessment. Additionally, in light of her core beliefs in formative assessment, she also seemed to highly value self-assessment as a means to increase students' awareness of errors and foster independence through the revision and editing of their own work.

When Batool was asked to share her beliefs about the effective types of writing assessment to assess students' performance, she stated: "I believe formative assessment is the best way" (Batool, II) to measure students' performance. In Excerpt 7.5, aside from her preference for formative assessment, Batool compared the two types of assessment—formative and summative—while explaining her reasons for favoring formative assessment, stating:

Excerpt 7.5:

Formative assessment, like, continues assessment, because if there's only, like, a summative assessment, the ones that you only give them a midterm or a final, the students will not be able to learn from their mistakes because it's really normal to make mistakes. Like, I want them to make mistakes during continuous assessment. I mean through the whole/ throughout the trimester because they also want to learn from their mistakes. So, what's my point? From my point of view, if you give them only, like, summative assessment once or twice per trimester, it's gonna be/ it's not fair for the students, not even for me, to judge their level. So, I believe formative assessment is the best way. (Batool, II)

Batool asserted that relying solely on summative assessments, such as midterms or finals, might hinder her ability to judge the students' level or their ability to learn from their errors. By employing formative assessments throughout the trimester, Batool appeared to aim to create an environment where errors and mistakes are seen as a natural part of the learning process, stating

that “it’s really normal to make mistakes.” Batool’s viewpoint seemed to center on the idea that formative assessment allows students to grow and develop by learning from their errors, while infrequently held summative assessments can unfairly assess their progress or level.

Once again, Batool’s preference for formative assessment clearly reflected her commitment to creating a learning environment that encourages students to make errors and learn from them. This preference seemed to closely interact with her core belief in enabling students to derive knowledge from their errors. It became evident that Batool highly values the role of errors as a valuable learning tool in the development of writing skills. Both Excerpts 7.3 and 7.6 show that making mistakes as an integral part of the learning process is fundamental aspect of Batool’s writing assessment belief system.

Another piece of evidence highlighting her appreciation for errors emerged when I asked her about the assessment methods she would use if she were to have her own writing class. She responded,

Excerpt 7.6:

Honestly, I love writing tasks. I would keep/ I want the students to write as much as they can, like, throughout the trimester, write paragraphs about anything, graded or not. Like, if you give them only writing, they will learn writing. I want them to try and make mistakes. It is okay. They will learn and get better and better, so I will make them write a lot. (Batool, II)

Batool’s response in Excerpt 7.6 highlights two aspects of her writing assessment belief system. Firstly, it shows her genuine enthusiasm for tasks that require students to write, emphasizing the importance of encouraging students to write extensively throughout the trimester, regardless of whether the writing is graded or not. This approach demonstrated her belief in the value of practice and continuous improvement. Additionally, it reflected Batool’s attitude towards mistakes as a natural part of the learning process.

Furthermore, when Batool was asked to discuss her assessment methods in the initial interview, she addressed various methods, stating, “Honestly, I like to use more than one way”

(Batool, II). During the interview, I did not need to provide her with a specific list of methods in the question; she promptly answered the question and referred to both self- and peer assessment. She also elaborated on her beliefs regarding feedback, as well as her beliefs in practice concerning the implementation of those writing assessment methods.

Firstly, the initial interview data showed that Batool appeared to highly value the implementation of self-assessment. She mentioned on several occasions during the interviews that she frequently encourages students to make errors, assess them, and learn from them, as identified in Excerpts 7.5 and 7.6. Additionally, Batool deliberately employed the concept of ‘self-assessment,’ stating, “Honestly, I always encourage them to self-assess. I encourage them to be more self-independent” (Batool, II). She elaborated on her response and explained her beliefs and beliefs in practice regarding self-assessment, stating:

Excerpt 7.7:

So, I always tell this to the students from the first task. I would point out most of the mistakes that they have written in the first draft. But then, I will teach them to be more independent. So, I will tell them, “Later on, I’m not gonna point out every single mistake that you have made.” (Pause). Like, for example, whatever mistakes they have in the first draft, I will point out, and then I will ask them to check the rest of the mistakes themselves. So, this way, like, I save time, and I teach them to be more independent. (Batool, II)

In Excerpt 7.7, Batool’s response revealed a thoughtful and strategic approach to writing assessment, particularly feedback. She seemed to believe in providing detailed feedback on the first draft, helping students identify and rectify their errors. However, what appeared to set Batool’s approach apart was her emphasis on fostering independence in students. By gradually reducing the amount of feedback she provides and encouraging students to take responsibility for their revisions, Batool seemed to believe in instilling crucial skills for self-assessment and improvement.

On another occasion, when she discussed her beliefs and beliefs in practice about feedback, Batool addressed the topic of self-assessment once more. She mentioned that she

instructs her students not to rely solely on her feedback but to also identify their errors independently before submitting their final drafts, explaining:

Excerpt 7.8:

I don't want them to depend completely on the things that I point out in the feedback sessions because I tell them, "Not every time I will be/ for example, sometimes I might not see the mistakes that you have written here, done here, in the first draft, but I might see them in the final draft. So don't depend on the things that I give you in my feedback completely. No, I want you to go beyond, revise, and edit many times yourself." (Batool, II)

In her first follow-up interview, she also referred to this core belief. When I asked her about her approach to written feedback, she stated, "I'm trying to point out all the mistakes, because this is, like, their first draft. Maybe in the second assignment, I'll leave some stuff for them to point out" (Batool, FI 1). Subsequently, in the second TAP session, there was an instance in which Batool verbalized, "[Refers to a student's grammatical mistake] she should know this. I'll leave it. I told them about this" (Batool, TAP 2). Batool chose to not correct the student's mistakes concerning subject-verb agreement and expected the students to notice the mistake and correct it by herself.

Furthermore, when discussing the factor of time constraints, as in Excerpt 7.7, Batool repeated that she teaches her students to self-assess their paragraphs "to save time and to teach them to be more independent" (Batool, II). In essence, Batool's belief in self-assessment aimed not only to save time but also to empower students to become more self-reliant learners who can critically assess and improve their own writing. Her advocacy for self-assessment seemed to be rooted in the belief that students ought to be responsible for their own learning and that they are capable of enhancing their writing skills independently.

Secondly, Batool also seemed to value peer assessment implementation in her assessment practices. Her concept of peer assessment also appeared to center on teaching students to identify and correct errors independently. She explained her beliefs and beliefs in practice regarding peer assessment, stating:

Excerpt 7.9:

I also, sometimes when I ask them to write a short piece of writing, I ask them to peer check because I told them/ I always tell them this “Sometimes, you can’t spot the mistakes that you’ve written yourself, but when someone else reads your writing, they would find out more mistakes”. And they would/ of course, this is not for major graded tasks. But as I told you, when I give them any piece of writing other than the graded task, I ask them to peer review. And peer check is a really effective way I’ve noticed this. The students get excited and learn how to spot mistakes. So, yeah, but again, peer review or peer feedback, if it is not a graded task. (Batool, II)

Batool’s response in Excerpt 7.9 demonstrated her belief in the promotion of peer assessment for non-graded tasks. She seemed to place great value on having students review each other’s work, particularly for the purpose of learning error identification and emphasizing the significance of learning from them. Her approach appeared not only to promote a sense of collaboration and shared responsibility among students but also to heighten their awareness of assessment and evaluation. However, in contrast to self-assessment, Batool appeared to believe that peer assessment is most suitable for assignments with lighter grading, while she perceived self-assessment as a valuable skill applicable to all types of tasks that students should acquire.

Finally, in response to the question about assessment methods, Batool also discussed feedback. She explained that part of using “more than one way” involved integrating both verbal and written feedback, stating, “I use both oral and written feedback. Like, I use them a lot, most of the time, the written and the oral feedback” (Batool, II). However, she did not elaborate further on her feedback beliefs until I specifically asked her about them, which will be discussed further in the following sections.

7.4 Criteria for Evaluating Writing: PYP Rubric Influence

When Batool responded to the question about what she would focus on when assessing students’ paragraphs, she seemed to share her beliefs in practice within the PYP context. She stated, “Honestly, everything...” (Batool, II). She elaborated on her response, explaining,

Excerpt 7.10:

Honestly, everything, organization, formatting, and even content/ because I have already explained to the students the concept of editing and revising, and what do they mean? So, they know how to correct their own

mistakes and to focus on everything. I don't want them only to focus on the grammatical and spelling mistakes and this stuff. No. I want them also to focus on the content and content itself. Are all ideas related to the topic sentence? Is their writing going smoothly and flowy? So, for me as a reader to read it in a smooth flow way or not. So, I'm concentrating on all sides, all types of mistakes. (Batool, II)

Batool's complete response in Excerpt 7.10 demonstrated a comprehensive approach to evaluating students' writing, aligning with the criteria outlined in the PYP rubric, such as paragraph organization and formatting. Initially, it was ambiguous whether the criteria she mentioned originated from her own assessment belief system or the rubric, as she did not reference additional criteria beyond those in the rubric. Nonetheless, her detailed discussion on the significance of assessing not just grammar and spelling but also the substance of content—how students' ideas relate to the topic sentence and their coherence—suggested her personal belief in the importance of content evaluation. In the end, Batool reiterated this holistic focus by stating, “So, I'm concentrating on all aspects, all types of mistakes,” indicating that the rubric might indeed inform her comprehensive evaluation strategy. Batool also emphasized the comprehensive evaluation of students' compositions because she instructs them on these criteria, focusing on revision and editing techniques. Thus, she seemed to expect her students to attend to all elements taught in the English II course, mirroring her own focus on evaluation criteria that were dictated by the course curriculum.

7.5 Feedback: The Integration of Written and Verbal Feedback

Batool's belief system included various peripheral beliefs and beliefs in practice concerning feedback and grading, which seemed to be integrated with her core belief of formative assessment as a tool for student improvement and the significance of promoting self-assessment and error correction.

Excerpt 7.11 shows Batool's discussion of her belief in practice regarding providing feedback and grading students' papers. She explained that she invests time in reading and

evaluating each sentence carefully to be ‘fair,’ especially in the first drafts of their writing tasks. However, by doing so, she seemed to believe that students have a “chance to improve” and that they will learn not to repeat their errors. She explained:

Excerpt 7.11:

Honestly, it takes me a long time because, you know, I try to be as fair as possible. So, I read each and every single sentence I tell them, “Now it’s your chance to improve.” I tell them/ especially in the first draft, or practice paragraph. “I would be very picky about every single thing” because they know if it’s a practice, I would concentrate on every single word, every single mistake, everything. And I know that إن شاء الله ((if God wills it)) later on, they won’t repeat these mistakes. So, whenever we have a practice paragraph or first draft, I concentrate a lot, especially the first draft, because it’s their chance to improve before things get finally graded and can’t be changed. (Batool, II)

Batool appeared to believe that offering a comprehensive assessment in the first drafts helps students identify areas needing remedial action and reduces future errors. She elaborated on her response: “So, when I give them the graded task, they already learned from the mistakes, they wouldn’t repeat the mistakes that they have made” (Batool, II). Interestingly, it seemed that Batool believed that students inherently make mistakes and that once those mistakes are corrected or pointed out, students will immediately learn not to repeat them.

Additionally, Batool perceived feedback as ‘an interactive session.’ In Excerpt 7.12 below, it seems at first that she was referring to the interactive nature of verbal feedback sessions. However, the rest of her response indicates that she also values the integration of both written and verbal feedback, wherein she typically provides written feedback to her students and then, allocates specific time for each student to meet and review her written feedback comments together. Her response indicates that she believed in supporting her written feedback with an interactive verbal session for each student. She explained:

Excerpt 7.12:

Yeah, feedback, to me, is an interactive session. The students need not only, like, go over or through the comments that I have written, and that’s it. No, it’s/ I want it to be an interactive session. I sit with every student, but I try to make it very organized. I ask them, and I make a schedule table of students’ names. I call their names, student by student. I give certain timing because, you know, we only have a short time. So, for example, each student, let’s say five to ten minutes. I sit with her, and we go over all the comments together. (Batool, II)

Furthermore, Batool's core belief in the value of engaging the students in error identification and self-assessment appeared in the discussion of her verbal feedback practices. She stated, "In this feedback session, *they* [[students]] don't only read the comments that I have written, they realize other mistakes that I did not point out" (Batool, II). In Excerpt 7.13 below, she provided examples of how students usually identify other errors in their writing during this verbal feedback session, stating:

Excerpt 7.13:

When I point out a few mistakes/ let's say grammatical or capitalization, or let's say punctuation mistakes, they find other ones. So, they tell me, "Miss, I have a mistake here," and then when they read more, they say, "Miss, I want to change the placement, the word choice of this word, the idea can be rewritten, the sentence structure." (Batool, II)

To clarify, at the end of her response, Batool underscored the significance of interactive verbal feedback in fostering self-awareness and self-assessment. Her verbal feedback sessions appeared to focus on aiding students in independently identifying errors and understanding how to improve their writing. She stated, "So, I make sure it's an interactive feedback session because I want the students not only to go over it but also to know their mistakes and how to correct their writing to make it better" (Batool, II).

Batool also discussed her beliefs in practice related to her written feedback approaches. She seemed to believe in employing a combination of written feedback comment approaches, including highlighting, commenting, and using error codes. She also mentioned using positive reinforcements, which she believed can positively affect students' motivation and enthusiasm.

Excerpt 7.14:

Okay. You will see my feedback that I use error codes, highlight, write comments to the students. So, yeah, it varies... So, my feedback varies between short sentences, questions, error codes, and of course, I use a lot of motivational words for the students, and they get really excited about this part. (Batool, II)

She further elaborated on her beliefs in practice of implementing those different written feedback approaches for various types of errors. For minor mechanics and grammatical issues like spelling, punctuation, and sentence structure, she seemed to believe in using error codes. For

more substantial concerns related to the content and clarity of ideas, she believed in asking questions and seeking clarification from the students, stating:

Excerpt 7.15:

So, yeah, it varies because I see some minor mistakes, the easy ones. For example, for spelling, I already gave *them* [[Students]] the error codes table. They know the shortcuts very well, and the abbreviations refer to what? So, I use error codes for spelling, punctuation, sentence structure, like, very simple things. However, if I find a mistake in the content itself, let's say the idea is not really clear, or the student is not conveying the things that she wants to deliver correctly. Here, I write a comment, and I ask a question, "Is this what you believe? Is this a main point? Is this a supporting sentence that is related to that? If so, you have to make it clearer, like, a very short explanation. (Batool, II)

7.6 Grading and Rubric Use: The Importance of Fairness

When asked about her beliefs and belief in practice regarding grading, Batool acknowledged its importance for "fairness" and for comparing students to each other. However, she appeared to believe that assigning specific grades induces stress in students. Therefore, she strongly emphasized the importance of never assigning grades to students' practice writing pieces, repeatedly using the word 'never.'

Excerpt 7.16:

Honestly, grades are important for fairness. You can compare the students to each other with grades. But, when I give them an external piece of writing, a free piece of writing, that they write, I never give them a grade. Never. I never. Because I don't think that/ you know, because grades are attached to stress for the students, they focus on grades, and they want a specific GPA, and they have to do something about it. It's either this or that. They get really stressed about these things. So, for me, I want them to learn. So, whenever I give them anything to write besides the graded tasks, I never put a grade, only comments, positive comments, the things that they need to improve, but never, never a grade. (Batool, II)

Batool's rationale seemed to stem from the belief that grades may lead to an excessive focus on grades and GPA, detracting from the actual learning process. Consequently, she believed in creating an environment that is more conducive to learning by refraining from grading extra writing practices and instead offering constructive feedback and positive comments.

Lastly, Batool conveyed her beliefs in practice regarding rubric's use. She underscored the significance of making students aware of each rubric criterion as a way to prevent grade deductions. Batool detailed her approach, "I project the rubric on the board for the students to see. And, I go over every single criterion, each one. I tell them, 'If you want to get the full mark,

I want you to follow those criteria” (Batool, II). It was unclear whether this belief was a result of her own views or influenced by PYP assessment policies, especially since other participating teachers reported that they are required to review the rubric with their students prior to writing tasks.

Batool seemed to emphasize avoiding unnecessary grade deductions for her students, likely shaping her belief in practice of discussing the rubric before the writing tasks. In Excerpt 7.17 below, she explained that the main reason for insisting on adherence to the rubric criteria was to prevent grade loss, particularly in final writing exams evaluated by teachers unfamiliar with the students’ writing. Her belief in practice reflected a commitment to clarifying the grading process, ensuring students understand the criteria for assessing their writing, and providing students with explicit guidelines for success to enhance their accountability.

Excerpt 7.17:

I usually tell them this “You have to follow the rubric. You have to follow the rules. Why? Because you don’t know. Maybe, maybe, just maybe, someone else will grade your writing in the final exam. And she doesn’t know, the teacher herself, that you are a really good student, and you made an honest small mistake, and she would forgive you about this one or ignore this one. No, I want you to follow the rubric because once you do, you are sure that whoever corrects your writing, will give you the mark, the score that you deserve”.
(Batool, II)

7.7 Ecological Factors Influencing Batool’s Writing Assessment Beliefs and Practices

The analysis of Batool’s data revealed that her self-directed language learning experiences were instrumental in shaping her core beliefs regarding self-assessment, aligning closely with her assessment practices. Her individual learning journey significantly influenced her emphasis on fostering student motivation and autonomous error correction. These principles were evident in her written feedback and the assessment approaches observed in her classrooms.

Furthermore, curricular and time limitations were predominant factors that Batool identified, posing significant challenges to the implementation of her assessment beliefs. The dense syllabus curtailed the feasibility of additional, non-graded writing activities that could

provide insights into students' comprehension and constrained the time available for detailed verbal feedback. Other contributing elements to the belief-practice tension encompassed the assessment rubric, although Batool did not explicitly reference these in her discussions.

7.7.1 The Influence of Batool's Learning Experience on her Writing Assessment Beliefs and Practices

Batool's English learning experience narrative captured the essence of a self-motivated learner who embarked on a learning journey within the confines of a public school system that offered only the rudiments of the language. As mentioned in Excerpt 7.2, she stated, "I've never been like in any private or international school," which contextualized her initial exposure to English as devoid of private educational environments. She elaborated that her attraction to English was immediate upon its introduction in middle school, not merely as a linguistic system but as a gateway to a "new culture and everything." This perception of English as an alluring and holistic experience rather than just a school subject propelled her self-directed learning.

Despite the public school system's emphasis on foundational elements—"the basics, ABCD and like, the basics, small conversations"—Batool's intellectual curiosity prompted her to venture beyond the prescribed curriculum. She reflected, "Most of my learning was self-learning," highlighting her proactive approach to language learning. Consequently, Batool reported her initiative in expanding her knowledge by buying books in English and starting "to read them to know beyond the course book that they gave us" (Batool, II). This initiative illustrated her intrinsic motivation to push the limits of her educational setting. Her drive to self-educate, spurred by an inner desire to explore the vast facets of English, emphasized an essential aspect of her pedagogical beliefs about language learning—learner autonomy.

Thus, Batool’s experience stood as a powerful testament to the transformative impact of self-directed learning in language mastery, reflecting her core belief in developing students’ writing skills through self-assessment and the autonomous practice of identifying and correcting errors to foster progress and improvement. Batool’s response to a specific question about the influence of her learning experience underscored the idea that students can develop through motivation, independent practice, and self-assessment. She elucidated,

Excerpt 7.18:

Honestly, I am not sure/ my learning experience, like, influenced my assessment. But I think you have to keep your students motivated all the time. Because if you’re motivated, you will learn, look for stuff in English, read books. You will try to learn more independently and, like, not rely only on schools or classes. (Batool, II)

Batool’s narrative indicated a nuanced relationship between her past autonomous language learning experiences and the emphasis she placed on motivation in her teaching and assessment practices. Initially hesitant about the influence of her learning experience on her assessment beliefs, she eventually supported the idea that motivation was crucial for fostering learner autonomy and self-determination. She had engaged in motivated learning behaviors, such as purchasing books and actively directing her learning process. Therefore, it could be inferred that her own educational journey subtly influenced her instructional emphasis on sustaining student motivation as a means to facilitate autonomous learning.

Furthermore, her actual approach to written feedback — emphasizing motivational comments and positive reinforcement— aligned with this insight, reflecting the broader impact of her learning experiences on creating harmony between her assessment beliefs and actual feedback practices. During her TAPs, Batool consistently aimed to provide students with encouraging comments after she completed corrections for each of the student’s six paragraphs. For example, she verbalized, “Okay. I’ll write something motivational” (Batool, TAP 1). In another instance, she said, “Motivational, motivational. Okay. I’ll write [[Types on student

paper]] Excellent {student name}. This is a very good piece of writing. Very well-written, organized, and structured. Keep it up, sweetie” (Batool, TAP 2).

Moreover, the analysis of the 20 student paragraph samples confirmed Batool’s use of positive feedback. Each sample contained an encouraging comment at its conclusion. While most were extensive, such as “VERY good job {student name}! I can see the amount of effort you put when writing this paragraph; bravo =)” and “Good job {student name}! I appreciate the effort you put when writing this paragraph. Excited to read the final draft =) Keep up the good work, sweetie!”, others were brief, often limited to phrases like “good job” and “Thank you.”

Furthermore, Batool’s propensity to integrate self-assessments within her pedagogy appears to have been influenced by her personal experience with self-directed learning. Her core belief in actively involving students in the process of error identification and self-assessment is reflected in her actual assessment practices. For instance, during her TAP sessions, Batool articulated,

Excerpt 7.19:

P: [[Reads student paper]] “E-books have availability more than textbooks because (Pause) on the Internet you can buy them at any time, also textbooks it is likely/ (Pause) it is likely to sell out.”

P: (Pause) Okay. I think this is run-on. (Pause) Yeah, run-on. (Pause) Okay. Two subjects. (Pause) Subject-verb agreement.

P: [[Types on student paper]] run-on.

P: I hope she can fix this. Maybe I can tell her in class. (Batool, TAP 1)

In Excerpt 7.19, while Batool identified three grammatical errors in the student’s sentence, she deliberately addressed only the issue of the run-on sentence. Her intention seemed to encourage the student to recognize the remaining errors independently or to address them in a forthcoming individual verbal feedback session. This pedagogical choice aligned with the belief articulated in Excerpt 7.11, where Batool professed the importance of identifying all errors in students’ first drafts. Nonetheless, in actual practice, she prioritized the student’s autonomous error discovery over comprehensive immediate correction. This pedagogical strategy indicated a stronger

commitment to developing students' self-assessment capabilities rather than adhering strictly to her belief in exhaustive error feedback at the draft stage. It is plausible that she intended to cover all errors comprehensively in a subsequent verbal feedback session, as indicated by her remark, "Maybe I can tell her in class."

During her second observed class, Batool reinforced the critical role of student engagement in revising and editing their work in response to her written feedback, prior to their one-on-one verbal feedback session. She instructed, "Ladies, when you receive your feedback, please focus on the comments and also other mistakes... Make sure to read your paragraph carefully before you come to class" (Batool, OBS 2). By encouraging students to address errors not explicitly identified in their first drafts, Batool was cultivating their capacity for autonomous error detection and correction.

Batool's classroom assessment practices further reflected her commitment to self-assessment; she facilitated independent brainstorming and outlining by students, refraining from providing guidance or suggestions. This approach was also exemplified when she encouraged a student struggling with sentence formation: "Try. Try to write it, even if it's wrong. It's okay. You can fix it later" (Batool, OBS 2). This incident underscored Batool's overarching assessment belief that values the cultivation of students' self-regulatory skills and their ability to learn from mistakes.

7.7.2 Tensions Between Batool's Writing Assessment Beliefs and Practices

Curricular focus and time constraints emerged as the primary factors causing discord between Batool's pedagogical beliefs in assessment and her actual practices. She observed that the English II curriculum's focus on non-writing disciplines impinged upon the time available

exclusively for writing instruction. This, in turn, constrained her capacity to incorporate supplementary writing tasks and allot substantial time for personalized verbal feedback.

Batool's commentary in Excerpt 7.20 below underscores the profound influence of curricular constraints, including the prescribed writing course book and syllabus, on her instructional approach to writing. While she recognized the curriculum's requirement for explicit instruction in reading and grammar as foundational to teaching writing, she suggested that an overemphasis on these elements, already addressed in other courses, might detract from the primary goals of effective writing instruction and practice. She elucidated further,

Excerpt 7.20:

Honestly, you know, the explicit part of the book is important. But sometimes, I feel we're not teaching them writing here. Like, as you can see in the book, it has grammar, it has reading parts, many other things. And we have to teach these parts. So, I would give them more time for the writing itself. And less time for the things that are already familiar to the students. Because now they're taking past tense in the writing course, present tense. Things that are not useful for them to know about in the writing course. They have a grammar course, so I assume that they're already familiar with these things. They could take these grammar lessons/ they can take them from the other courses and apply them in the writing course. Here we should concentrate more on writing, academic writing, and practicing writing more and more. Students want to practice more, but because the chapters are very long sometimes, we don't have the time. We need a chance to write, so only I have to/ It doesn't really make sense to be teaching writing explicitly and less practically. (Batool, II)

Batool's response indicates that curricular obligations compelled her to allocate a disproportionate amount of time to components she deemed secondary, like grammar rules, rather than to writing practice, which she held in higher regard. This divergence between her preference for regular writing practice and the curricular demands caused tension between her ideal pedagogical strategy and her actual instructional practices.

In a subsequent segment of the initial interview, Batool revisited the restrictive effect of time constraints on affording students supplementary writing exercises to evaluate their progress.

She expounded,

Excerpt 7.21:

I think the trimester is very short. We only have 10 teaching weeks, not counting the exam weeks. Like, we have a lot of things to do. So, I believe we don't have the advantage of giving the students lots of things to assess them, lots of practice, writing practices. Like, I think the course book itself has a lot of things to assess the students. But honestly, I don't have time to do that, to give the students/ to tell them to write a piece of

whatever. Like, to practice writing based on the things that I have taught. Just to see what they've learned, to see if they know the rules that we have taken today. So yeah, I think here, in this course, we have a lot of things, quizzes, many tasks. But they are all graded. We have many things. I don't think I'll be able to give them more, more free stuff, like free writing. (Batool, II)

In Excerpt 7.21, Batool articulated concerns regarding the brevity of the trimester, noting the constraint of having merely 10 teaching weeks, not accounting for the examination period. This compressed schedule imposed a significant burden due to the extensive syllabus and numerous assessments to be administered. As a result, she acknowledged that such time restrictions hindered her capacity to integrate supplementary non-graded writing evaluations or exercises, which could provide insights into the students' understanding of the material. Consequently, she found herself obligated to prioritize graded evaluations, such as quizzes and writing tasks, over her pedagogical preference for offering abundant writing practice.

Furthermore, Batool illuminated the substantial impact of time constraints on her practices of providing verbal feedback. She underscored,

Excerpt 7.22:

So, you know, I believe for any writing course, you have to give extra time because the feedback session itself needs time because students/ they ask a lot. They want to learn from every single mistake that they have done. So, they usually take more than the feedback session and their time. Especially that student's number is a lot in one class. So, they take time. Like honestly, I want to give them time, I want to, but I can't. (Batool, II)

Batool articulated the critical need for sufficient time allocation in writing courses, specifically for conducting one-on-one verbal feedback sessions that occupy a significant segment of the instructional period. She valued detailed verbal feedback as a fundamental component of assessment, acknowledging that students often present multiple queries and exhibit a keen interest in rectifying their mistakes. This aspect of her formative assessment belief reflected her dedication to enhancing student proficiency through feedback. Nonetheless, she conceded that the practical challenge of catering to a large class exacerbated the demands on her time, creating a tension between her belief in verbal feedback and the actual time available. This juxtaposition

of Batool's assessment beliefs against contextual constraints like time limitations encapsulated the broader struggle inherent in reconciling assessment ideals with real-world teaching conditions.

However, Batool appeared to have adapted her approach to mitigate the impact of time constraints on her verbal feedback practices, as demonstrated in Excerpt 7.8. In this Excerpt, she elucidated her strategy of encouraging students to become more self-reliant by revising and editing their work independently. Her rationale for this approach was twofold: "to save time and teach them to be more independent" (Batool, II, Excerpt 7.8). Furthermore, Batool's one-on-one verbal feedback practices indicated that she usually allocated specific time intervals, such as ten or five minutes, for each student for their verbal feedback sessions. This allocation of time can be seen as a deliberate adjustment to align her assessment practices with the available time, thereby striving to maintain the efficacy of verbal feedback within the constraints of the teaching schedule.

A divergence between Batool's professed assessment beliefs and her actual feedback practices was apparent in the focus of her written commentary. Despite professing a comprehensive approach to writing assessment — encompassing "everything, organization, formatting, and even content," as noted in Excerpt 7.10. The examination of 20 samples of Batool's students' paragraphs indicated her predominant emphasis on mainly linguistic form. This inconsistency might be attributable to the influence of the PYP rubric. As mentioned earlier, the rubric might have oriented Batool's attention to believe she was addressing all the evaluation criteria it covers, such as content. However, this interpretive discrepancy was evident in the analysis, suggesting a gap between her perceived and actual feedback practices. Figures 7.1 and 7.2 present the areas that Batool principally highlighted in her written feedback. In both figures,

Batool's attention was primarily directed towards form, specifically adhering to the rubric's evaluation criteria concerning mechanics and grammar.

Figure 7.1

Example of Batool's Written Feedback Focus in the First Task

The Differences and Similarities Between Textbooks and E-books

There are three similarities and differences between e-books and textbooks. First, there are a lot of summaries between e-books and textbooks. For example, they are similar on several pages, and all of them are written clearly and understandably, and can take them to any place you want, also all kinds of books for adults and children, which are related to cooking, are available in e-books as well as textbooks. Finally, they have a lot of differences between e-books and textbooks. For example, you can save e-books on your phone or computer, but this is not safe for your books because when any problem occurs, all books may be deleted, unlike textbooks, they will remain with you for a longer time, but may be damaged if you do not preserve them. E-books are cheaper and can be available faster than books. Textbooks, as well as textbooks, expire faster and are not available until after a long period, websites offer a free simplified version for purchasing the e-books, unlike textbooks, but you can borrow textbooks for a long time. There are summaries and differences between e-books and textbooks. In my opinion, I prefer e-books.

WW :Commented

WW :Commented

WW :Commented

SP :Commented

no comma :Commented

MW :Commented

Run-on :Commented

Run-on :Commented

CP :Commented

MW :Commented

Run-on :Commented

WW :Commented

Run-on :Commented

full stop :Commented

WW :Commented

not clear :Commented

There has to be a concluding word :Commented

SP :Commented

Thank you Amal =) :Commented

Figure 7.2

Example of Batool's Written Feedback Focus in the Second Task

A true friend

True friends are people who are close to you and help you when you need them. First of all, real friends stick around at any time and any place when you need them. They are present in the good time and the bad time, for example, friends who exist in good times and shares happiness but don't come in the bad times and share sad feelings aren't real friends because she/he just looking for fun, not for a good relationship. Second, real friends always like your success and never like your failure. The person who supports you to achieve your dream and reach your goal that we named a real friend. For example, my best friend encouraged me to enter a nursing major and when I told her about my admission to the university and I saw her happiness for me, so I knew at that moment she is a good friend. Finally, true friends who can rely on and give them full confidence. When you do something you are afraid of what people say, but you don't feel that feeling with your real friends because you are sure that they will not understand wrongly. All in all, real friends that really love you and you can rely on everything. In my opinion, it isn't easy to find true friends, on other hand, you can be a true friend to another.

CP :Commented

SV :Commented

no contractions :Commented

no contractions :Commented

pronoun :Commented

fragment :Commented

encourage you to succeed :Commented

they never wish you failure :Commented

VT :Commented

SS/ complex sentence :Commented

WW :Commented

SS :Commented

WW :Commented

not clear :Commented

Good job ya Haya! I appreciate the effort you put when writing this paragraph. Excited to read the final draft =) Keep the good work up sweetie!

7.7.3 Harmony between Batool's Writing Assessment Beliefs and Practices

As a result of her learning experience, the primary alignment between Batool's beliefs and her assessment methods resided in her commitment to motivating students towards self-learning and assessment, as reflected in her classroom assessments and feedback strategies.

Additional evidence of alignment was found in the analysis of Batool's written feedback, which confirmed the consistency between her professed assessment beliefs and her implemented assessment techniques. Batool reported the utilization of a multifaceted feedback strategy that encompassed highlighting errors, annotating with comments, and applying error codes for specific mistakes. She further asserted the efficacy of positive reinforcement in bolstering student motivation and engagement, as shown in Excerpt 7.14. This multifaceted feedback strategy was corroborated by the annotated student work displayed in Figures 7.1 and 7.2, where Batool systematically identified errors, employed error codes such as SP for spelling and VT for verb tense, and interspersed affirmative comments, thereby actualizing her stated feedback beliefs.

Additionally, Batool articulated the importance of complementing written feedback with individualized verbal feedback sessions for each student, as delineated in Excerpt 7.12. In her instructional practice, she meticulously planned and communicated the timeline for distributing the first drafts, complete with written feedback, and scheduled subsequent one-on-one discussions to address errors. She specified, "On Wednesday, I'll give you your first drafts back. On Thursday, I'll call your name and discuss your mistakes. I'll post the schedule on Teams" (Batool, OBS 2), thus highlighting the structured approach she adopted to ensure each student received tailored feedback within the constraints of the course schedule.

7.8 Summary of Batool's Case Findings

Figure 7.3 presents the findings from Batool's case study, revealing the interplay between her assessment beliefs, actual practices, and influencing ecological systems. The left side of Figure 7.3 delineates Batool's core belief in formative assessment and the integral role of self-assessment in enhancing student error detection and autonomous learning—a belief heavily informed by her self-directed learning experiences. This personal history affirmed her belief that learners are inherently capable of independent growth and posited the purpose of assessment and the teacher's feedback as pivotal in steering this self-guided learning trajectory, as depicted by the purple bidirectional arrow. Consequently, the interplay between Batool's own learning experiences and her core belief in self-assessment significantly influenced her pedagogical practices. She consciously refrained from intervening in students' independent brainstorming and editing processes, motivated them to revise and edit on their own, deliberately overlooked certain errors in students' first drafts to encourage self-discovery, and provided comments that served to encourage and motivate them. Additional beliefs she held included the acceptance of error-making as part of learning, the importance of frequent writing practice, the utility of peer assessment in non-graded activities, and the significance of individual verbal feedback to complement written feedback.

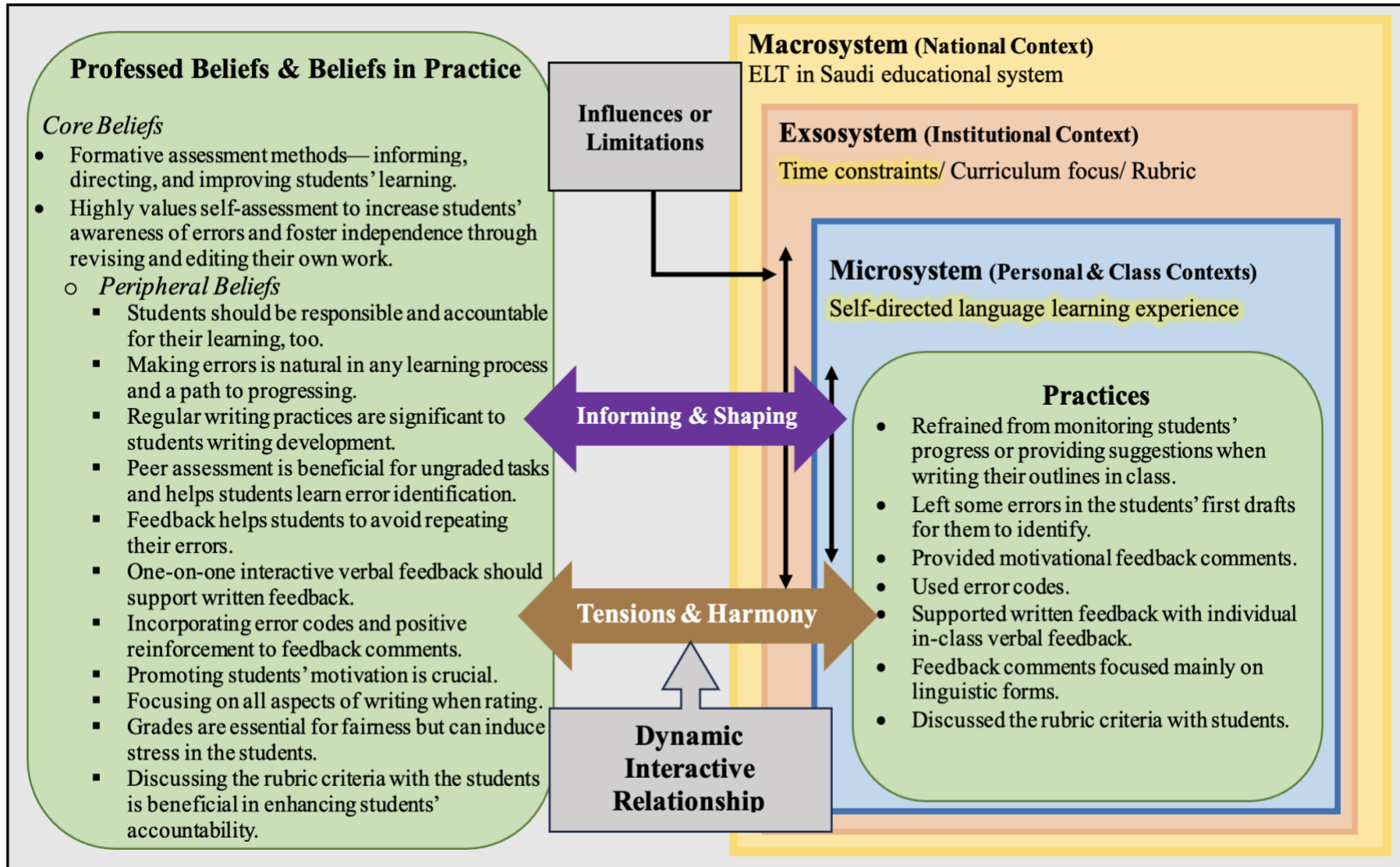
The interaction of curricular and time constraints, within the exosystem, as delineated in Figure 7.3, posed considerable challenges to the enactment of Batool's assessment beliefs. The syllabus' density and grammar-centric orientation curtailed opportunities for extensive direct writing practice and constrained the allotment of one-on-one verbal feedback sessions. Although not explicitly mentioned by Batool, the rubric was identified as another exosystem factor that possibly led Batool to believe she was focusing her feedback on all aspects of writing. However,

in her actual practices, she focused mainly on linguistic forms, mirroring the rubric's evaluation criteria of grammar and mechanics. The interactive relationship among Batool's systems, and the tensions arising therein, underscored the complexity of actualizing professed assessment beliefs within the practical constraints of the institutional setting. Despite these tensions, Batool's commitment to student improvement through the integration of written and verbal feedback remained evident. She adjusted and adapted her practice by allocating specific time for one-on-one verbal feedback sessions with each student, aligning these practices with her feedback beliefs within the given time constraints.

At the macrosystem level, Batool critiqued the disconnect between ELT practices in schools and higher education within the Saudi educational context, advocating for a more synergistic approach that tailors English instruction to the specific academic disciplines of students. However, this critique did not directly bear on her assessment beliefs or practices.

Figure 7.3

Summary of Batool's Case Findings



CHAPTER 8: THE CASE OF NOOR:

Assessing Meaning and Unity in Informal Writing

8.1 Background for Noor

Noor is an English lecturer who has been teaching at the University for four years. She holds a bachelor's degree in English language from KSA and a master's degree from the US, specializing in education, boasting a teaching experience of five years. Noor's educational journey commenced at another PYP in a public university, where she began her career as a teaching assistant. In this role, she taught foundational English language skills to both beginners and advanced students, covering essential skills such as reading, writing, listening, and grammar. She explained,

Excerpt 8.1:

I taught in {university name} for one year for beginners and advanced students. It's, like, levels. We had at {university name}, like, four levels, from one to four. The fourth level is for advanced students, and level one is for beginners. So, I taught both the beginners and the advanced level, but I'm not sure if I taught the intermediate levels. I taught the whole course, like writing, grammar, reading, and listening. It was integrated all in one book. (Noor, II)

In the University, Noor has been teaching fundamental English skills to undergraduate medical students as part of their PYP curriculum. Over her four years at the University, Noor has exclusively taught writing courses to a varied range of student levels, from the unified and nursing sections. At the time of data collection, she was teaching a writing class consisting of 23 students from the unified section. Noor engaged with her students for three hours every week, concentrating on enhancing their writing abilities within the framework of the English II course.

8.2 Teaching and Learning Writing: Informal Writing through the Integration of a Multimedia Learning Approach

Noor's beliefs concerning the teaching of writing centered on cultivating a relaxed and engaging atmosphere, fostering informal writing experiences for students across various genres, such as journaling, reflection, and commenting. She also appeared to believe in integrating

reading and listening skills with writing abilities, employing multimedia platforms like Teams, Blackboard, Instagram, and YouTube. In her own words:

Excerpt 8.2:

So, I wanna make *it* [[the writing course]] more casual. Maybe I'm gonna create something on Teams or Blackboard, but I'll give *them* [[students]], like, different topics or different articles, or something to read, like a small passage, and then they can comment and reflect on it, or, like, YouTube videos to comment on... The other thing is maybe I'm gonna, like, use/ also in a casual way/ for example, let's say, Instagram. So, I'll show them a post on Instagram and ask them to comment and reply to it. I'll try to choose some of the posts that are very interesting to comment on. So, I can, like, also assess their writing and how they can develop writing for daily life in an engaging, casual way. (Noor, II)

By offering students diverse topics, articles, reading passages, and videos, Noor intended to facilitate informal writing experiences. Her approach seemed to encourage students to engage in commenting and reflecting on a range of content. As a belief in practice, she mentioned employing Instagram posts as writing prompts, selecting posts that are captivating and interesting. Her perspective on teaching writing appeared to aim at personalizing the experience of learning to write in English, making it more relevant and applicable to students' everyday lives.

Furthermore, Noor advocated for an approach to teaching and learning writing that involves continuous free writing. She explained,

Excerpt 8.3:

I can use, like, other ways. For the first time, like, at the beginning of the course, let's say I'll give them a task about writing a journal, something to read or watch and write about it in a journal, or, like, write about anything they want. So, not daily, but on a weekly basis. (Noor, II)

Noor firmly believed in integrating journaling tasks throughout her writing classes, enabling students to express themselves and reflect on their experiences or the content they have encountered. This approach underscored her belief that the foundation of learning writing lies in incorporating a personal and expressive form, allowing students to document their thoughts, experiences, and emotions.

When discussing teaching academic writing genres, Noor highlighted that the process of learning to write commences by encouraging students to engage in consistent writing about

diverse topics. According to her, this approach serves as the foundation upon which students refine their academic writing skills. She responded,

Excerpt 8.4:

If students start by writing about anything, any topic, like, they read some stuff and comment on them. Writing about their feelings, ideas, opinions. I feel, like, they can write about anything later. But they have to start with all of this stuff. (Noor, II)

Additionally, she emphasized the significance of intertwining reading and writing within informal teaching methods. In her perspective, students must read and engage in personal reflections on reading materials, deeming it a fundamental skill in the journey of mastering academic writing. Notably, Noor's beliefs regarding teaching and learning writing intricately align with her beliefs on writing assessment, a connection that will be delved into further in the subsequent discussion.

8.3 Writing Assessment Purposes: Improving Students' Learning Outcomes

When asked about her beliefs on the purposes of writing assessments, Noor made a clear distinction between the purpose of writing assessments within the PYP framework and her personal viewpoint. Both these articulated purposes underscored formative objectives aimed at enhancing students' learning and fostering their future academic accomplishments.

In Excerpt 8.5 below, Noor indicated her belief that within the PYP writing courses, the primary purpose of writing assessments is to prepare students for the culmination of their academic journey – writing research papers in English. As per her understanding, students are expected to progress from composing paragraphs to crafting essays, a trajectory designed to improve their capabilities for undertaking comprehensive research papers at the end of their university years. She explained,

Excerpt 8.5:

To prepare them for writing research at the end. That is what I know that they have to do at {university name} and what they're doing. They ask the students to write paragraphs, then essays, and all of this stuff to prepare them for writing research. Doing a research paper at the end. (Pause) What I know is they, like, ask

the students to write research, but not in the English courses. No. But in other courses. I think there's a course about research at the end. (Noor, II)

When specifically asked about her own beliefs about the purpose of writing assessments,

Noor expressed,

Excerpt 8.6:

Assessment helps students to improve their English. They're gonna develop themselves, but honestly, till now, I don't believe in judging them on their grammar or spelling mistakes and all of this stuff. So, what I would focus on is coherence—the unity and all of this stuff, but not the spelling and the grammar. Whenever I look at the paragraph and understand it, that's all I need. So, it's more, like, focusing on the meaning and the unity. (Noor, II)

According to Noor, assessments are crucial in improving students' English language and enriching their learning. She also elaborated on her evaluation criteria, which are centered around coherence, unity, and meaning, a belief that will be detailed in the subsequent section.

8.4 Criteria for Evaluating Writing: Focusing on Meaning Over Mechanics

In the concluding statement of Excerpt 8.2, Noor emphasized her primary objective in assessing students' writing: evaluating “how they can develop writing for daily life in an engaging, casual way” (Noor, II). It generally appeared that she views the teaching and assessment of writing through an informal lens, prioritizing criteria that assess the ability to construct meaningful content. This belief is illuminated in Excerpt 8.6, where Noor expounded on her evaluation criteria. To her, understanding the paragraph's meaning is paramount; it is the sole criterion of significance. Noor voiced her reluctance to evaluate students based on grammar or spelling inaccuracies. Instead, her attention is steadfastly fixed on coherence, unity, and the holistic meaning encapsulated within the students' written responses.

When specifically addressing how she would assess students' writing, Noor distinguished between her personal beliefs and the requirements imposed by the PYP, highlighting the contrast between her ideals and the practicalities dictated by the curriculum. She underscored her

preference for evaluating students' writing based on criteria related to meaning, favoring a relaxed and informal approach over strict academic standards. She explained,

Excerpt 8.7:

I believe they should start with writing a journal regardless of the word count and all this stuff that we focus on in {university name} like spelling, grammar, organization/ but not less, for example, than half a page or something like that. So, I feel, like, it's not gonna be so formal or academic. It's gonna be like a casual way. And that way, I'm gonna assess their writing... So, I would focus on major mistakes. I'm not gonna take marks off for the grammar. But if it's a major mistake in grammar, like, affecting the meaning, I'm gonna, like, take it into consideration. And as I told you, like, the coherence and all of this stuff, but not gonna focus on the minor things or the spelling mistakes. So, I feel, like, they're not that important for me to focus on because most people make these mistakes {laughs}, so it's not gonna mess up their writing. (Noor, II)

Her response in Excerpt 8.7 again shows her belief that students should commence their writing journey with journal entries. Her belief in journal writing appeared to influence her evaluation criteria, emphasizing the significance of meaning and logical connections (i.e., coherence) rather than formalities like word count or specific mechanical aspects. Consequently, she criticized the excessive attention given to minor grammatical and mechanical errors within the PYP, arguing that these common mistakes should not overshadow the essence of meaningful expression in students' writing. Her rationale appeared to stem from the belief that minor mechanical issues are common among all learners and do not markedly diminish the overall quality of writing.

In other segments of the interviews, Noor emphasized the importance of coherence and cohesion as crucial criteria for evaluating and conveying meaning. For instance, in Excerpt 8.8, when describing the journal task that she desired to incorporate into her writing classes, she highlighted another specific area of evaluation: transitional signals.

Excerpt 8.8:

... let's say I'll give them a task about writing a journal, something to read or watch and write about in a journal, or like, write about anything they want. So, not daily, but on a weekly basis. And I'm gonna work with them on transition signals. So, they gonna focus on them. I'll give them a variety of these signals so they can use them. So, they would know that I'm gonna assess the transition signals... transition signals create unity and make the meaning better (Noor, II)

Furthermore, in her first follow-up interview, when probed about the significant errors she had specifically considered during her first TAP session and the assessment of the students' compare

and contrast paragraphs, Noor clarified, “I see everything. Also, like, transition signals are very important, especially in the compare and contrast paragraph” (Noor, FI 1).

On the other hand, Noor reported that within the PYP, she assessed students’ writing according to the rubric (see Table 3.3 for a copy of the PYP rubric). In Excerpt 8.9, she detailed her evaluation process, focusing on criteria such as spelling, grammar, sentence structure, paragraph organization, coherence, and adherence to the task instructions. She explained,

Excerpt 8.9:

I look for the spelling, grammar, sentence structure itself, and the organization of the paragraph. If there’s something that breaks the unity or something, I can, like, focus on that. Also, the word count. So, I follow the rubric. So, we should specify and give feedback on all these things. And if the students, like, follow the writing task prompt itself. Also, if they follow the instructions and all of this stuff, we gonna give them a grade on it. (Noor, FI 1)

It could be inferred that Noor’s writing teaching beliefs, which embrace a real-life and multimedia approach, profoundly influenced her standards for evaluating students’ writing. Noor believed in an informal teaching style, incorporating diverse activities like reflecting on readings or videos, journaling, and engaging in online discussions. It appears that she favored a contemporary and learner-centered approach that prioritizes engagement and meaning. Consequently, her evaluation criteria beliefs also emphasized meaning and the development of students’ foundational writing skills for everyday situations, all within a lively and relaxed educational atmosphere. Therefore, scrutinizing mechanics in students’ journals or Instagram comments did not make sense to her. What held paramount importance to her was whether the meaning was effectively conveyed or not. When specifically asked about her reluctance to focus on mechanical errors, Noor’s response was resolute: “Because I got the meaning already” (Noor, II).

Another piece of evidence comes from Noor’s second TAP session. While correcting a student’s paper, she encountered a grammatical error in the sentence “After fasting a long day, people gathers on one meal for the iftar.” She remarked, “I hate this, this is boring” (Noor, TAP

2). When asked to elaborate on this comment, Noor laughed and said, “{laughs} this is just boring stuff. All of this stuff, the topic. I don’t know, the way we correct. I feel students need to write in a different way” (Noor, FI 2). Noor appeared to be torn between her beliefs about how writing should be taught and assessed, and the PYP mandated curriculum. Although she reported she adhered to the rubric criteria, she expressed a desire to transform the entire dynamic of the course.

8.5 Writing Assessment Methods: Emphasis on Formative Methods

One of the elements of Noor’s writing assessment belief system encompassed beliefs about the methods used to assess students’ writing. Primarily, Noor believed in the high value of using writing tasks that integrate reading and listening, such as journaling and reflection, as shown in Excerpts 8.2 and 8.3.

Moreover, Noor advocated for the use of direct formative assessment methods in evaluating students’ writing, emphasizing the importance of continuous engagement in writing tasks. She emphasized,

Excerpt 8.10:

I think we have to give them, like, writing tasks, one after another. Many writing tasks where they write a lot. So, they could improve their writing. Give them feedback on all of this stuff. But I think giving them only one exam or two without writing. It’s not gonna improve their skill. They cannot/ It’s gonna be challenging for them. Because we will not be able to give them enough feedback or comments about their mistakes.
(Noor, II)

Noor appeared to emphasize the importance of offering students multiple direct writing assignments, allowing them to enhance their writing skills through consistent practice. She also seemed to advocate against relying solely on one or two exams without regular writing practice, deeming such an approach ineffective. According to her, this method would lack the provision of adequate feedback and impede students’ skill development. Her belief in formative methods,

such as practice and feedback, for developing writing skills appeared to align with her purpose in writing assessment: enhancing students' learning.

Another element of Noor's beliefs about assessment methods was written and verbal feedback, which she deemed as "the most important part" (Noor, II). While detailed discussions about her feedback beliefs are reserved for the subsequent section, Noor clarified her approach to in-class verbal feedback during the interview, explaining that it is the job of the students. She explained,

Excerpt 8.11:

I also give verbal feedback, but I work on that as a group during class time. Like, I show the students, for example, three of the written papers as a sample for them to see. And I show them, like, the papers, so they can correct them, and they can give the feedback. I don't give them the feedback. They gonna give it to me in a verbal way. So, they take my role, the teacher's role. So, it's more like I make them think about themselves, like self-assessment to some extent, but in groups. (Pause). So, maybe peer assessment and self-assessment. (Noor, II)

In Excerpt 8.11, Noor stated her beliefs in practice of providing in-class verbal feedback to students. However, she involved the students actively in the process, promoting engagement and collaboration. By showcasing sample papers and encouraging students to critique and provide feedback, Noor appeared to integrate aspects of self-assessment and peer assessment as students reflect on their work and that of their peers.

8.6 Feedback and Grading

Noor perceived feedback as the most essential part of writing assessment as it helps students to improve and develop their writing skills. She clarified,

Excerpt 8.12:

I feel like the most important part is giving feedback to the students because I feel, like, when you highlight something, and you have a comment on it, I feel like the students get the idea of why it's wrong and how to improve it and develop it. طبعًا ((for sure)) how to improve it and develop it in her own way. I'm not gonna give her the answer for sure, but I'm gonna, like, guide her to the correct answer, the correct sentence. (Noor, II)

Noor attributed significant importance to providing indirect written feedback to students, a method where she points out errors without directly correcting them. She held the belief that

pinpointing specific issues and offering corrective feedback comments help students comprehend the rationale behind their errors, facilitating their writing improvement. She stressed the significance of guiding students toward the correct forms rather than furnishing ready-made corrections. During the first follow-up interview, when discussing her written feedback approach in the first TAP session, she reiterated her commitment to the indirect method. She elaborated,

Excerpt 8.13:

So, when I check the paragraph itself, as you can see, I have parts *here* [[on Teams]] to write comments. So, I just underline or highlight the word or the sentence that I feel like it's incorrect and add a comment on it... So, the student, when she, like, sees the error itself, the highlighted one. She can rewrite it and re-correct the sentence itself or the word itself. And she resubmits it as a final draft, or even if she wanna, like, make sure and ask me beforehand. She is welcomed. (Noor, FI 1)

Noor emphasized the value of indirect feedback, underlining that students could learn significantly by recognizing and correcting their errors independently. However, she also maintained the belief that providing a certain form of feedback was contingent upon the unique dynamics between individual students and their teacher. Noor clarified,

Excerpt 8.14:

I think it depends on the students and their relationship with their teacher. Like, some students come to my office asking for verbal feedback. They need more, want more. Some don't want that. So, they just see the written feedback. Like, sometimes, some students are weak. They need extra, so it depends on/ some students are more open and ask a lot of questions. So, I understand, and I tell them, "If you want more, come to my office." (Noor, II)

Excerpt 8.14 implies that Noor believed there is no one-size-fits-all approach or form to providing feedback. Instead, she advocated tailoring the feedback approaches to meet individual students' needs and preferences. Some students might prefer written feedback, while others might prefer verbal feedback. For instance, she emphasized the importance of giving direct feedback comments to weaker students at times, wherein she identifies errors and provides corrections.

Noor's approach to addressing the challenges low-level students face in understanding indirect feedback seemed to combine correction with guidance. She elaborated,

Excerpt 8.15:

Sometimes, I do *that a lot* [[providing the correct forms]] for the weak students. Sometimes, I feel like they will never get it in any way. So, I'm trying to help her, but at the same time, I give her, like, one example, like, correction of her mistake. So, I usually tell them, "You have another one or a different one so that you can correct the similar mistakes in the paragraph." Like, I just show her an example. So, she can get the idea herself. So, I correct one of those mistakes, as an example, and they can apply it to others. (Noor, II)

Noor's feedback method for weaker students involved presenting a specific example of a corrected error, guiding them to comprehend the error and its correction within the context of their writing. By employing this strategy, she seemed to believe in scaffolding students' understanding, enabling them to recognize similar errors and promote self-correction in subsequent work. As a result, her preference for the indirect written feedback method appeared to be more deeply ingrained and essential to her assessment belief system than her inclination toward direct feedback. Still, her willingness to adapt her feedback methods according to individual student needs occasionally seemed to lead her to combine both approaches.

Furthermore, Noor made a clear distinction between what she would prioritize if she had the autonomy to design her own writing class and what was mandated within the PYP framework. On the one hand, she emphasized that given the freedom, she would concentrate on "coherence and cohesion." She also emphasized that she would focus on the students' "development of the vocabulary itself. And the use of transition signals. Like, focusing on how to use them and to be more advanced one to make the meaning rich" (Noor, II). This emphasis on evaluation criteria that enhance the meaning and logical flow of written texts aligned seamlessly with her broader assessment belief system, where coherence and cohesion were pivotal aspects of her feedback focus.

On the other hand, in her actual feedback practices within the PYP, Noor reported a clear commitment to aligning her feedback with the rubric's grading criteria. She outlined this approach in her response:

Excerpt 8.16:

I focus right now on the things that they're gonna, like, be graded on. So, yeah, for example, the vocabulary itself, do they use, like, more advanced ones? I'm gonna focus on the things, like, focus on the coherence and content itself. And if they're following the instructions or not and on all of this stuff. So, I'm not focusing on the things that are out of the rubric. (Noor II)

Noor's explanation emphasizes her concentration on elements directly tied to the specified assessment standards in the PYP, such as evaluating the use of advanced vocabulary in line with the lexical resources criteria of the rubric, (see Table 3.3 for a copy of the rubric). Notably in Excerpt 8.16, she explicitly mentioned her decision to focus aspects like "coherence and content." However, as previously noted, Noor reported a belief in practice of focusing on the rubric criteria. Evidence from her TAP sessions revealed that she also concentrated on criteria related to mechanics and grammar, with remarks about "fragments," "spelling," and "comma rules" (Noor, TAP 1 & 2). For instance, in the following segment from a TAP session, her focus was primarily on mechanical and grammatical criteria:

Excerpt 8.17:

P: [[reads student topic]] "Ramadan mouth/ Ramadan mouth?"

P: {laughs} Mouth. Okay, spelling. I will write spelling.

P: [[continues reading]] "In this holy month, all Muslims all over the world fasting during the sunset hours." Fasting? [[types on student paper]] Verb form.

P: [[continues reading]] "Ramadan not only to fasting, (pause) but also to be closer to Allah."

P: [[types on student paper]] preposition. (Noor, TAP 2)

Grading Beliefs. Noor's perspective on grading reflected a preference for qualitative evaluation, favoring broader grade ranges such as A, A-plus, or B over specific numerical grades. She explained her stance as follows:

Excerpt 8.18:

I wish I could, like, give an A, A-plus, or like B, and this stuff, not like a grade by itself, like 91, 92, and all these specific numbers...because it is enough for the students and me, like, to know their range, instead of focusing on a specific grade for each number of mistakes... It also makes my job hard. Like, for example, I don't like to deduct marks. I'm not gonna deduct a mark for anything. Yeah. I'm trying my best way to give them, like, a push to give them an extra grade or something. But in a way that's, like, fair enough for the students. So, like, a range is better because I don't have to, like, think about giving a mark or not. (Noor, II)

Her belief seemed to emphasize grasping the overall performance of students rather than rigidly quantifying their errors. She also highlighted a possible reason for this belief, which was the challenge she faced in assigning precise numerical grades and her reluctance to penalize students

for their errors. By opting for a broader grading scale, she could sidestep the complexities of detailed mark deductions, allowing her to focus on acknowledging students' effort. It can also be argued that her emphasis on the meaning and unity of students' writing justified her grading belief of recognizing progress and achievements within a broader grading spectrum instead of reducing students' work to mere numerical values.

8.7 Ecological Factors Influencing Noor's Writing Assessment Beliefs and Practices

The findings from Noor's case study revealed various factors that contributed to the formation of her writing assessment beliefs and influenced the connection between her reported beliefs and actual practices. Firstly, her educational background, including her experiences as an EFL learner, graduate studies in the US, and participation in professional development workshops, seemed to have significantly shaped her perspectives on writing assessment methods and criteria. Additionally, challenges such as curricular constraints, limited teacher authority, and time constraints were identified as key agents causing discrepancies between her assessment beliefs and the practical implementation of her rating, pedagogical, and classroom practices.

8.7.1 The Influence of Noor's Learning Experience on her Writing Assessment Beliefs

One of the major findings in Noor's case was the significant influence of her language learning experience and professional development courses on her writing teaching and assessment beliefs and beliefs in practice.

Noor's journey learning EFL writing was marked by her immersion in free writing, a method she utilized for self-expression. She chose to document her thoughts on her phone, underscoring her need for a secure space to articulate her feelings. This preference highlighted the personal and meditative nature of her writing learning process. She explained,

Excerpt 8.19:

For writing, I used to write, like, not a journal but a free writing. In my notes and this stuff, whenever I feel like I wanna write something about my feelings and this stuff. So, I started writing in my notes. (Pause). But

not paper-based, but on my phone, so I like using notes and this stuff to write. I don't like the paper. I feel like everyone can see it {laughs}. So, I prefer to write on my phone. Type on my phone. Cause it's like secure for me {laughs} because I don't want anybody to know and see {laughs}. So, I still do it, and I kind of consider it, like, a kind of meditation. (Noor, II)

Regarding academic writing, Noor's exposure to this form of writing occurred during her bachelor's years at a university in Saudi Arabia. There, in an educational setting, she learned the intricacies of crafting essays and research papers.

Noor's narrative about her English writing learning journey highlighted her strong belief in the significance of personal initiative and individual comfort within the context of L2 writing education. This belief seemed to resonate in her teaching and assessment practices. Her consistent focus on informal assessment techniques, like journaling tasks and online comments, coupled with the creation of a supportive writing atmosphere where minor errors held little importance, suggested a clear connection between her learning experiences and her approach to assessing writing skills.

Moreover, it was inferred that her reliance on expressive writing using her phone and her emphasis on safeguarding her emotions signified her belief in the profound impact of emotions on the process of learning to write. This became evident when Noor was asked about her sentiments concerning teaching and assessing writing; her response was poignant and deeply felt:

Excerpt 8.20:

{laughs} To be honest, I hate, hate, writing {laughs} ... both as a teacher and a learner. Because I don't/ I like to write. Like, a free writing style is okay because I won't be judged by others. I mean, when I was a student, for sure. But now, as a teacher, I feel like/ I feel students feel that way. And I have to judge and rate them for sure, so I'll look for grammatical mistakes, the structure, and all of this stuff, you know? Coherence, unity, and all of this stuff. So, they feel, like, "We're not gonna write something perfect." (Noor, II)

In Noor's response, her candid and humorous tone implied her complex relationship with writing as a teacher and a learner. Her laughter suggested a discomfort, highlighting her personal struggle with being evaluated and critiqued as a student. Despite her aversion, she acknowledged the importance of a free writing style, indicating her understanding of the emotional vulnerability

students might feel from being academically assessed on every error. This emotional reflective stance likely influenced her writing teaching and assessment methods, emphasizing a supportive atmosphere where students can express themselves without fear of judgment. Consequently, it could be concluded that her own learning experiences have shaped her beliefs, focusing on informal, approachable ways of teaching and assessing writing, prioritizing content and meaning over fixating on every mechanical and grammatical error.

Regarding professional development and training experiences, Noor's response indicated a level of engagement in workshops related to learning and assessing writing skills. She stated,

Excerpt 8.21:

When I got my master's, we had like workshops and also, like, doing assessments for other students. So, I corrected papers and all of this stuff. I assessed students, but they were not mine. But at least I did these kinds of things. So, I have like some knowledge about how to correct... The workshop was about correcting papers, correcting them using error codes, and this stuff. And what we can consider as a mistake and what other mistakes you can, like, just leave them cause they're not gonna affect the writing itself... It was, like, about a weeklong. A week of a workshop. The other workshop that I attended was, like, about three days in a row. It was about writing an essay. Yeah, it was about writing an essay. How to write an essay. (Noor, II)

Noor reported that she participated in two workshops that enhanced her assessment practices.

The first, spanning a week, provided in-depth training in paper correction techniques, encompassing error codes and differentiating between mistakes and errors in writing. The second workshop, held over three days, specifically concentrated on improving essay writing skills.

In addition to the practical experience gained during her master's studies, Noor highlighted another significant influence on her beliefs concerning writing evaluation criteria. She elaborated,

Excerpt 8.22:

So, I used to, like, study writing in a formal way at {university name}. So, I feel when I got my master's, I found teachers correct students' papers according to the meaning and, as I said before, the coherence and all this stuff. So, I felt like, okay, we don't have, like, to work on the students' writing and focus much on the grammar or things. So, we can focus on the errors, not on the mistakes. We can't look at everything as an error. I feel like there is a difference between an error and a mistake. If there's something that doesn't affect the paragraph itself, it's a mistake. So, it's not an error. Why deduct marks? So that's why? (Noor, II)

Noor's response reveals a significant transformation in her approach to evaluating students' writing, shaped by her master's degree learning experience. She described how, during her bachelor's years, she had been exposed to a formal and traditional evaluation method, identifying every error. However, her perspective underwent a substantial shift when she studied in the US. There, she observed her teachers employing different assessment approaches, prioritizing meaning, coherence, and context over strict adherence to grammatical rules. These contrasting learning experiences served as an 'aha' moment for her, reshaping her perspective on the writing evaluation criteria she would prioritize as an EFL writing teacher.

According to Noor (see Excerpts 8.21 and 8.22), 'errors' impact the paragraph's overall structure and meaning, while 'mistakes' are minor issues that do not significantly affect the content. This distinction was shaped by knowledge obtained from attending the error correction workshop. Noor elaborated on the tangible impact of this learning experience in enhancing her efficiency in evaluating students' papers. This newfound skill not only streamlined the correction process but also reduced the time spent on rating. She explained,

Excerpt 8.23:

To be honest, the error correction, the coding that we used, I felt, like, kind of helped me to assess the students' papers. In the past, it took me a while to correct students' papers, but now I feel like it's not a piece of cake. But I feel like I'm not consuming all the time to correct one student's paper. So, it made me get used to it. (Noor, II)

Additionally, the workshop on essay writing had a transformative impact on Noor's ability to maintain focus and unity within a topic, as she discussed,

Excerpt 8.24:

The other thing, the one I took about the essays, I learned how not to be out of topic, so not breaking the unity and this stuff. So, I feel sometimes, even if I know how to write English and I know how to write specifically about a topic, I feel sometimes, I get out of topic somehow. So, now, I know how to explain to the students not to be off-topic. I remember that workshop and my experience... So, that was really beneficial, really. (Noor, II)

Learning how to prevent drifting off-topic provided valuable insights, enabling her to guide her students effectively. Noor's ability to incorporate these lessons into her own writing and

subsequently into her teaching practices demonstrated the practical and profound impact of her professional development experience. Consequently, when addressing the significant challenges her students face in writing, she noted, “I feel like the complicated thing for the student is how to be on topic. Cause they seem like going out of topic most of the time” (Noor, FI 2). It appeared that her learning experience guided her to notice this issue in students’ writing and perhaps led her to focus on helping them overcome it.

8.7.2 Tensions Between Noor’s Writing Assessment Beliefs and Practices

Throughout Noor’s discussion of her beliefs and practices, she consistently drew a distinction between her personal beliefs and the practical constraints imposed by the PYP in terms of writing assessment. Notably, she stressed her preference for evaluating writing based on its substantive content, prioritizing meaning and coherence over minor mechanical errors. This approach conflicted with the PYP’s insistence on strict adherence to all predetermined rubric criteria. Another notable disparity emerged in her perception of the purposes of writing assessment. While Noor viewed assessment as a means “to improve their English,” she contrasted this with the PYP’s perspective, which she perceived as aiming “to prepare them for writing research at the end” (Noor, II, Excerpts 8.5 & 8.6). This consistent differentiation underscored her intentional effort to distance herself from the PYP context. Consequently, this distancing might reflect the underlying tensions between her personal beliefs regarding writing assessment and the practices mandated by the PYP, a conflict primarily influenced by curricular constraints, limited teacher authority, and time constraints.

In one response, Noor addressed both curricular constraints and the limited autonomy teachers have in shaping the curriculum. She stated:

Excerpt 8.25:

Honestly, we’re not choosing the rubric. We’re not choosing the topics. We’re not choosing whatever, even the writing style. If I wanna, like, ask the student to do something, I have to follow the instructions, word by

word. Even if they don't suit me, oh, I have to follow them. So, you know, they are all done deal, ((enough))
 خلاص like even the task instructions. I just do it. (Noor, II)

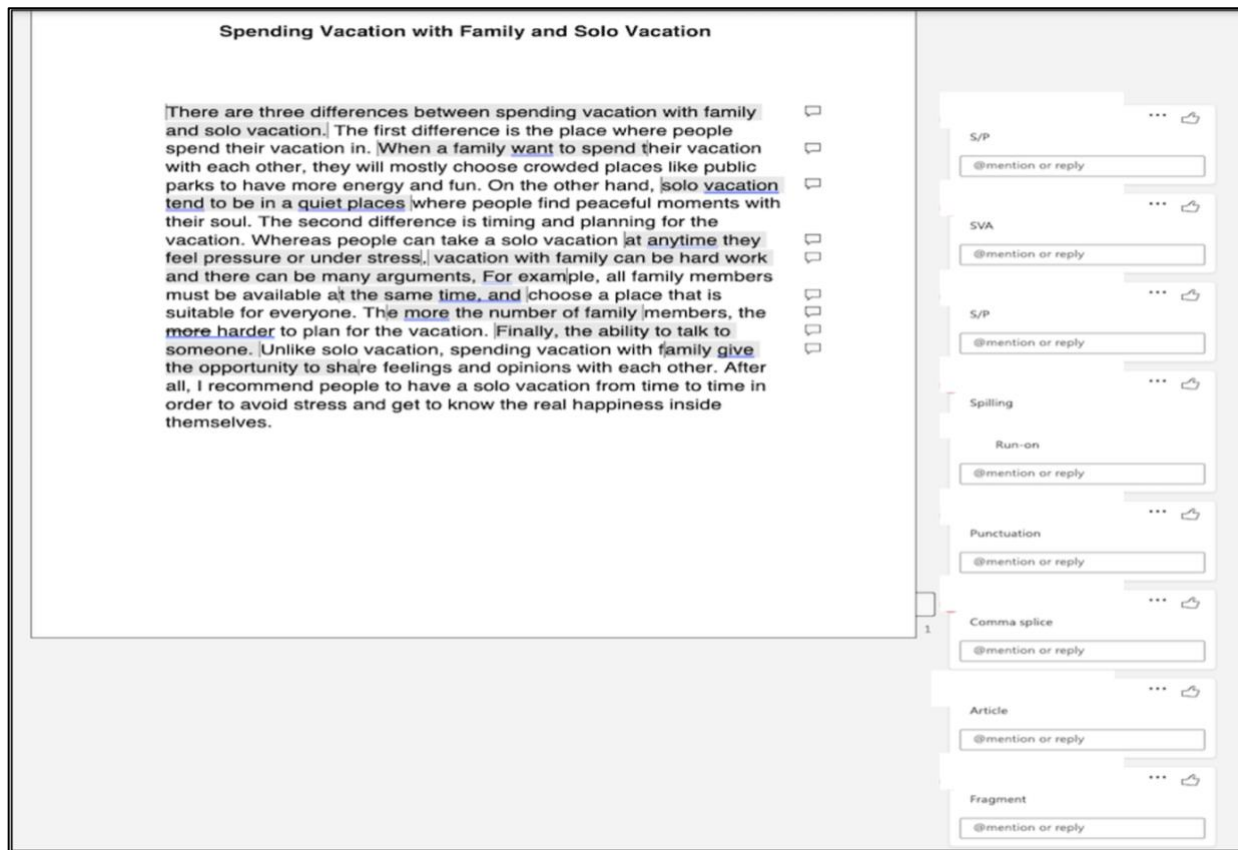
In Excerpt 8.25, she emphasized the lack of autonomy in selecting rubrics, topics, or even writing styles. Noor expressed her frustration, highlighting that teachers were bound to follow instructions meticulously, even if they did not align with their teaching and assessment beliefs. She underlined the inflexibility of the system, noting that all instructions, including task guidelines, are predetermined and non-negotiable, leaving teachers with little room for creativity.

These limitations noticeably influenced the alignment between her assessment beliefs and her assessment practices. Noor held core beliefs that writing is to be taught in a relaxed, informal manner, with the goal of encouraging students to explore free writing genres like journaling, reflecting on reading materials, and commenting on digital platforms (Noor, II, Excerpts 8.2 & 8.3). These pedagogical beliefs aligned with her core assessment beliefs, emphasizing a focus on major errors that can impact meaning over minor mechanical ones or word count (Noor, II, Excerpt 8.7). However, a significant misalignment became evident when Noor discussed her actual assessment practices in the PYP. She candidly admitted, "I'm not focusing on the things that are out of the rubric" (Noor, II, Excerpt 8.18). She must assess students based on all criteria outlined in the rubric, including mechanics, grammar, word count, organization, and content. This disconnect between her beliefs and the constraints imposed by the standardized assessment requirements was palpable in her actual rating and class practices.

Firstly, Noor's written feedback comments were specifically centered around distinct sections of the rubric, particularly emphasizing grammatical and mechanical aspects, as demonstrated in Figure 8.1.

Figure 8.1

Example of Noor's Written Feedback in the First Task



In Figure 8.1, Noor directed her nine feedback comments exclusively toward errors falling under a specific rubric category, namely Grammar and Mechanics. These comments addressed issues like spelling, subject-verb agreement, run-on sentences, punctuation, comma splices, articles, and fragments. Across all students' papers, Noor's primary focus remained on ensuring grammatical and mechanical accuracy. However, in some instances, additional comments touched upon aspects such as word count, with Noor advising a student with a paragraph over 250 words to "make it shorter, please." Another comment delved into formatting concerns when a student presented each sentence on a new line, prompting Noor to remark, "Please, all should be in one paragraph." Interestingly, remarks related to the rubric's content

criteria (unity, coherence, logic, and quality of ideas) were notably scarce. Across the twenty paragraphs from the two tasks, not a single comment addressed coherence or ideas, for instance. The absence of such comments could be attributed to Noor's earlier focus during in-class outlining sessions on evaluating students' quality of ideas, a topic that will be further explored in the subsequent section.

Furthermore, the classroom observation provided clear indications of Noor's emphasis on meticulously explaining the instructions and procedures involved in the writing tasks, despite her belief in avoiding a rigidly fixed approach to writing. In the observed session, she explicitly stated to her students:

Excerpt 8.26:

T: Girls, you have to follow the [[Task]] instructions. We will do the class brainstorming and the outline today. The first and final draft at home. Remember you must have three main ideas and conclusion, and for sure, a topic sentence. (Noor, OBS)

In her classroom, Noor consistently ensured that students followed the guidelines outlined in the task instructions. She explained the feedback procedures and the timing of its delivery, as mostly specified in the guidelines. Additionally, she explained the rubric to her students, a practice emphasized by the course coordinator. Her classroom approach highlighted her dedication to maintaining clarity and consistency in the assessment process, aligning her teaching methods with the prescribed guidelines and expectations set by the PYP.

Another ecological factor that Noor discussed was time constraints. The overwhelming influence of limited time was palpable as she portrayed several times its influence on her teaching and assessment practices. She first discussed its influence on her pedagogical approach. She explained,

Excerpt 8.27:

So, I use, like, the textbook, and that's all. And even if I wanna, like, do an activity, there is no time to do any extra activities. We have a shortage of time, and, like, we have to push them to do the tasks, virtual tasks, all of this stuff. So, I can't, like, use any activity that I learned before to apply in my class. (Noor, II)

Her frustration was evident as she explained the limitations of her pedagogical approach, primarily relying on the textbook with no room for additional activities. She also elaborated on the influence of time on her rating practices:

Excerpt 8.28:

The challenges for me, I think, I feel like, just time. That's all. There's nothing challenging me. But, like, the *deadlines* also [[for providing feedback for students' first drafts]] and all of this stuff. It takes me a while to correct students' papers. So, I feel like it's time-consuming. So, like, time, deadlines/ the only things that I feel like challenging me...Because I feel I correct so fast just to finish. (Noor, II)

In Excerpt 8.28, she expressed deep frustration and constraint, underscoring time as her foremost challenge. The imposition of deadlines, particularly concerning the time allocated for providing feedback to students' first drafts, seemed to intensify her struggle. This time pressure significantly restricted her ability to evaluate students' papers thoroughly and properly.

Lastly, Noor voiced a genuine concern about the limited time allotted for students to engage with their writing tasks and practice at home through additional assignments. She expressed,

Excerpt 8.29:

والله ((swear by God)) I wish we have, like, plenty of time to write and work on their paragraphs. Like, even if we don't have enough time because we are focusing on following the schedule and the course and the coursebook, I'm okay with that. But I feel like if we give them more time to work on the tasks, or like more writing practice, it's gonna be like good for them. Cause they will have plenty of time to practice. (Noor, II)

Despite the constraints imposed by the prescribed schedule and coursebook, Noor underscored the significance of granting students sufficient time to enhance their writing abilities through dedicated writing task completion. This emphasis might suggest that Noor believed the allocated time for the writing paragraph tasks was inadequate for both comprehensive student evaluation and effective learning.

8.7.3 Harmony between Noor's Writing Assessment Beliefs and Practices

It appeared that translating Noor's fundamental writing assessment beliefs into her practice necessitates significant modifications to the entire writing course curriculum and

substantial changes to the assessment policies within the PYP context. Her core beliefs about teaching writing in a “casual” manner would prove challenging to implement within a PYP framework geared toward preparing students for academic studies. Hence, it could be inferred that Noor’s beliefs seemed better suited for contexts where writing is taught in less formal settings, such as instructing general writing skills in an English institute. Consequently, the alignment between Noor’s writing assessment beliefs and her practices was confined to assessment domains where she had more room to implement her own methods without strict PYP directives. These areas included written assessment feedback approaches and classroom assessment practices.

Firstly, Noor’s written feedback practices were in harmony with her belief in the effectiveness of indirect feedback to help students comprehend the underlying reasons for their errors and enhance their writing skills. Across the 20 student paragraphs she assessed as part of this study, Noor predominantly employed this indirect approach. She pinpointed the specific locations of the errors and provided comments elucidating the nature of the issue without directly correcting it, as illustrated in Figure 8.1. However, in cases where a student’s paragraph exhibited significant weaknesses, she did provide direct corrections, such as suggesting to “remove this word,” or advising to “use the plural form,” as demonstrated in Figure 8.2. The application of a direct feedback approach for weaker students was also consistent with her reported feedback beliefs in practice, when she mentioned, “I do that a lot for the weak students” (Noor, II, Excerpt 8.15).

Figure 8.2

Example of Noor’s Written Feedback in the Second Task

Happiness is a sensation of enjoyment, happiness, and non-fear. There are two sorts of short-term happiness: those that remain for a short period of time due to one of the reasons of happiness, such as taking a walk to one's favorite location, and those that vanish as soon as one returns home. Long-term happiness is happiness that lasts for a very long period because it has long-lasting causes. For instance, a girl may have years of happiness if she marries a young man she adores.

The causes of happiness differ from person to person, but there are some actions that will increase the degree of pleasure in accordance with the human nature in which he was damaged and hence the construction of happiness for him, including the following actions:

People are happier and more content when they travel and explore the nation and its stunning locations. Being outdoors and surrounded by beautiful scenery while engaging in some sports, such as walking or running, is quite satisfying. Also, a special sort of chocolate can make some people happy, demonstrating how eating certain foods can improve human happiness.

In conclusion, people might feel more confident and happier by doing favors for others and volunteering their time.

Every person has a unique character; what one person finds pleasant may disturb others, and vice versa. What irritates others may make him feel comfortable and delighted. Comfort and contentment are the cornerstones of this life. Everyone has the right to pursue their happiness in life.

Header

You mentioned 2 sorts of happiness not only the short term
@mention or reply

wrong usage of due to here
@mention or reply

Rewrite the sentence
@mention or reply

use plural form
@mention or reply

replace it
@mention or reply

It could be concluded that the alignment between Noor's beliefs concerning feedback approaches (both indirect and direct) and her practical application was evident, reflecting the flexibility she had in implementing these methods. In contrast, her beliefs about employing a variety of letter grades contradicted the PYP's established numerical grading policies and procedures.

Furthermore, some of Noor's classroom practices aligned with her beliefs about writing assessment, particularly when she instructed students to concentrate on the quality of their ideas during the outlining. Noor's observed class, lasting one hour and fifty minutes, was divided into two main segments. The initial part involved discussing task instructions, procedures, and the rubric. Subsequently, she allocated 50 minutes for students to brainstorm and outline their first

paragraph task. While discussing the outlines with her students, Noor stressed the importance of meaning construction and logical idea development, emphasizing,

Excerpt 8.30:

T: Okay. In writing the outline or brainstorming, I don't care that much about the structure, the grammar, spelling. Just your ideas. I need your ideas to be clear and logical. I'm not gonna grade you. Whatever that you have, I don't think about the grammar, or the structure. Just the ideas that you wanna write. Okay? (Noor, OBS)

This emphasis on ideas, not grammar or structure, suggested that Noor found an opportunity within the PYP assessment framework to highlight the significance of constructing coherent and meaningful ideas in the writing tasks.

During the brainstorming and outlining sessions, Noor encouraged students to seek clarification by inviting questions. Notably, she refrained from circulating around to provide individual assistance, a choice possibly rooted in her belief that verbal feedback should be tailored to the student's needs and their rapport with their teacher, as expressed in her statement, "depends on the students and their relationship with their teacher" (Noor, II, Excerpt 8.14). In her responses to the students' questions, she reinforced the notion of logical idea division. For example, when a student sought her feedback on the three main points in her paragraph, Noor advised her to avoid repetition, highlighting the importance of clarity in ideas presentation, explaining,

Excerpt 8.31:

L: [[student came to the teacher and asked]] Miss, can you see my points?

T: Yes. [[reads the outline silently]]. Okay, you have two repeated ideas, like, look here you talked about that/ "eating in a restaurant is expensive," and then you said, "many cheap restaurants." يعني لازم ما تكون الفكرتين متناقضة ((the two ideas must not contradict each other)) (Noor, OBS)

Moreover, Noor advocated for employing verbal corrective feedback as a collective endeavor within the class, where students collectively participate in an editing session, assuming roles akin to that of the teacher. She elucidated, "I also give verbal feedback, but I work on that as a group during class time...I don't give them the feedback. They gonna give it to me in a

verbal way” (Noor, II, Excerpt 8.11). Her actual assessment practices in the classroom aligned with this belief. She informed her students that their outlines would be presented on the board in the next class, encouraging them to engage in a group editing exercise:

Excerpt 8.32:

T: So, I'll take the outlines from you and upload them on Teams, and next class, we will do the group editing.

L: Miss, can I upload my outline?

T: Yes, you can. But listen, try to make it clear so we can see it. (Noor, OBS)

This emphasis on content, evident during the students' outline activities and potentially in subsequent group editing sessions as referred to by Noor, might explain why her written feedback comments on students' paragraphs did not specifically focus on content. It could be concluded that beliefs related to feedback approaches (i.e., indirect, or direct) aligned with Noor's actual practices, as she has the freedom to do so, unlike beliefs regarding grading the students using a range of letters, which contradict the PYP numerical grading policies and procedures.

8.8 Summary of Noor's Case Findings

Figure 8.3 encapsulates the essence of Noor's case, delineating the integral elements of and interplays within her belief system regarding writing assessment, her actual practices, the various ecological systems exerting influence, and their interrelationships. The left side of Figure 8.3 shows that Noor's belief system was centered around her core belief that teaching should focus on informal writing and integrating reading and listening through multimedia channels such as Instagram and YouTube. She believed that the genesis of academic writing is embedded in the mastery of informal writing genres, including journaling and interactive engagements with digital content. This core belief profoundly informed her assessment beliefs and beliefs in practice. Noor advocated for writing assessments to enhance students' writing competencies for everyday use, urging for assessments to be immersive, prioritizing meaning and structural

coherence over grammatical precision. Her assessment beliefs extended to her feedback approach, favoring indirect feedback to prompt students' self-correction and grading strategies that evaluate the conveyance of meaning rather than quantifying errors.

Figure 8.3 further illustrates that macro-level factors appeared to exert minimal direct influence on Noor's beliefs and practices. However, the purple bidirectional arrow in the Figure indicates how her microsystem, especially her educational background, profoundly influenced her beliefs and beliefs in practice—stemming from her time as an EFL learner, graduate education in the US, and professional development on error correction. Firstly, Noor's journey of self-learning writing was characterized by her immersion in free writing and self-expression. This shaped her preference for teaching academic writing, beginning with informal genres such as journaling. Secondly, her graduate studies in the US shifted her assessment emphasis to a more holistic and meaning-focused approach to writing assessment, influenced by native English teachers who valued this over strict grammatical accuracy. Lastly, an error correction workshop refined her understanding of 'errors' versus 'mistakes' and the application of error codes, improving her efficiency in evaluating students' papers.

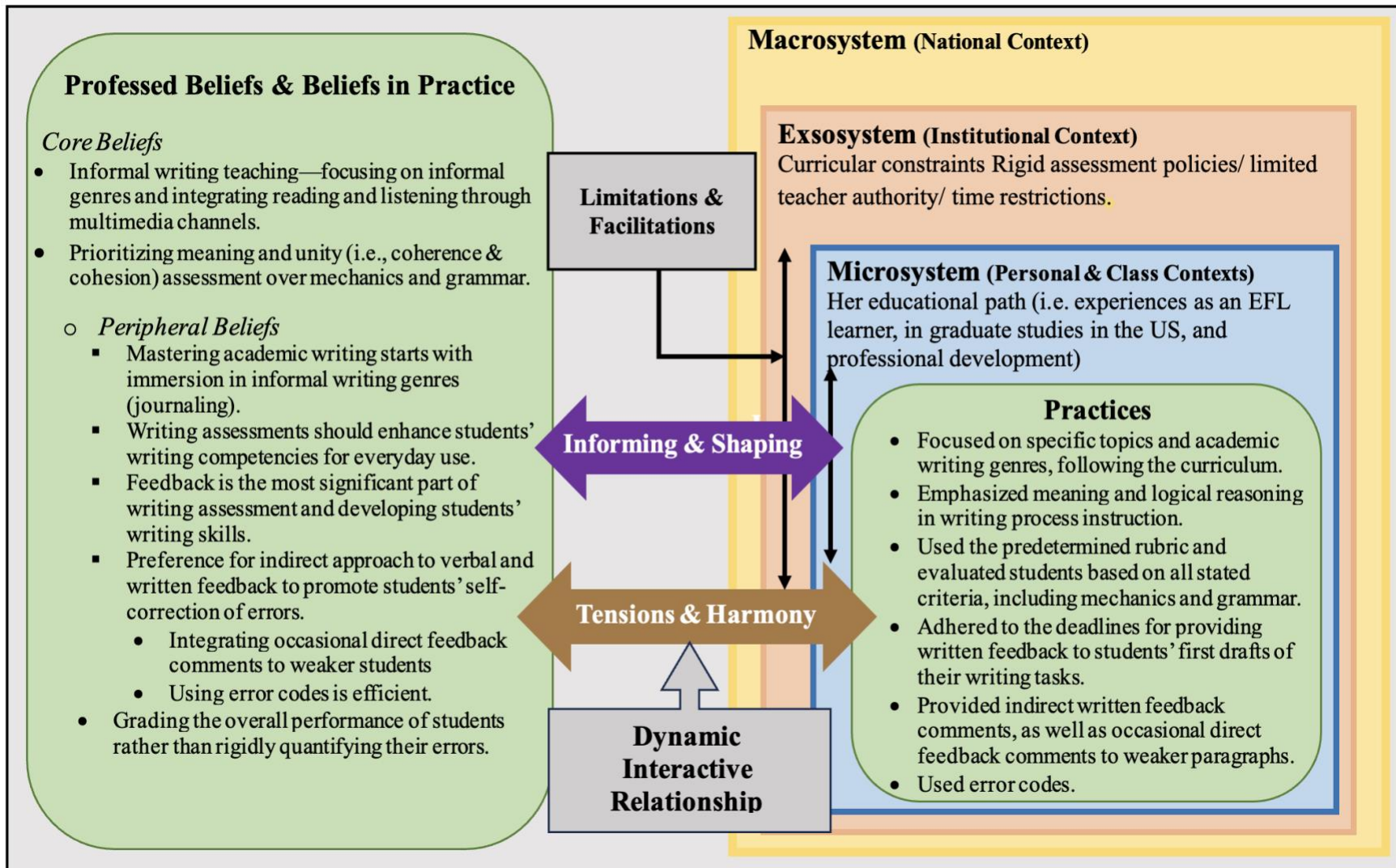
Exosystem factors such as curricular constraints, rigid assessment policies, limited autonomy, and time restrictions caused a misalignment between Noor's assessment beliefs and her actual practices in the PYP. These constraints led to her focus on specific topics and academic writing genres, reliance on detailed rubrics for mechanics and grammar, and adherence to strict deadlines, diverging from her preferred emphasis on engaging in informal writing practices and meaning-focused assessment.

The harmony between Noor's beliefs and practices is depicted as occurring in the absence of external ecological pressures. Her belief in fostering students' ability to self-identify

errors through indirect feedback is mirrored in her actual feedback practices, which emphasize the type of error made, occasionally supplemented with error codes, without furnishing the correction. Noor's classroom practices also resonate with her belief in the significance of meaning and logic in student writing, advocating for an emphasis on the quality and coherence of ideas during the writing process. Moreover, her workshop learning experience on error correction facilitated her sporadic use of error codes in her grading practices.

Figure 8.3

Summary of Noor's Case Findings



CHAPTER 9: THE CASE OF RANA:

An Affective Approach to the Teaching and Assessment of Writing

9.1 Background for Rana

Rana, a dedicated English lecturer aged between 30 and 35, has been teaching in the University for four years. She had teaching experience over seven years and holds a bachelor's degree in the English language from KSA and a master's degree in TESOL from the US. Her teaching journey began immediately after her undergraduate studies in 2014, where she delved into teaching English across various skills and age groups in different private English institutes.

She has been an integral part of the English department teaching faculty at the University since 2016, with a brief hiatus for her master's studies, making her experience a total of four years. She explained her journey, stating,

Excerpt 9.1:

It's been, like, seven years. I graduated in 2014 and immediately started teaching English at English institutions. Then, I got accepted at {university name} and started working here from 2016 till 2017. Then, I did my master's at {American university name} and returned in 2020. So, it's, like, four years of experience, I guess, here at {university name}. (Rana, II)

During her teaching journey, Rana taught all four essential English language skills: reading, listening, writing, and speaking. Notably, she took pride in her diverse teaching experience, covering various age groups. She enthusiastically stated, "I taught all levels, from kindergarten to 50, 60, to 70-year-old female students. All in different English institutions. So, you know, yeah, I taught all levels. All ages. And I'm very proud of that" (Rana, II). In the PYP, Rana has taught reading, grammar, and writing for unified and nursing sections. She elaborated that she has primarily been teaching writing, a responsibility she was assigned for the past three years since her return from the US.

Unlike other participants, Rana was exposed to English from an early age. Rana reflected on her early exposure to the English language, highlighting a significant learning period in an

international elementary school. She emphasized the strength of her English foundation, stating, “I have a solid English background... learned all the basics from that international school” (Rana, II). According to her, this experience equipped her with the essential skills of reading and speaking fluently. However, after four years, her education trajectory changed when she transitioned to a public school, leading to a temporary pause in her English learning for approximately two years. Rana resumed her English education in middle school, continuing her language journey until her graduation in 2014. Rana pointed out that her writing abilities were significantly shaped during high school. She expressed, “I feel like high school defined my writing skills because we had to write, you know, about ourselves, a paragraph about ourselves” (Rana, II).

At the time of the study, Rana was teaching a writing class consisting of 24 students from the nursing section. She engaged with her students for three hours every week, concentrating on enhancing their writing abilities within the framework of the English II course.

9.2 Teaching and Learning Writing: An Interactive and Student-centered Class

Rana’s beliefs about teaching and learning writing were highlighted when she discussed her ideal L2 writing class, which reflected her interactive and student-centered approach to teaching writing. She explained,

Excerpt 9.2:

I would ask the teachers to be super active with their students during the class and avoid the traditional way. Because it will kill the students... I mean by traditional way, like, when the teacher only teaches the content, and the students do nothing. So, I would encourage a lot of practice during the class instead of, you know, explaining a lot of content because *here* [[PYP]], we just explain a lot of content. And we don’t have time to do a practice. I don’t even give students a break sometimes... So, definitely avoid the traditional methods and focus on a lot of practice rather than just content. It could be teaching the content through practice... For example, the teacher can teach them about the logical division of ideas paragraph immediately after that; she can ask the students to do something with that information... I would ask them to write freely. Don’t, don’t limit their ideas to specific topics or a specific word count. Just say, “Okay, you’re learning about this paragraph. Pick the topic you want to talk about, and let me see you talk about it.” And the class has to be a very, very, very interactive class. Focuses on practice, different practices, rather than content. (Rana II)

In Excerpt 9.2, Rana emphasized the crucial role of writing teachers in actively engaging students, urging a departure from teacher-centered traditional methods that might hinder students' creativity and involvement. She seemed to advocate for a dynamic and interactive classroom environment where practical writing engagement takes precedence over inundating students with excessive content. For Rana, interactive learning means empowering students to explore topics freely, unrestricted by specific themes or word limits. Rana's beliefs on teaching content through practice seemed to align with modern pedagogical approaches, emphasizing interaction, task-driven class, and active student participation.

Furthermore, Rana underscored the importance of granting students greater autonomy in managing their tasks in the PYP. She highlighted the significance of empowering students to determine the timing of their quizzes and assignments according to their schedules. She emphasized, "They have to change the timing of their tasks and quizzes. We have multiple quizzes on the same day. They should let students decide when they take their quizzes and tasks" (Rana, II).

Notably, in both Excerpts above, Rana's critique of the specific PYP framework at her university stands out prominently. Her dissatisfaction with the assessment policies and curriculum remained steadfast throughout the interviews, underscoring a consistent discontentment. Right from the beginning of the initial interview, Rana set the tone, emphasizing her commitment to candidness: "As I promised you, and as I told you, I'll talk freely, and I'll be very honest with you" (Rana, II). In subsequent sections, I will explore Rana's writing assessment beliefs, which have emerged mainly through her critique of the English II course within the PYP context, in more depth.

Beyond her critical stance, it is essential to highlight the emotional depth of Rana's accounts, particularly concerning her students and her role as an educator. For instance, Rana commented when answering a question about grading, stating,

Excerpt 9.3:

Listen, my students, زي مايقولون خط أحمر بالنسبة لي ((As they say, are a redline for me)), especially in writing. They are my priority. And when I teach writing, I get very emotional, and I take it personally with them. (Rana, II)

This sentimental depth in Rana's responses intricately shaped her interaction with the interview questions, weaving through every facet of her reported beliefs and beliefs in practice. Grasping this profound affective connection proved indispensable in understanding Rana's writing assessment belief system. Furthermore, it potentially illuminated the origins of her markedly critical perspective on the PYP context. This exploration of Rana's critical and emotional stance is a focal point in discussing this chapter's findings.

9.3 Writing Assessment Purposes: Formative and Summative

Rana's beliefs concerning writing assessments purposes seemed to encompass two distinct purposes. Firstly, she emphasized formative goals, focusing on providing feedback to enhance students' learning and progress. Secondly, her beliefs indicated an emphasis on AoL, involving the evaluation of students' learning outcomes based on the materials taught.

Rana's beliefs concerning writing assessment purposes centered on the idea that providing students with constructive feedback is vital. She appeared to view feedback not merely as an evaluative tool but as a catalyst for progress and learning. She explained,

Excerpt 9.4:

To be honest with you, I feel giving students feedback is the most important part of writing assessment. It's very beneficial and helpful for their progress and learning... When I give them feedback, they edit and revise their writing, and then, you know, they will learn from their mistakes. (Rana, II)

According to Excerpt 9.4, Rana believed that the core objective of writing assessments is rooted in the feedback she provides to students. She emphasized its role in engaging students in the

essential practices of editing and revising their work. Rana seemed to believe that through this iterative process, students gain insights into their errors, fostering a deeper understanding of their writing strengths and areas needing improvement.

Additionally, when I asked her about the influence of assessment on students' learning and performance, Rana focused on the emotional aspect of assessment. She emphasized that assessments should help students progress and learn without instilling fear of judgment. She responded,

Excerpt 9.5:

I do not want my students to suffer, and I always, always/ from day one/ and you will notice that/ I focus on their psychological side. I tell them, "Ladies, I appreciate your ideas. Don't be afraid. I'm here to help you. I'm here to give you feedback." I always, always give them a lot of positive, positive, you know, motivational talks because I want them to write. I don't want them to fear me or to fear the course, especially from day one. I really need to tell them/ because I think it's very important. I tell them, "Once you enter my class, imagine you're a white dove {Laughs}. No fears. No concerns. It's okay. Give me a wrong answer. I will accept it. You're learning English." I always tell them, "I'm not judging your English. Who am I to judge your English?" I tell them, "You know, just write your ideas. Okay? I'll give you my feedback. Stick to the structure. The structure is easy, and if you need any help, just contact me". Sometimes, I even, you know, voice message them explaining something about the course because some students feel too shy to ask questions during the course. I'm super flexible with them because I don't want to limit their ideas when they write. You know, we were students once, and we know how it feels. (Rana, II)

Although Rana's extended response in Excerpt 9.5 slightly deviates from stating a specific belief regarding the influence of assessment on student learning, it highlights her view of the potential negative emotional impact of writing assessments on students. Her emphasis on the emotional well-being of her students fundamentally shaped her assessment beliefs. Rana appeared to perceive writing assessment, particularly feedback, as an empathetic and nurturing task. By emphasizing the psychological aspect of assessment, she not only encouraged open communication but also created a safe space for her students to express themselves and learn without apprehension of evaluation. Rana's commitment to fostering a supportive learning environment was evident through her constant reassurance and positive reinforcement for her students. Furthermore, she reiterated her crucial role in providing feedback and assisting her

students, a responsibility that she seemed to take “personally” (Rana, II, Excerpt 9.3), and this aspect will be discussed further in the subsequent sections.

Her beliefs about AoL, which involved evaluating students’ learning based on what they had been taught, will be further explored in the following section.

9.4 Criteria for Evaluating Writing: Ideas and Students’ Psychology

Rana’s approach to evaluation criteria centered on the student’s grasp of the instructed writing course content. However, in other parts of the data, it seemed that she preferred evaluating the quality of students’ ideas more than other aspects of writing.

In Excerpt 9.6 below, Rana’s emphasis on assessing every writing or grammatical rule the students have learned reflected her belief and belief in practice in evaluating their understanding of the taught material. She responded to the follow-up interview question about her evaluation criteria for students’ written paragraphs, stating,

Excerpt 9.6:

Oh, on everything, because I’m focusing on every rule they’ve learned. You know, it’s their writing course. They have to follow a specific structure, specific rules. (Rana, FI)

Her response in Excerpt 9.6 underscores her commitment to evaluating students’ writing comprehensively based on the content of their writing course. By emphasizing the necessity for students to adhere to the course-specific requirements, she seemed to communicate a belief in maintaining academic rigor and adherence to established PYP guidelines. Excerpt 9.6 may also suggest that Rana did not perceive writing assessment purposes as solely formative; she viewed it as an AoL, evaluating students’ outcomes based on what she has taught. For instance, during her one TAP session, Rana conveyed her belief in focusing her evaluation criteria on course materials.

Excerpt 9.7:

P: [[Reads student paper]] “They can eat at any time and without any delay. Although it does take time for preparation. On the other/”

P: Fragment.

P: [[Types on student paper]] I've told them about this many, many times. She should be careful when using 'although' at the beginning. (Rana, TAP)

Excerpt 9.7 highlights her focus on assessing the grammatical rules she taught on in class, such as the potential fragments resulting from using subordinate conjunctions at the beginning of the sentence. In another part of her TAP, she verbalized such focus and emphasis, articulating, "I've explained this in class" (Rana, TAP).

Moreover, it can be deduced from Rana's discourse concerning the discrepancies among teachers in assigning grades to students' final writing exams that she valued the importance of evaluating the substance and quality of students' ideas. She explained,

Excerpt 9.8:

When I mark [[the final exam]], I can give a student a 100%. And the other teacher would say, "No, she doesn't deserve 100. She deserves 80". So, it's super unfair for the students. Super unfair. Honestly, I don't know. It's just because we're not judging their ideas. Again, we are not judging their ideas. (Rana, II)

In Excerpt 9.8, Rana elucidated a critical concern within the PYP grading procedure during final examinations, underscoring her discontent with the inconsistency observed in evaluating students' submissions. Her discourse on the omission of appraising students' ideas within the PYP framework emphasized her belief in the significance of assessing the substance of ideas rather than solely allocating grades grounded in predetermined criteria. Her frustration highlighted her profound aspiration to administer a fair assessment rooted in the merit of students' thoughts and creativity, an evaluation she believed that the instructor teaching them should conduct.

Remarkably, students' psychological preparedness emerged as a significant consideration in Rana's evaluation criteria. This observation was further elucidated in her elaboration on Excerpt 9.8, where Rana delved into the complexity of student performance dynamics during exams. She elaborated,

Excerpt 9.9:

I feel like I know that some of my students are excellent, but because of the exam time or because they are afraid, they don't do well in the end. Like, they had a midterm before, they had a quiz before, so the students couldn't concentrate enough. So, we should consider a lot of factors. We should think about the students, to be honest with you. And marking depends on the way you've trained your students. So that's why we have a lot of disagreements on the grades. (Rana, II)

In Excerpt 9.9, Rana emphasized her belief in the multifaceted nature of student performance, urging an awareness of the psychological factors influencing them during exams. Her belief underscored the importance of a sympathetic evaluation process, acknowledging external pressures, exam timing, and student anxieties as influential elements. Furthermore, Rana reiterated her belief in the importance of assessing students according to the content and methodology imparted by their instructor. Her discussion of grading and the PYP rubric also indicated her emphasis on evaluating students based on their teachers' judgment. This theme will be explored in more detail in a subsequent section.

9.5 Writing Assessment Methods: Direct and Indirect Methods

Rana discussed her assessment methods after I specifically asked her about them. Rana elaborated on what she believed were the best methods to assess students in the PYP framework (i.e., indirect methods) and what the best methods would be if she had her own class (i.e., direct methods). Additionally, she appeared to believe in the effectiveness of peer and self-assessment.

In the PYP English II curriculum context, Rana articulated the suitability of employing indirect assessment methods as the most effective way to evaluate students' comprehension of the writing course content. She elaborated on this perspective, emphasizing the utility of MCQ quizzes in assessing students' grasp of grammatical rules and paragraph structures, as the course lacked content quality judgment. She explained,

Excerpt 9.10:

Quizzes. Quizzes. They're very useful because everything that they learn here, they will be tested on... You know, like, multiple choice questions test them on what grammatical rules they've learned, what is the structure of a paragraph because, after all, we're not judging their ideas. Right? So, they have to follow a specific structure for writing a paragraph. It's required from them. It's their course to write a topic sentence. Then, three main points. Each main point has supporting details and then the conclusion. So basically, they

have a structure, a specific structure. And then they just fill this structure with their ideas. I don't care, to be honest with you. I told them, "I'm not judging your ideas. Write whatever you feel is good for your topic." Yet, I feel it's still, you know, limits their creativity. (Rana, II)

Rana seemed to support the assessment of creative expression. However, her belief in the inherent limitations posed by the structured approach of the course made her acknowledge the appropriateness of MCQ questions in assessing students' outcomes within the English II curriculum. As a result, her belief seemed to have shifted to evaluating adherence to the prescribed assessment format rather than the substance of students' ideas.

However, when I asked about her assessment methods she would use if she had her own writing class, Rana emphasized her preference for direct assessments of writing. She clarified,

Excerpt 9.11:

For *here* [[PYP context]], a lot of quizzes are fine, but for me, I would only ask them to write, write whatever they want...I mean, I would teach them about paragraph structures and everything, but I won't limit their ideas and creativity. I will ask them to write many paragraphs with different structures. (Rana, II)

In Excerpt 9.11, Rana diverted from the prevalent indirect-oriented assessment methods commonly used in the PYP context. Instead, she advocated for a more direct approach to assessment, granting students greater freedom in their writing. She envisioned a classroom that emphasizes teaching fundamental skills, such as diverse paragraph genres, while simultaneously nurturing unrestrained creative expression. While the precise reasons behind Rana's perception of the PYP writing paragraph tasks as constraining students' idea expression were not explicitly elucidated, it can be inferred that she considered these limitations to arise from limited writing topic choices and word count restrictions. This perspective is evident in Excerpt 9.2, where she emphasized, "Don't, don't limit their ideas to specific topics or a specific word count" (Rana, II, Excerpt 9.2).

An additional belief regarding assessment methods, specifically peer assessment, became apparent when Rana was asked explicitly about diverse methods, including peer and self-assessment. She promptly affirmed her belief in the effectiveness of self-assessment and peer

assessment, emphasizing, “I do use self-assessment and peer assessment, and they are very useful” (Rana, II). Further elaborating on her approach, she described a specific peer-assessment practice implemented in her classroom. In this exercise, Rana randomly paired students and instructed them to exchange paragraphs for assessment, emphasizing that the selected practice paragraph was not intended for grading but for peer evaluation. She emphasized her goal of developing students’ ability to “spot mistakes and edit them” in this process (Rana, II). Rana further connected the success of this peer-assessment practice based on student feedback, noting, “I received a lot of great feedback regarding this. They told me, ‘Miss, we learned a lot because we spotted our friends’ mistakes’” (Rana, II).

Furthermore, during the peer-assessment practices, Rana emphasized, “I told them, ‘Write a paragraph.’ I gave them, like, ten topics to give them more freedom to pick whatever they feel comfortable with. And I asked them to write a paragraph about it” (Rana, II). This statement underscores the earlier point that Rana believed in the association between students’ free expression and the number of topics provided for their writing tasks. She seemed to believe that providing the students with ten topics is better for their creativity and learning than providing four topics, as dictated by the English II curriculum.

9.6 Feedback: Critique of PYP Feedback Policies

Rana’s initial discussion of her feedback beliefs was connected to her critique of the existing PYP feedback policies. She seemed to assert that the current feedback practices, which involved written feedback given to students on their first drafts and subsequent final grading on their final drafts, were not sound. According to her, written feedback should ideally be offered concurrently with or following the assignment of final grades. She explained,

Excerpt 9.11:

I don’t like how we give the students feedback. I know feedback is very important. But what’s the point of, you know, standing for two hours, explaining every single rule to them on how to write a paragraph? After

that, I ask them to write and then give them my feedback, but I will not give them a final grade now. Why? You know, it's just a waste of time, and it's spoon-feeding. To be honest with you, I will just ask them to write, but I will not give them feedback. I will tell them, "You will receive your final grade, and maybe after your grade, I will give you feedback." You know? Because it's unfair to the teachers, it will teach them to rely on me, "Yeah. I don't care. Anyway. She will give me feedback." (Rana, II)

Rana questioned the effectiveness of providing written feedback to the students after spending extensive time teaching them the writing course content, deeming it a futile effort akin to "spoon-feeding." Rana highlighted her concern about fostering dependency among students, leading them to rely on her written feedback rather than developing independent skills. This belief hinted at Rana's emphasis on promoting self-reliance, especially since she argued that students will face challenges in the future if they become accustomed to depending on their teachers. She stated, "And they will face a lot of challenges because they will not be getting feedback. Maybe they will get feedback in the first two years. After that, what?" (Rana, II).

The second belief concerning feedback that Rana held was related to the modality of feedback delivery. Rana expressed a strong preference for personalized, one-on-one verbal feedback sessions over written feedback, believing in their greater effectiveness. She elucidated this preference by sharing her own tendencies, stating,

Excerpt 9.12:

You know, I am talkative. I like to express my feelings and thoughts by speaking.... I never, I never like to express by writing, even to my friends, you know, I send them voice messages {laughs}. So, let me give the students verbal feedback instead of written feedback over and over again. Let me decide (Rana, II).

Her preference for verbal communication initially appeared to be a personal choice. However, upon reporting her practices of verbal feedback, it became apparent that another factor contributing to her belief in its efficacy was the students' reflections on these verbal feedback sessions. This point will be discussed below.

Rana described a specific practice where she individually addressed students by their names, employing a pen to highlight their errors and engage in detailed discussions about them. Notably, Rana emphasized considering the students' psychological well-being, acknowledging

their tendency to feel shy or stressed during feedback sessions. To mitigate their anxiety and ensure they absorbed the feedback, she allowed students to audio-record these sessions. In her explanation, she emphasized,

Excerpt 9.13:

...sometimes students get really shy and stressed. I don't know, but they just freeze when I give them feedback. I know for sure they will forget most of it. So, I focus on the psychological part so much, as I told you. So, I tell them, "You know what, I want you to listen to me, come close to me. I will give you feedback. However, record it. Because I want you to focus on my feedback and not worry about forgetting it." (Rana, II)

Rana's beliefs about the focus of her written feedback demonstrated adherence to the PYP evaluation criteria, the taught material, and a personalized, motivational element. She asserted,

Excerpt 9.14:

I give them feedback on everything I taught. And to be honest with you, I also give them some motivational comments. And that's me, I guess. I write, "Oh my God, thank you. Your ideas are great. I love that idea. However, you should elaborate on this more." I am just doing it because I know that if I encourage them, they will be very happy. They will write better. And that's my goal. (Rana, FI)

In Excerpt 9.14, Rana demonstrated her commitment to evaluating students based on the entirety of her teachings, aligning with the PYP criteria and the taught material. Moreover, her inclusion of "motivational comments" indicated a deliberate effort to inspire and encourage her students. She appeared to endorse the belief that positive reinforcement promotes enhanced writing abilities, emphasizing her objective of nurturing skills through encouragement.

When asked about her approach to written feedback, she seemed to incorporate direct and metalinguistic approaches. She articulated,

Excerpt 9.15:

I use a lot of motivational comments, you know, they help... I highlight their errors and give them a comment about it. I write, like, "Please explain more." "You can't add this here. Remove it." I tell them/ like, (Pause) I sometimes explain why their sentence is wrong. I give them an explanation about it. (Rana, II).

Her TAP session highlighted her adherence to indirect, direct, and metalinguistic approaches to written feedback. For instance, in Excerpt 9.16 below, Rana was noted to identify errors, specifically articles and capitalization mistakes, and she provided not only the correct forms but also explanations of the grammatical rules or guidance for correcting these errors:

Excerpt 9.16:

P: [[Reads student paper]] “Eating at the restaurant consumes more time. For example,/”

P: “The restaurant.” Which restaurant?

P: [[Types on student paper]] “Which restaurant? Remember, when you use the article ‘the,’ this means you are referring to something specific. For example, ‘the writing class we took in school.’ So, for your second main point. It is better to write, ‘Eating at a restaurant consumes more time.’”

P: [[Continues reading student paper]] “For example,If someone wants to eat at restaurants/”

P: Capitalization.

P: [[Types on student paper]] We do not capitalize letters after commas. Also, please leave a space after the comma. (Rana, TAP)

Rana did not articulate the specific names of the feedback approaches she employed nor the reasons for her chosen methods. Nevertheless, her commitment to assisting her students was assumed to be a driving force behind her multifaceted feedback strategy. She seemed to provide corrections and explanations sometimes as a means of support and instruction. For example, Rana expressed her dedication to giving verbal feedback with such enthusiasm that it once resulted in her losing her voice for three days. She humorously remarked, “{laughs} I swear I couldn’t talk for three, three days straight because I give them feedback with all my heart {laughs}” (Rana, II). This connection between her beliefs and beliefs in practice will be further examined and substantiated in the context of her actual feedback practices later in the chapter.

9.7 Grading and Rubric Use: Teacher-centered Rating Approach

Rana shared her beliefs about grading in connection to her beliefs about rubrics. She strongly critiqued the existing grading system and the use of fixed rubrics in effectively evaluating students’ writing. She argued,

Excerpt 9.17:

To be honest with you, I don’t believe in this system of grading... يعني ((I mean)) a lot of unfair things happen...the rubric is not effective. Before you ask, I am telling you it’s not effective. To be honest with you, I would say it works 10% out of 100. I will be very specific. You cannot, you cannot use a fixed rubric to evaluate someone’s writing. You cannot. You cannot do that because it’s super unfair. A lot of students get a lot of marks out of it, you know? Because it’s like 3, 2, 1. Like, she will get the marks because it’s structured. Follow the structure, and that’s it. I would love to suggest to the coordinators that just let the teacher evaluate her students. Let her do it. We need to prioritize teacher’s judgment. For example, the first task is out of 10. let the teacher decide what the student’s grade is. That’s it. (Rana, II)

Rana vehemently voiced her skepticism about the effectiveness of using rubrics to evaluate students’ writing abilities. She argued that a fixed rubric system, with its rigid structure of

assigning marks based on predefined criteria, is inherently flawed and fundamentally unfair. Rana emphasized the limitations of this approach, highlighting how students often receive marks based on a simplistic scale that fails to capture the nuances of their writing skills. Instead, she advocated for a more flexible and teacher-centric evaluation approach, suggesting that teachers should be able to assess their students according to their judgment and not rigid policies.

In response to a question about what she would prefer to do for rating students' written paragraphs, Rana again articulated a strong preference for a personalized grading approach, dismissing the use of rubrics, stating, "Just the teacher grading without a rubric, because you know your students better than anyone. You know what you've taught them... You've been with them from day one, you know their levels" (Rana, II). Rana appeared to believe in the significance of a teacher's intimate understanding of their students, asserting that this deep knowledge allows for a fairer evaluation. Her stance reflected a belief in the value of subjective and impressionistic teacher judgment, emphasizing the importance of considering students' needs and individual development.

9.8 Ecological Factors Influencing Rana's Writing Assessment Beliefs and Practices

In examining Rana's data, a significant revelation was found regarding the profound impact of her personal history (i.e., teaching and learning experience) on her beliefs and beliefs in practice concerning writing instruction and assessment. The impact of both her teaching and learning experiences appeared to take an affective nature.

Additionally, as noted above, Rana's responses during the interviews are marked by a distinctive perspective. In her discussions, she persistently critiqued the PYP assessment policies, at times making it challenging to differentiate her personal beliefs from her critical evaluation of the PYP. This complexity highlighted the numerous factors influencing the

application of her beliefs about writing assessment into her actual practices. These ecological agents significantly shaped Rana's approach to writing assessments within the PYP framework. Noteworthy among these constraining elements were the rigid assessment policies (particularly regarding rating and feedback mechanisms), time constraints, levels of student proficiency, and, ultimately, the limited authority granted to teachers.

9.8.1 The Influence of Rana's Teaching and Learning Experience on her Writing Assessment Beliefs and Practices

When asked about her emotions regarding the teaching and assessment of writing, Rana candidly confessed her reluctance to teach writing, a sentiment she explicitly expressed. She stated, "I'm gonna be very honest with you...I do not like writing... I feel like, you know, it's not my cup of tea because I've always been great at teaching grammar" (Rana, II). She preferred teaching grammar, a subject where she felt confident and creative. According to Rana, writing instruction curtailed her creativity, leading her to approach it less enthusiastically than teaching grammar.

Rana clarified that her passion for teaching writing persisted due to her students' positive responses and reflections on her instructional methods. Rana cited four instances where students' favorable feedback bolstered her enthusiasm for teaching writing and heightened her sense of responsibility in exploring different assessment methods. For instance, she stated that students' positive responses validated the effectiveness of her assessment strategies, such as one-to-one verbal feedback and peer assessment, both of which she referred to as "experiments" (Rana, II). The success of these experiments, as indicated by the students' feedback, motivated her to incorporate them more extensively into her writing class.

Moreover, during Rana's description of her verbal feedback practice (Excerpt 9.14), she interrupted the explanation and emphasized her dedication to trying different assessment approaches, stating,

Excerpt 9.18:

Remember, I will do anything for my students. I mean it. I'm not trying to be perfect or that you look at me as an ideal teacher. No. I would do anything for them to make them better writers in English because it's a huge responsibility for me to teach writing. After all, it's me who will prepare them to write in the future. So, I feel, um, I don't know. I sometimes feel very anxious regarding this because I feel like it's a huge responsibility. (Rana, II)

Rana portrayed herself as an “emotional” teacher, deeply invested in her students (Rana, II, Excerpt 9.3). This emotional involvement appeared to underscore her constant need for affirmation regarding the efficacy of her teaching and assessment methods. Moreover, her strong sense of responsibility, stemming from her emotional engagement with her students, seemed to intensify, compelling her to approach teaching with a deep sense of obligation. Consequently, it could be inferred that her teaching experiences, particularly the positive interactions with her students, played a pivotal role in shaping her assessment decisions and choices of specific assessment methods and their implementations.

Rana's pedagogical approach displayed a distinct emotional dimension that appeared to stem from her personal learning experience. Her emotional attitude toward assessment seemed to intertwine with her learning experiences. This connection between her personal experiences with writing and her role as an assessor became evident in her statements, exemplified by, “I'm super flexible with them because I do not want to limit their ideas when they write...we were students once, and we know how it feels” (Rana, II, Excerpt 9.5). This expression underscores the impact of her prior learning background on her empathetic approach to teaching and evaluating her students' writing abilities.

Rana's focus on the emotional aspect of writing assessment seemed to result from a particular negative learning experience while taking a writing course in her undergrad. She

clarified, “When I teach writing, I get very emotional, and I take it personally with them. Because, you know, I had a bad experience as a student taking a writing course. So, I don’t want my students to suffer” (Rana, II). Rana elaborated and provided significant insight into her personal history, revealing an experience that profoundly impacted her attitude toward writing. The incident involved her cousin’s failure in a writing course at university, which created an intense fear within Rana. This fear, compounded by her cousin’s negative beliefs about the writing course, triggered Rana’s strong aversion to the subject. When Rana was in the position to take the same course exam, she could not write. She expressed, “I froze, and I cried. I remember I cried a lot, you know, till the teacher talked to me. She calmed me down... You have no idea how super scared I was? I was just so scared” (Rana, II).

This particular experience appeared to have enhanced her empathetic understanding of students undergoing similar challenges. For instance, it likely led her to critique the PYP practice of confining students to specific topics and word limits, a sentiment she expressed while recounting the incident, “When the teacher distributed the papers and asked me...and she said, ‘Write a paragraph about this topic.’ I couldn’t write. I couldn’t remember anything about the topic. It was something about obesity” (Rana, II).

This incident might have further instilled in her the belief in granting students the autonomy to determine the timing of their tasks and examinations (Rana, II, Excerpt 9.9). While implementing this practice might pose practical challenges in the PYP context, where exam schedules and task deadlines are coordinated among numerous students from various sections, Rana’s deep empathy for her students could have led her to overlook these difficulties.

Moreover, statements such as “I don’t want them to fear me or to fear the course” (Rana, II, Excerpt 9.5) indicated that she perceived assessment as a risk that could threaten the student

emotionally. She might tell herself if I felt this way during my writing exam, my students must feel the same. Her learning experience might have heightened this belief, reinforcing her commitment to alleviating any apprehensions her students might have about writing courses assessments.

9.8.2 Tensions Between Rana's Writing Assessment Beliefs and Practices

Throughout the interviews, Rana consistently criticized the PYP, faulting various aspects, from limited teacher autonomy to rigid assessment policies and curricular restrictions. These factors often caused a discrepancy between her assessment beliefs and her ability to execute them in practice; a tension echoed in previous sections. For instance, although Rana preferred teacher evaluation that emphasizes the quality of students' ideas without relying on a rubric (Rana, II, Excerpts 9.9 & 9.16), she was obliged to conform to the established rubric. Her actual practices were largely rubric-focused, as observed in her feedback and ratings. Figure 9.1 illustrates an instance of Rana directing her feedback comments to specific criteria on the rubric.

Figure 9.1

Example of Rana's Written Feedback in the First Task

<p>First Draft #1</p> <p style="text-align: center;">The Differences and Similarities Between Textbooks and E-books</p> <p>There are three similarities and differences between e-books and textbooks. First, there are a lot of similarities between E-books and textbooks. For example, they are similar in pages number, and all of them are written clearly and understandably, since I can take them to any where I want, also All sorts of books for adults and children which are related to cooking, are available in E-books as well as textbooks. Finally, they have a lot of differences between E-books and textbooks. For example, saving E-books on your phone or computer, but this is not safe for your books because when any problem occurs, all books may be deleted. Unlike E-books, textbooks will remain with you for a longer time, but they may get damaged if you do not preserve them. Also, E-books is cheaper and can be available faster than textbooks. Websites offer a free simplified version for purchasing E-books. Unlike E-books, textbooks expire faster and are not available until after a long period, but you can borrow them for a long time. There are similarities and differences between E-books and textbooks. In my opinion, I prefer E-books.</p> <p>Word count: 212</p>	<p>do not leave double space after periods. :Commented</p> <p>wrong use of the subordinating conjunction. please use a correct subordinating conjunction, or connecting word. :Commented</p> <p>Dear, this is a good supporting details. Please rewrite it more academically by paying attention to sentence structure, using more academic words. Just edit it dear. :Commented</p> <p>fragments :Commented</p> <p>sv- agreement :Commented</p> <p>dear, is this your conclusion? if yes, please add a transition signal to show your readers that this is the end of your paragraph as I explained to you in class regarding concluding your paragraph. :Commented</p>
--	--

In Figure 9.1, Rana’s feedback specifically addresses the rubric criteria, focusing on aspects such as mechanics and grammar (including punctuation and sentence structure), content (including the use of transition signals), and lexical resources (including formal register).

In the analysis of 20 samples of Rana’s assessment work, a consistent approach was evident in her application of the rubric’s diverse criteria. Her overall evaluative comments addressed each of the five rubric criteria comprehensively. With respect to format, her feedback included specific directives, such as, “Please pay attention to the format. This is not the same font type & size.” In terms of paragraph structure and organizational criteria, her comments were detailed, for instance:

Excerpt 9.19:

Dear, I know this is your second main point & it is pretty good. However, you need to refer to your book & what I taught you in class on how to write a main point and THEN its supporting details. You need to separate your main point from its supporting details by using transition signals, word connectors, and end punctuation to make it easier for your readers to understand your paragraph. For example, write it as this “Secondly, Internet connection problems can affect e-books”. THEN WRITE YOUR SUPPORTING DETAILS. (Rana, Feedback Comment)

Additionally, she provided feedback on content relevance, such as “This breaks the unity. Please relate your example more to your main point above.” She evaluated lexical resources with comments like “This is not very academic.” Furthermore, her feedback on mechanics and grammar was targeted, as seen in remarks such as “Dear, please place an end punctuation here. Consider this as your first main point,” and “Dear, you already have a main verb here, ‘organize,’ so delete the auxiliary verb ‘is.’” Consequently, Rana’s actual assessment practices appeared to align more with the stated evaluation criteria of the English II course than her assessment beliefs.

Additionally, Rana’s ability to apply her beliefs about feedback was constrained by the policies in place. Contrary to her belief, she did not support giving feedback before final grades

(Rana, II, Excerpt 9.11). Nonetheless, due to feedback policy requirements, she was compelled to provide students with written feedback on their first drafts before assigning a grade.

However, an additional factor that appeared to create tension between Rana's beliefs about assessment and her ability to enact them in her practices was time constraints. As mentioned earlier, Rana articulated her strong belief in the benefits of one-to-one verbal feedback but indicated that time restrictions prevented her from applying this to all student writing tasks. She stated, "Honestly, I don't think I will use this kind of verbal feedback again. I know it is helpful, but it took a lot of time ... I feel I can do it for the first task but not for all tasks. There is no time." (Rana, II). Notably, Rana linked the limited time available to other constraining factors such as feedback policies and timelines, student proficiency levels, and professional commitments, which significantly impede her capacity to implement individual verbal feedback.

Firstly, Rana's elaboration elucidated the burdensome nature of the assessment policies in the PYP, revealing a system that inadvertently promoted dependency rather than independence among students, a point also evident in Excerpt 9.11. She articulated a cycle of exhaustive marking and feedback that not only was demanding for her, with phrases like "why do I have to mark twice, twice, and give [[written]] feedback," but also believed in prompting a dependency among students who "know that we will give them feedback." This continuous loop of submitting, receiving, and re-submitting work for almost 25 students was described as "very exhausting" and "too much," emphasizing the disproportionate workload relative to the time allotted. She clarified, "You do this for each task, and for each student. No. At least for the first one. It would be fine" (Rana, II). Moreover, Rana perceived the number of tasks and the iterative nature of the written feedback process as not just a challenge for the teachers but overwhelming

for students, too, suggesting an inefficient system that strained all participants. Consequently, Rana's preference for allocating time to provide students with verbal rather than written feedback was not fulfilled due to the stringent assessment policies mandated for all writing instructors within the PYP context (Rana, II, Excerpt 9.12).

Secondly, Rana amply explained that the students' foundational understanding of English drastically varied, with some lacking even the basic grasp of 'is,' 'am,' and 'are,' which is pivotal before advancing to constructing paragraphs. Rana also clarified that the one-size-fits-all writing course approach is ineffective, yet the institutional expectations did not align with the pedagogical reality. She explained,

Excerpt 9.18:

They don't consider that she [[student]] doesn't know the basics. They don't consider the level of the course. They don't care about who to admit. They just don't care. They're just like, 'Okay, teach her writing.' And then they blame the English department for the student's level...Actually, when we tell them about the level of the students, they just say keep going, keep going. But I have to take my time if I'm gonna keep going. (Rana, II)

Rana highlighted a systemic issue: the disregard for the starting skill level of students, the underestimation of time necessary for proper instruction, and a curriculum that does not accommodate these disparities, thus limiting her ability to conduct meaningful writing assessments based on her choice.

Lastly, Rana's elaboration vividly portrayed the tension between her dual roles as an educator and an administrative quality officer, a responsibility she was assigned without seeking it. This additional administrative duty had significantly encroached on her time, limiting her ability to provide the level of feedback she valued in her assessment practices. She stated, "I feel I'm not as free as I was before," indicating a loss of autonomy over her schedule due to the demands of "a lot of paperwork, admin work." The exhaustive juggling between marking her students' paragraphs and fulfilling administrative tasks like "filling forms and sending them to the head of departments" made her feel "a little bit overwhelmed" (Rana, II). Rana's account

underscores how involuntary administrative duties could inadvertently create a two-fold workload, thereby possibly impinging upon the quality of the teaching and assessment she could have delivered.

Rana concluded her reflection on the constraints of her role with a candid acknowledgment of the core issue: her lack of autonomy. This perceived absence of authority emerged as the critical factor that compounded the misalignment between Rana's assessment beliefs and her actual practices. She explained,

Excerpt 9.19:

To be honest with you, I feel like, I don't know, it's just/ teaching writing courses at this university. This university is a bit challenging because I'm not free to do, you know, what I think is best for my students. I just follow the coordinator's orders. I can't do more. I can't do less. It's what it is. (Rana, II)

Her response in Excerpt 9.19 captures a sentiment of resignation, as Rana explained she was bound by rigid institutional policies, which left her no room to tailor her teaching to the needs of her students or to exercise her professional judgment in assessment. Her use of the phrase "It's what it is" further resigned her to accepting the status quo, however misaligned it may have been with her pedagogical ideals. The ultimate limiting factor, then, was not merely the administrative workload or the feedback policies but the overarching limitation on her professional role, which ultimately dictated the extent to which she could enact her beliefs in her assessment practices within the English II course at the University. For instance, Rana seemingly endorsed using direct assessment methods, perceiving them as the most fitting for evaluating writing skills. Nevertheless, she preferred not to constrain students with a limited range of topics or a strict word count, as required by the English II course requirements. However, the rigidity of the curriculum and the limited authority granted to the teachers prevented her from offering additional topics in the graded writing tasks. Consequently, she reported limiting this practice to ungraded practice writing. A consistent theme throughout the interviews was a desire for

increased teacher autonomy in the assessment process, allowing for judgments based on personal pedagogical beliefs (Rana, II, Excerpts 9.9 & 9.16). This stance frequently underscored the tension between the institutional directives for assessment and Rana's own writing assessment beliefs and practices.

9.8.3 Harmony between Rana's Writing Assessment Beliefs and Practices

In Rana's data, it was observed that there was considerable alignment between her stated beliefs in practice and her actual practices. For instance, Rana articulated a practice of addressing "everything" in the rubric criteria (Rana, II, Excerpt 9.5), and she demonstrated this by equitably focusing on each criterion rather than prioritizing one over another. This comprehensive approach to evaluating all aspects of the students' paragraphs was exemplified in the analysis presented in the previous section.

Furthermore, Rana did not categorize her feedback techniques explicitly as direct, indirect, or metalinguistic when discussing her feedback approach. Nonetheless, her feedback examples demonstrated her use of these methods. For instance, as illustrated in Figure 9.1, she provided explicit corrections like "do not leave double space," identified errors by type such as "subject-verb agreement" and "fragment," and gave metalinguistic advice like "add a transitional signal to show your readers that this is the end of your paragraph..."

The primary harmony between Rana's assessment beliefs and practices originated from her focus on her students' emotional and psychological well-being during writing evaluations and feedback (Rana, II, Excerpts 9.14 & 9.15). She consistently provided supportive feedback to lower students' affective barriers. Her feedback often included supportive comments and terms of endearment like "dear" and "sweetie." Furthermore, she actively acknowledged the positive aspects of her students' writing, such as well-structured supporting details—a practice

substantiated by feedback examples in Figure 9.1. In another case, she enthusiastically praised an error-free paragraph and the student's application of the writing structures taught in class: "Dear {student name}, THANK YOU FOR THIS PARAGRAPH!!! WOW. I am so, so proud of you! You wrote it in the structure I taught you! & I was really happy reading this! Good job, sweetie!" This enthusiastic feedback exemplified her nurturing approach and her dedication to reinforcing the strengths in her students' writing efforts.

Moreover, the observation of Rana's classroom revealed her consistent efforts to support her students emotionally and psychologically. For instance, before the paragraph outlining session, she boosted their confidence by affirming their abilities: "Ladies, start writing your outlines. And I know you can do it. I know. So please say بسم الله ((In the name of God)) and start writing. And remember, I am always here for you" (Rana, OBS). Such interactions aligned with Rana's belief in alleviating any student apprehension and fostering a sense of closeness and trust (Rana, II, Excerpt 9.5). This nurturing environment was also evident through her in-class use of affirming language, consistently addressing students with terms of endearment such as "Dear" or "beautiful {student name}" (Rana, OBS).

9.9 Summary of Rana's Case Findings

Figure 9.2 summarizes the key findings of Rana's case, illustrating the elements of and interactions among her writing assessment belief system, actual practices, and the ecological systems that influenced them and their relationship. The left side of Figure 9.2 shows Rana's belief system was founded on the necessity of affective approaches in writing assessments to bolster student learning. Her beliefs and beliefs in practice were heavily driven by affection. Rana consistently emphasized fostering emotional bonds, demonstrating empathy, and caring for her students during the writing assessment process. These core beliefs extended to her peripheral

beliefs of advocating for student autonomy in choosing their own topics and word limits for writing assignments, deciding on the most suitable times for exams and quizzes, and providing positive and motivational reinforcement through verbal and written feedback. Rana also valued personal, supportive interactions in one-on-one feedback sessions. Her affective approach also appeared to inform her favoring impressionistic, teacher-based judgments that diverged from stringent assessment protocols and fixed rubric guidelines.

As depicted in Figure 9.2, the macrosystem factors (yellow box) did not seem to exert an influence on Rana's assessment beliefs and practices. However, factors from the microsystem and exosystem, as indicated by the bidirectional arrows, significantly informed Rana's core beliefs and created tensions between her beliefs and actual practices. These tensions did not arise in isolation but were the product of the interplay between various factors from different ecological systems.

Microsystem personal factors, such as Rana's own teaching and learning experiences, profoundly shaped her core beliefs. Positive student interactions and feedback motivated her to explore what was best for them, like peer and one-on-one feedback. A past negative experience in an English writing course in her undergraduate studies strengthened her belief in the importance of students' psychological preparedness, influencing her to cultivate a love for writing courses from the outset. Another microsystem factor, students' proficiency level, interacted with limiting exosystem agents, notably time constraints.

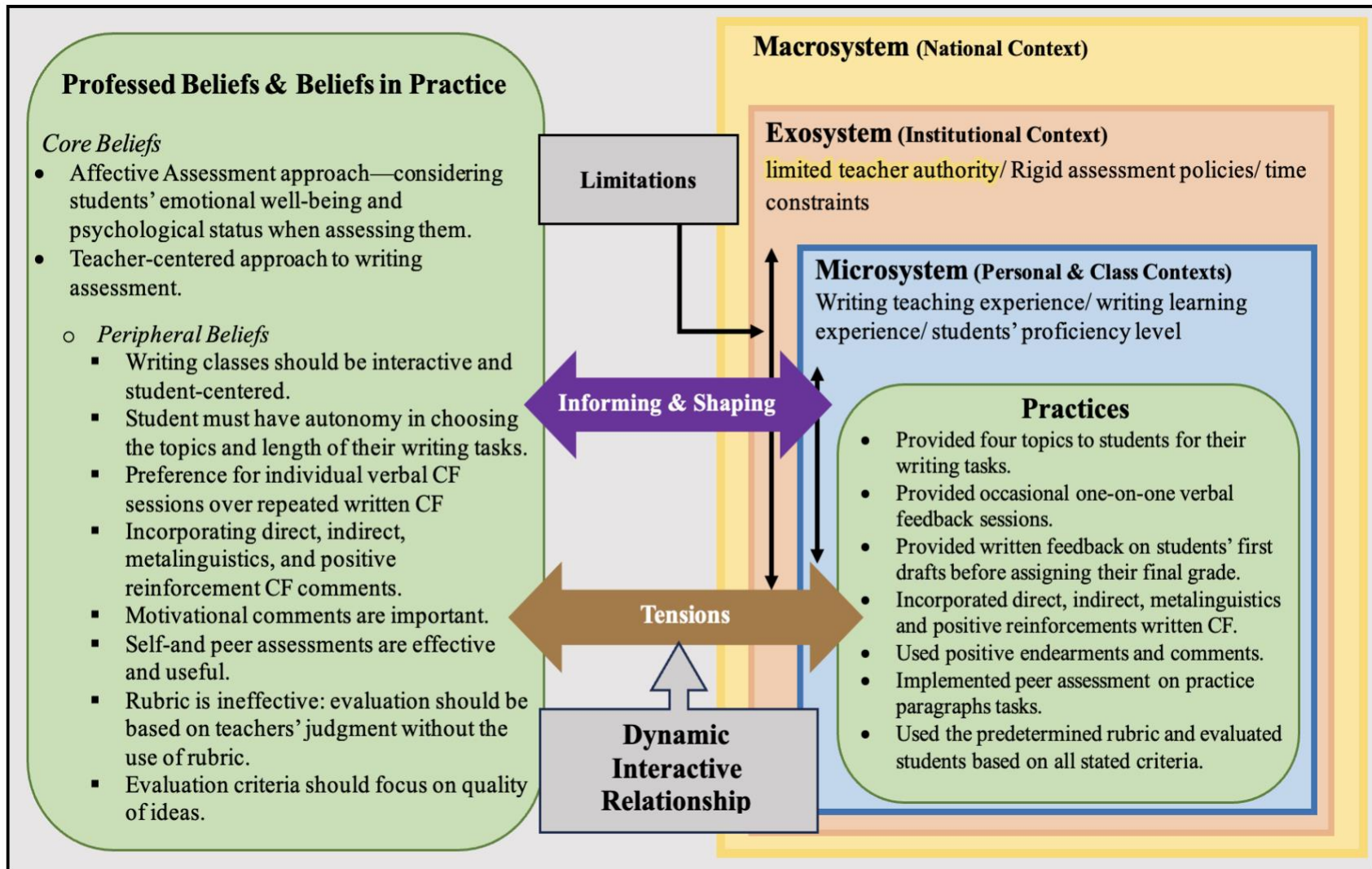
Exosystem agents were primarily responsible for misalignments between Rana's beliefs and actual practices in the PYP. Limited time, influenced by mandated assessment policies and students' proficiency level, restricted her ability to fully implement one-on-one verbal feedback, a practice she valued for its effectiveness. As highlighted in Figure 9.2., limited teacher authority

was identified as the primary source of these misalignments, forcing Rana to adhere to predefined evaluation criteria and rating policies rather than relying on her own judgment.

Harmony between Rana's beliefs and practices did not stem from facilitating ecological factors but from her intrinsic emotional drive (core belief). Her empathetic personality led her to push for and offer one-on-one verbal feedback, despite not being obligated to do so, and adopt a gentler tone in her written feedback, using terms of endearment and praise.

Figure 9.2

Summary of Rana’s Case Findings



CHAPTER 10: CROSS-CASE FINDINGS

This study used complexity theory (CT) to explore EFL teachers' L2 writing assessment beliefs and practices in a PYP at a Saudi university, alongside examining the interrelations among those beliefs, practices, and ecological factors. The research questions were as follows:

1. What are EFL writing teachers' beliefs regarding the assessment of L2 writing?
2. How do teachers' writing assessment beliefs relate to their actual practices?
3. What are the interrelationships between EFL teachers' writing assessment beliefs, actual practices, and contextual factors?

After reporting findings for each case in the previous chapters, this chapter synthesizes and discusses the findings from the cross-case data analysis of the five cases. The results are discussed based on themes connected to the study's research questions and theoretical framework.

Data concerning the first research question revealed one primary theme: the heterogeneity and the interactive nature of EFL teachers' belief systems regarding L2 writing assessment. The heterogeneity of the teachers' belief systems is reflected in the variety of beliefs the five teachers held concerning five interrelated domains the teaching and learning of L2 writing, purposes of assessment, methods of assessment, criteria for evaluating writing, and assessment processes and tools, including feedback, grading, and rubrics. This heterogeneity is further amplified by the variability in the teachers' core and peripheral assessment beliefs. The interactive nature of teachers' belief systems are reflected in how these beliefs interact with one another, typically with core beliefs influencing or informing the peripheral ones.

Analyses concerning the second and third research questions highlighted themes about the interactive relationship between teachers' stated beliefs (i.e., general beliefs and beliefs in

practice) and their actual practices. The findings demonstrated that the relationship between beliefs and practices was non-linear, characterized by tension (misalignment) and harmony (alignment) within and across the various domains of writing assessment beliefs. The tensions and harmonies were found to be contextual and significantly shaped by factors originating from the ecological systems. These included the macrosystem (national and social context), the exosystem (institutional context), and the microsystem (personal and classroom contexts).

10.1 Heterogeneity of Teachers' L2 Writing Assessment Belief System

Heterogeneity in CT refers to the concept that the elements within a complex system and their interactions are diverse in type and nature (Larsen-Freeman & Cameron, 2008). Such heterogeneity was observed in the five EFL teachers' belief systems about L2 writing assessment in two distinct ways: the diversity of the content of the teachers' belief systems and the nature of their beliefs. This section explores how teachers' belief systems were diverse in content that spans various domains of writing assessment, encompassing both core and peripheral beliefs, which interacted with each other.

10.1.1 The Content of Teachers' Belief Systems

The analysis of the five teachers' data revealed a spectrum of beliefs that were primarily associated with five domains of writing assessment: teaching and learning writing, purposes of assessment, methods, criteria for evaluation, and the use of processes and tools (i.e., feedback, grading, and rubric). Each teacher's belief content was unique, reflecting her personal belief system, especially in terms of defining quality writing. Nonetheless, there was a consensus among the teachers on certain beliefs within other writing assessment domains, particularly the methods and purposes of assessment. The subsequent sections will explore in greater depth the individual belief systems of each teacher in those various areas of writing assessment.

Beliefs about Teaching and Learning Writing. A salient component of the teachers' writing assessment belief systems was their beliefs concerning the teaching and learning of writing. While these beliefs were not directly related to assessment, they informed the understanding of the teachers' writing assessment belief systems. They influenced their perspectives on the nature of writing, which in turn affected how they believed writing should be assessed. The beliefs among the five teachers varied widely; they did not hold a uniform view of writing. Leena, Aya, and Noor explicitly expressed their beliefs about the process of learning writing. In contrast, Batool and Rana's beliefs were implicit, discerned indirectly through their discussions on various aspects of writing.

Leena adopted a skill-based approach to writing. She regarded academic writing as instrumental and technical, holding the belief that "writing is just like techniques, chunks. If you learn them, *you can go* [[you can write]]. Especially academic writing" (Leena, II, Excerpt 5.2). Leena believed students could enhance their writing by employing various strategies and techniques, such as utilizing phrase banks and word lists to refine their use of advanced and academic vocabulary. She also valued using written models for composing well-structured essays, including a topic sentence, three main points, and a concluding sentence. Furthermore, Leena viewed writing as a tool for fulfilling academic or professional objectives rather than for its intrinsic value. She believed that writing proficiency could be cultivated through focused practice and instruction, emphasizing that "[the] key factor in any writing course is practice. If the students don't practice, they will not be able to write" (Leena, II).

In contrast, Aya's perspective on learning and teaching writing was centered on fostering students' abilities to generate ideas and construct high-quality content. She believed the formation of the initial idea, especially the topic sentence, to be the most crucial and challenging

aspect of writing. She explained that writing is “how to get the idea clearly written in the introduction” (Aya, II, Excerpt 6.2). Aya supported student autonomy in choosing topics, particularly those that resonated with their studies in health-related fields, as she believed this would significantly increase their engagement and proficiency in writing. For Aya, writing was primarily about content generation, not just technical skills.

Batool did not explicitly articulate her beliefs regarding the learning or teaching of writing. Nevertheless, it appeared that she believed in the development of writing skills over time, through a process of learning from one’s error. She seemed to conceive writing instruction as a facilitating space where students are free to make mistakes and learn from them, thereby allowing them to track their progress by comparing their initial work in the first class to their improved writing at the course’s end.

Noor’s beliefs about the teaching and learning of academic writing focused on creating a relaxed environment that would first encourage informal writing across various genres. Noor advocated for a contemporary student-centered approach to writing that was more focused and incorporated multimedia tools, such as social media and video platforms, into the teaching and learning of L2 writing. She believed in encouraging students to write reflectively and comment on various reading and listening materials, aiming to make writing relevant to their daily lives. Her emphasis was on continuous, free form writing and journaling to help students express themselves and build a foundation for more formal academic writing. She saw consistent writing about diverse topics, coupled with personal reflection on reading materials, as essential to mastering academic writing skills, highlighting the intertwined nature of reading and writing in her teaching philosophy.

Rana's beliefs were shared more explicitly. She believed in an interactive, student-centered approach to teaching writing, favoring active student engagement over traditional, lecture-focused methods. She explained that she "would encourage a lot of practice during the class instead of...explaining a lot of content" (Rana, II, Excerpt 9.2). Rana advocated for student autonomy, suggesting students should be free to choose their writing topics and not be restricted by word counts. She also believed that students should have control over the timing of their quizzes and assignments to manage their schedules better. Rana's teaching philosophy was characterized by a substantial emotional investment in her students' success and a critical view of the existing PYP framework, desiring a more flexible and interactive educational environment.

Beliefs about Writing Assessment Purposes. The five teachers' beliefs regarding the purposes of writing assessments generally revealed a consensus on the formative nature of assessment, which aims to foster student development and learning. These beliefs, which varied in how explicitly they were stated, often intersected implicitly with the teachers' discussions of assessment methods and processes.

Leena's beliefs concerning the purposes of writing assessments were centered on the practical application of assessment methods. Reflecting a skill-based orientation, she advocated for direct writing tasks and writing practice as essential to evaluating students' writing performance, sidestepping MCQs, which she found ineffective for accurate understanding evaluation. She further emphasized the importance of direct assessment in tasks such as paraphrasing, where the actual writing skills could be directly measured, and critiqued the PYP's use of MCQs for this purpose (see Figure 3.1 for a copy of the MCQs in the paraphrasing task).

Batool's belief in the formative purpose of writing assessments also manifested implicitly through her discussion of providing students with constructive feedback essential for learning from their errors. She stated that students "won't know what mistakes they've made... unless you give them the proper feedback," highlighting the role of feedback in helping students recognize and improve upon their mistakes to prevent recurrence (Batool, II, Excerpt 7.3).

Similarly, Rana's beliefs, while implicit, seemed to encapsulate two purposes: enhancing learning through feedback and mitigating the emotional impact of assessments on students. "I feel giving students feedback is the most important part of writing assessment," she conveyed, valuing feedback as crucial for learning (Rana, II, Excerpt 9.4). She further prioritized creating an emotionally supportive assessment environment, "I do not want my students to suffer... I'm here to give them feedback," aiming to alleviate the fear of assessment and encourage expression (Rana, II, Excerpt 9.5).

In contrast, Aya and Noor explicitly addressed the purposes of writing assessments, providing a direct response to the question. Aya believed that the primary purpose of assessment was "to assess [students'] progress in class," placing importance on tracking and encouraging students' individual growth over time rather than their comparative performance (Aya, II). She advocated for assessments that assess students' progress within the frame of the taught content and contribute to genuine learning and opposed the grade-focused culture she perceived within the PYP.

On the other hand, Noor clearly distinguished between the specific objectives of writing assessments within the PYP and more general assessment goals. She stated that the PYP assessments were designed "to prepare [students] for writing research at the end," indicating a progressive development of writing skills as a preparation for future academic challenges (Noor,

II, Excerpt 8.5). Noor also saw assessments as beneficial for enhancing students' English language proficiency, asserting that "assessment helps students to improve their English" (Noor, II, Excerpt 8.6).

Beliefs about Writing Assessment Methods. In the domain of writing assessment methods, the comparative analysis of the five teachers' beliefs revealed a nuanced landscape. Their discourse was predominantly centered around assessment tasks and the value of self and peer assessments, while other methods, such as portfolios, were notably absent from the conversation. Their discussions exposed a shared preference for formative and direct writing tasks instead of summative and indirect ones, such as MCQs in quizzes and exams.

However, the teachers presented a dichotomy in their beliefs regarding self and peer assessments. Batool and Rana were staunch advocates for the efficacy of these methods in fostering students' learning. In contrast, Leena, Aya, and Noor espoused the belief that self and peer assessments should be seamlessly woven into assessment strategies without explicit directives to students. For instance, they contended that students inherently engage in self and peer assessments during the revision process following feedback or participating in collaborative activities to pinpoint errors.

Leena juxtaposed direct writing assessments with MCQs, advocating the former as essential to students' learning within the PYP framework for English II. She critiqued MCQs for failing to measure genuine learning, as students might not grasp the reasoning behind their choices. She emphasized the importance of including actual writing tasks in quizzes and exams, acknowledging that "focusing on writing assignments is the most important thing to assess the student's performance" (Leena, II, Excerpt 5.5). Yet, her articulation of self and peer assessments was sparse, suggesting a peripheral role for these methods in her belief system.

On the other hand, Aya believed that exams and quizzes were unnecessary and that students should be evaluated on their writing ability, valuing the use of few direct writing tasks over the use of many tasks of varied genres. She also criticized MCQs that tackled the identification of grammatical rules, questioning, “Why would [students] need to know the name of the rules? They should be able to apply those rules in their writing” (Aya, II). Moreover, Aya perceived peer assessment as beneficial, as students of different levels could aid each other, and self-assessment as an intrinsic activity that students engage when proofreading and improving their work based on her feedback.

Batool’s advocacy for direct formative assessments was predicated on her belief that they facilitate learning from mistakes, which she emphasized by asserting, “It’s really normal to make mistakes” (Batool, II, Excerpt 7.5). She believed in the importance of extensive writing practice, leveraging self-assessment for students to identify errors after receiving feedback, and fostering student independence and collaboration through peer assessment in non-graded assignments. Similarly, Noor emphasized direct formative assessments, advocating for a succession of writing tasks coupled with feedback as vital for writing development. However, she perceived self and peer assessments as part of verbal group feedback activities in class, when students critique drafts collaboratively, akin to a teacher’s role.

Rana shared her beliefs regarding writing assessment methods, focusing on both the PYP context and her own beliefs. She emphasized the practicality of indirect methods like MCQ quizzes within the PYP as the English II curriculum focused on assessing students’ understanding of grammatical rules and paragraph structure, arguing that MCQs “are very useful because... [in English II] we’re not judging their ideas.” However, she recognized the creativity-limiting nature of such methods (Rana, II, Excerpt 9.10). Rana personally preferred a more

liberated approach, focusing on direct writing assessments and encouraging students to explore practicing with various paragraph structures without constraining their ideas to specific topics or word limits. Moreover, Rana strongly believed in the value of self- and peer assessments in promoting students' editing skills.

Beliefs about Writing Evaluation Criteria. Teachers' beliefs about writing assessment criteria were diverse too. Leena emphasized grammatical precision, whereas Aya and Noor favored the clarity of content over mechanical and grammatical accuracy. Batool and Rana, though somewhat inclined to prioritize content, emphasized the importance of adhering to the course requirements, focusing their criteria on the taught materials of any course, including English II.

Leena's beliefs about writing evaluation criteria assessment strongly emphasized grammatical accuracy, which she viewed as crucial as content in student writing. Leena considered sentence structure the most significant area where students make errors. Leena explained that in her ideal writing class, she would prioritize teaching grammatical rules, such as comma usage and article application, over discussing organizational elements like topic sentences. She believed mastering grammar was essential for students, linking effective writing to the absence of grammatical errors. Her belief in a skill-based approach to teaching writing, with a grammar focus at its core, revealed Leena's belief that writing served primarily as a vehicle for grammar mastery.

In contrast, Aya held an opposing view, placing the highest value on content clarity in writing assessments. She argued against the PYP overemphasis on spelling, given the advent of spellcheck tools, and opted to assess the substance of students' ideas instead. As a belief in practice, Aya believed in evaluating the strength of students' main points, topic sentences, and

conclusions while being lenient towards spelling mistakes. This stance indicated that Aya prioritized the understanding of content over stringent grammar or spelling rules, emphasizing the importance of content mastery in L2 writing.

Echoing Aya's beliefs, Noor also stressed the significance of content and coherence over strict grammatical correctness in writing assessments. She favored a method that values meaningful content without harsh penalties for minor grammatical or spelling mistakes. Promoting a relaxed teaching style, Noor suggested journal writing that overlooks stringent formalities, focusing on significant errors only when they obscure meaning. This reflected her prioritization of idea expression and coherence above meticulous grammatical and spelling precision.

Batool, however, adopted a more holistic view, believing that assessment criteria should encompass "everything, organization, formatting, and even content" (Batool, II, Excerpt 7.10). Her comprehensive approach involved assessing "all types of mistakes," highlighting the importance of a well-rounded evaluation that considered how ideas relate to the topic sentence and the overall flow of the writing. Batool's criteria seemed to align closely with the English II curriculum's standards, suggesting a thorough commitment to the established evaluative framework of the English II course.

Lastly, Rana also focused on assessing students' comprehension of the material taught, examining "every rule they've learned" (Rana, FI 1). Her belief reflected a dedication to academic rigor and adherence to the specific rules of English II. In addition, Rana considered students' psychological preparedness and the pressures of exam timing and anxiety as critical factors in their performance, advocating for a compassionate and multifaceted approach to evaluating their writing.

Beliefs about Writing Assessment Processes and Tools. The participating teachers' beliefs about writing assessment processes and tools primarily centered on feedback, grading, and rubrics. When prompted about other tools, processes, and ideal writing classes, their responses appeared influenced by their teaching experiences in the PYP. All of these teachers began their teaching careers in PYPs, except for Rana. In Saudi PYPs, feedback, grading, and rubrics are key elements in writing assessment. Hence, it can be deduced that their writing assessment beliefs and practices are influenced by these processes and tools. The five teachers agreed on the importance of feedback in enhancing student writing abilities, yet they held diverse opinions on feedback forms, types, and approaches. Similarly, their perspectives on the effectiveness of analytical rubrics featuring specific criteria were diverse.

Leena advocated for an indirect and detailed approach to feedback in writing assessment, believing that it encourages student autonomy by identifying errors without providing corrections; as she mentioned, "I like to give them detailed feedback, specifying where they make mistakes, but I don't give them the correction" (Leen, II, Excerpt 5.8). Due to its permanence and clarity, Leena valued written feedback over verbal feedback despite acknowledging that verbal feedback can be more detailed. She reserved verbal feedback for students' preference for one-on-one sessions during her office hours, preferring written comments as the primary mode of communication. Despite reporting that she occasionally forgets to provide positive feedback, she was aware of its importance in balancing students' need for improvement with positive reinforcement. As a belief in practice, Leena believed in using error codes aligned with course materials for consistency and clarity; however, she opted for more comprehensive feedback comments when necessary. In terms of grading, Leena strongly believed in the use of rubrics to mitigate subjectivity and ensure fairness, stating, "With the

rubric, you will have fair grading,” reflecting her belief in the rubric as a tool for equitable evaluation (Leena, II, Excerpt 5.11).

In contrast, Aya valued one-on-one verbal feedback in the classroom for its immediacy and interaction, believing it significantly improves student writing. She emphasized real-time assistance, stating, “Verbal feedback really helps because I can tell them, ‘This is not very clear,’ ‘You can do it like that’” (Aya, II, Excerpt 6.6). Aya preferred this hands-on method as it allows for immediate correction and clarification and permits using the student’s L1, advantages she found lacking in written feedback. Aya also believed in a mixed approach to feedback, providing direct feedback on minor issues such as spelling, punctuation, and formatting and indirect feedback for content-related errors. For grading, she advocated for a holistic approach that prioritizes content clarity and student progress. Aya challenged the fairness of rubrics, suggesting they fail to reflect individual learning trajectories: “What would really help with the grading is if we consider every student individually... The rubric is unfair” (Aya, II, Excerpt 6.7). However, she acknowledged that rubrics could ensure consistency across teachers in competitive contexts, like the PYP, yet she strongly supported personalized, progress-focused assessment over strict adherence to rubrics.

Unlike Leena and Aya, Batool believed in integrating both written and verbal feedback, convinced that providing both yields the most significant learning benefit. She viewed feedback as an interactive process, dedicating time to personalized sessions that discuss written comments, and valued the power of dialogue to enhance understanding. Batool also favored an indirect feedback approach, encouraging students to independently identify minor errors, thereby promoting self-assessment and error recognition skills. She believed in using different written feedback types, like error codes for minor issues and suggestive comments for content clarity.

For grading, she avoided assigning grades on practice tasks to reduce stress, focusing instead on constructive feedback. Additionally, she supported the rubric use to prevent unnecessary grade deductions and advocated for ensuring that students understand the rubric criteria to be self-aware and earn the grades they deserve.

Noor also viewed feedback as the most critical aspect of writing assessment, as it helps students enhance and develop their writing skills. She stated, “I feel like the most important part is giving feedback to the students... I’m not gonna give her the answer for sure, but I’m gonna guide her to the correct answer, the correct sentence” (Noor, II, Excerpt 8.12). She favored indirect written feedback, where she would highlight errors and comment on them, enabling students to understand why something was wrong and how to improve their corrections. Noor believed in customizing her feedback to meet the needs of individual students, providing direct feedback and corrections to weaker students to aid their understanding, and then scaling back to indirect feedback. Regarding grading, Noor expressed that, given the choice, she would opt for qualitative evaluation over precise numerical grades. She preferred broader grade ranges, stating, “A range is better because I don’t have to think about giving a mark or not” (Noor, II, Excerpt 8.18), highlighting her preference for assessing overall performance rather than focusing on specific errors and criteria.

Rana criticized the PYP feedback policies, arguing that the practice of giving written feedback before final grades was ineffective and can lead to student dependency. She believed it was a waste of time, akin to “spoon-feeding,” and stated, “I will just ask them to write, but I will not give them feedback. I will tell them, ‘You will receive your final grade, and maybe after your grade, I will give you feedback’” (Rana, II, Excerpt 9.11). However, Rana preferred one-on-one verbal feedback sessions over written comments to personalize the feedback process. She was

mindful of students' stress and believed that students should record verbal feedback sessions, allowing them to focus on feedback without the fear of forgetting it. Rana also believed that incorporating positive reinforcements in verbal or written feedback should be a deliberate strategy to inspire and encourage students, thereby enhancing their writing skills. In terms of grading, Rana was against the fixed rubric system, calling it "super unfair" and ineffective, and suggested that "just let the teacher evaluate her students" (Rana, II, Excerpt 9.17). She advocated for relying on the teacher's judgment and knowledge of their students over the use of a rigid rubric to evaluate students' writing.

10.1.2 The Nature of Teachers' Beliefs

As indicated earlier, beliefs within a belief system can vary, encompassing implicit or explicit beliefs. Additionally, a belief's strength also determines its role as an 'attractor' within the belief system. An attractor, or stable belief, serves as a core belief that shapes and informs other peripheral beliefs and remains unchanged over a long period (Lucas, 2004). All teachers, except for Leena, held core beliefs that exhibited a strong presence when articulating their views, which consistently influenced the entire trajectory of their writing assessment belief systems.

Aya prioritized individual progress and the clarity of content over linguistic perfection. Her core beliefs seemed to inform and interact with her peripheral beliefs, fostering a holistic and empathetic writing assessment approach that values the student's learning journey and content comprehension above grammatical precision. Thus, Aya's peripheral beliefs highlighted her dedication to the developmental aspect of learning to write and that "the ideal assessment will consider that this is a class for second-language learners" (Aya, II, Excerpt 6.10). This extended to her preference for personalized verbal feedback and her critical stance on the PYP exams and rubrics that prioritize achievement over progress.

Batool firmly advocated for the formative use of assessment, particularly self-assessment. She contended that formative assessment plays a crucial role in guiding and enhancing student writing and asserted that students must take ownership of their learning. According to her core belief, assessments should facilitate self-directed learning, fostering autonomy and enabling self-correction. Batool articulated, “I always encourage them to self-assess...be more self-independent... I would point out most of the mistakes that they have written in the first draft. But then, I will teach them to be more independent” (Battol, II). Her core belief informed several of her peripheral beliefs. She believed in the benefits of both indirect feedback and positive reinforcement for boosting student independence and motivation. Furthermore, Batool perceived error-making as integral to the learning journey, advocating for extensive writing practice to allow students to recognize and correct their own mistakes independently. She also stressed the significance of clarifying rubric criteria to students to promote their accountability in the learning process.

Noor centrally believed in the value of informal writing genres as a foundation to teaching academic writing, employing multimedia tools to integrates listening and reading. She stressed the need to cultivate writing skills relevant to real-life situations and believed in assessments that emphasize meaning and coherence over rigid grammatical precision. Consequently, Noor’s peripheral beliefs concerning grading and evaluation criteria emphasized the importance of effective communication and qualitative feedback rather than precise numerical scores, in assessing student overall performance rather than focusing on detailed grading.

Rana’s teaching and assessment methods were firmly anchored in an affective approach, as she focused on students’ psychological well-being, explaining, “I focus on their psychological side. I tell them, ‘Ladies, I appreciate your ideas. Don’t be afraid. I’m here to help you. I’m here

to give you feedback” (Rana, II, Excerpt 9.5). Her distinctive assessment beliefs were characterized by fostering an emotional connection, demonstrating empathy, and caring for students’ feelings. Consequently, Rana believed in supporting student autonomy, encouraging personalized topics, and using motivational reinforcement in verbal and written feedback. This affective perspective informed Rana’s peripheral beliefs in a more subjective and relationship-focused evaluation process, which stood apart from conventional objective assessment practices.

Finally, Leena’s core beliefs seemed less defined in a state akin to a basin of attraction—not fully crystallized within her belief system or substantially shaping it. This could be attributed to her eight years of experience in the PYP, which might have introduced a level of instability in her personal beliefs regarding writing assessment. This instability might account for aligning her core writing assessment beliefs with the English II curriculum, which emphasized grammatical and mechanical precision. Leena centrally believed in a skill-based approach, emphasizing grammatical accuracy as a central part of writing competence and a primary indicator of student achievement. This belief seemed to underpin her peripheral beliefs of focusing her evaluation criteria mainly on grammar and her endorsement of the PYP rubric as a more structured and technical tool for writing assessment.

10.2 Interactions Between Teachers’ Beliefs, Practices, and Contextual Factors

Teachers’ writing assessment belief systems, actual practices, and ecological systems are each characterized as CASs. They consist of various subsystems and exhibit non-linear interactions both internally and among each other (Larsen-Freeman & Cameron, 2008). As indicated above, teachers’ beliefs about writing assessment were heterogenous; they involved a spectrum of core and peripheral beliefs interacting with each other. Similarly, teachers’ actual practices were not isolated actions; they encompassed a range of connected activities such as

feedback focus and approaches, as well as grading practices, which include the application of rubrics. Furthermore, ecological systems were found to be interconnected, where factors from one system amplified the effects of factors from another system.

This concept of interaction within CASs suggested that influences were reciprocal, creating a complex network of interrelated elements. These interactions were primarily characterized by harmony or tension. Various ecological agents influenced teachers' L2 writing assessment beliefs, practices, and the relationship between them, ultimately shaping their pedagogical realities. For instance, factors within the microsystem influenced the formation of teachers' core beliefs within their writing assessment belief systems. In contrast, exosystem factors appeared to restrict or facilitate the alignment between teachers' writing assessment beliefs and practices.

10.2.1 Factors within Ecological Systems

As indicated above, ecological systems are inherently complex, consisting of multiple concurrent levels that engage in interactive interplay. Within the overarching framework of ecological systems—namely the macrosystem, microsystem, and exosystem—it became evident that these levels were not isolated entities; instead, they exerted mutual influence upon one another. For instance, fixed assessment policies and curricula were developed by a curriculum committee outside the teachers' direct environment. The course coordinator acted as the medium to distribute these policies to teachers, dictating the use of pre-established quiz and task templates, rubrics, and curriculum guidelines. Such rigid hierarchical structures, along with students' limited proficiency level, imposed time restrictions that curtailed teachers' ability to implement their beliefs about assessing writing effectively. Moreover, within the PYP, teachers had limited authority, significantly affecting their choice of assessment methods, tools, and

grading criteria. This constrained autonomy hindered teachers from fully applying their preferred approaches to writing assessments. Therefore, it was imperative to acknowledge that while these factors were categorized within specific ecological systems for this study, such a classification was not inflexible. Instead, it served as a means to elucidate these contextual factors' impact on the relationship between teachers' beliefs and their actual practices.

10.2.2 The Influences of Macrosystem Factors on Teachers' Beliefs and Practices

At the macrosystem (national context) level, factors or agents appeared to have a minimal direct impact on the five teachers' writing assessment beliefs, practices, or the interplay between the two. Nonetheless, some teachers held perceptions regarding these factors that shaped their opinions about some aspects of the national education framework. Specifically, Leena and Batool discussed the effects of ELT policies within the Saudi educational system on students' English language progression in higher education. Leena observed that students grapple with English courses in the PYP due to a lack of early English language education and insufficient exposure during their foundational schooling. This shortfall in language acquisition was seen to extend into their undergraduate experience and further. Leena remarked, "Students graduate from high school without basic grammar" (Leena, II).

Batool, conversely, argued that English language instruction in schools should seamlessly transition into and support higher education. She advocated for a curriculum where university English instruction builds upon what was taught in high school, emphasizing the need to teach more advanced aspects of English that align with students' specialized fields of study. Batool explained, "English *here* [[in the University]] should be an addition to what the students took in high school, not a repetition... We're supposed to teach them something new, things that they would need the most, things related to their medical field, to their major" (Batool, II). Although

Leena and Batool's perspectives did not seem to affect their writing teaching or assessment practices directly, it highlighted their beliefs regarding students' English language proficiency and their critical view of ELT in KSA.

10.2.3 The Influences of Exosystem Factors on Teachers' Beliefs and Practices

Factors at the exosystem (institutional context) level were the primary cause of tensions between the beliefs and practices of the participating teachers regarding L2 writing assessment. The most impactful external agents for the five teachers included curricular limitations, rigid assessment protocols, constrained teacher authority, and time restrictions.

Firstly, the central cause of contention between most teachers' beliefs about assessing writing and their actual practices were the English II curriculum, assessment approaches, and limited teacher input in deciding or modifying the curricular and evaluation methods and policies. For instance, both Leena and Aya preferred direct writing assignments to evaluate writing competencies effectively. However, the PYP's focus on MCQs for assessing writing skills hindered the application of their assessment beliefs in their practices. As a result, Leena's class instructional methods, for example, revolved around preparing students for MCQs despite her beliefs that such techniques did not accurately assess writing abilities. Similarly, Noor's ambition to instruct writing in a more casual and spontaneous setting, fostering student engagement in free writing genres like journaling, was markedly constrained by the PYP's stringent curriculum, which offered minimal flexibility for creativity and individualized teaching and assessment approaches.

Additionally, the demands of the PYP assessment guidelines led to a tension between the teachers' writing assessment beliefs and their practices. For instance, Aya placed a high value on the subjective evaluation of each student's unique development and favored more encompassing

grading categories that could capture individual progress. However, the PYP's detailed scoring guide and numerical grading framework compelled her to evaluate students' accomplishments differently, necessitating a precise and uniform application of the rubric's criteria. Even though Aya and Noor were proponents of prioritizing content and the generation of ideas in writing, the requirements of the PYP rubric obliged them to concentrate on spelling, grammar, and format.

The PYP's specific feedback policies, which mandate that teachers provide written feedback on students' first drafts, created a discrepancy between the teachers' feedback beliefs and the feedback they actually provided. For instance, Rana advocated for offering feedback simultaneously with or after grades, equating providing feedback before the final grade to over-assistance. On the other hand, Aya saw more value in one-to-one verbal feedback sessions rather than written feedback. Nevertheless, both teachers were obliged to conform to the PYP feedback regulations and deliver written feedback on students' first drafts before grading.

Finally, the PYP's rigid curriculum and evaluation policies, in conjunction with the limited authority of teachers, induced additional issues like time limitations. Leena, Batool, and Aya noted that the dense syllabus and the numerous graded assessments prioritizing MCQs often restricted their time for conducting peer assessments or offering detailed verbal feedback on direct writing tasks. Batool also highlighted that the strict time requirements set by the syllabus and course coordinator hindered her from employing her preferred teaching methods, such as extensive ungraded writing exercises.

Despite tensions between beliefs and practices due to external pressures, teachers showed adaptability to these challenges. CASs may self-modify in response to environmental changes; this is evident in studies of teacher cognition, which illustrate how teachers' belief systems, informed by their historical and social contexts, are flexible and adapt to situational constraints

(Feryok, 2010; Zheng, 2015). This was the case in the present study too. For example, Aya preferred to assess student progress impressionistically rather than through a rubric or numerical grades. However, she recognized that within the context of the PYP, rubrics were essential for ensuring equitable assessment across various classes and instructors. As a result, she adjusted her assessment practices to this environment by employing a systematic method that involved counting mistakes and subtracting points according to the number of errors. Similarly, Rana favored direct writing assessment but recognized the PYP's structured approach and the practicality of using MCQs to assess student outcomes within its curriculum. As a result, her primary evaluation criteria beliefs adapted to the context by focusing on evaluating students' adherence to the prescribed assessment format and criteria.

While most teachers voiced critiques of the PYP's curriculum and assessment policies, certain elements actually facilitated the integration of specific writing assessment beliefs into their practices. Leena, for example, supported the use of rubrics mandated by the PYP's policy because it aligns with her belief in the value of using rubrics to assess student writing. Aya found that a policy mandating the display of writing task instructions and explicating each step was beneficial for students' task comprehension. Batool also saw value in the policy of explaining rubric criteria to students. Therefore, it could be posited that some practices mandated for all PYP writing teachers helped them see the value of these methods and might have even nurtured the development of new beliefs within their L2 writing assessment belief systems.

Additionally, the lack of exosystemic pressure in certain aspects of writing assessment afforded teachers the liberty to incorporate their assessment beliefs into their practices. For instance, while adhering to the PYP mandate of providing written feedback on the first draft of students' writing tasks, teachers could still apply their personal beliefs about effective feedback

forms. Aya, believing strongly in the value of one-on-one verbal feedback, made it a point to move among her students in class as they worked on their paragraph outlines, offering personalized oral guidance. In a similar vein, Batool integrated individual verbal feedback sessions to discuss her written comments with students, dedicating time to each one.

Furthermore, teachers' beliefs regarding written feedback approaches faced no opposition from stringent policies. As a result, there was a congruence between their professed beliefs about feedback approaches and their actual practices. For example, Batool avoided providing corrections, trusting in the importance of students discovering and fixing their own mistakes. The leeway granted in feedback practices thus fostered a closer alignment between teachers' beliefs and their actions. The probable reason for this gap in policy may be the perception of feedback as an instructional tool rather than an assessment measure, which leads to fewer restrictions in this area.

10.2.4 The Influences of Microsystem Factors on Teachers' Beliefs and Practices

The openness of a belief system, as described by de Bot (2008), Larsen-Freeman and Cameron (2008), and Lucas (2004), are also attributed to its sensitivity to initial conditions. This sensitivity indicates that minor changes in early stages and interactions within cognition or the environment can lead to the formation and evolution of certain beliefs. The microsystem factors, particularly teachers' personal learning and teaching experiences, played a pivotal role in shaping their core beliefs concerning writing assessment. Consequently, the strong stance of these beliefs was the primary cause of harmony between teachers' beliefs and their actual practices.

Leena's experience with learning English influenced part of her core beliefs about writing assessment, which were rooted in a skill-based approach to writing instruction. Her pedagogical

philosophy emphasized practical tools like phrase banks, vocabulary lists, and writing models. These tools, beneficial in her own writing journey, were evident in her teaching practices. In her class, she encouraged her students to incorporate advanced vocabulary from reading courses and provided exemplary analyses to bolster their writing capabilities. This alignment was also reflected in Leena's class assessment practices, which involved thoroughly monitoring and reinforcing the student's grasp of such technical abilities before developing paragraph outlines. She actively aided students in dissecting written models and enriching their comprehension of paragraph construction, including writing topic sentences, supporting details, and conclusions.

Aya's perception of herself as an L2 learner significantly shaped her core beliefs about writing assessment, emphasizing progress rather than grammatical accuracy and recognizing that both she and her students are "still learning" (Aya, II & TAP 2). Her empathetic stance, informed by her L2 learning experiences, led to an assessment approach that values ideas and learner progress. In practice, she provided personalized verbal feedback and ensured students' works addressed prompts clearly. Moreover, Aya's grading practices balanced fairness within the PYP, where students compete for positions in medical school, and the generosity required to support non-native speakers, recognizing their ongoing language learning journey. To this end, she methodically tallied mechanical and grammatical errors, exhibited leniency towards certain mistakes, and undertook three rounds of reviewing students' paragraphs to guarantee equitable and comprehensive evaluations.

Batool's English learning experience has significantly influenced her core beliefs about writing assessment, particularly regarding the importance of learner autonomy. Her philosophy highlighted the value of self-directed learning, a concept she embraced during her own education in a Saudi public school. There, she cultivated a strong appreciation for English, taking the

initiative to learn outside the classroom by purchasing English books and practicing English through texting with friends. These beliefs were in harmony with her assessment practices, in which she promoted student motivation and engagement by incorporating self-assessment and independent error identification and correction. She employed positive feedback to bolster student learning and achievement. For instance, she deliberately left specific errors in student-written paragraphs for them to detect and amend, thereby reinforcing self-learning.

Noor's learning of English writing began with informal personal expressions jotted down on her phone. This habit cultivated a belief in the significance of informal writing as a foundational step in learning to write effectively. Her academic achievements, including a master's degree from the U.S. and her participation in assessment-related workshops, also shaped her understanding of writing assessment. She learned to differentiate between significant errors affecting meanings and minor mistakes that do not affect meaning. These experiences reinforced her belief in creating an informal, supportive environment that encourages students to focus on freely expressing their ideas in writing without the pressure of grammatical perfection or academic constraints. In her actual assessment practices, Noor prioritized the clarity and logic of ideas over grammatical structure during the outlining phase of the writing tasks in the classroom, emphasizing the development of coherent and meaningful ideas.

Finally, Rana developed a pedagogical approach prioritizing students' emotional well-being, influenced by her own educational experiences that were marked by anxiety and a fear of writing due to past negative experiences as an undergraduate student. Accordingly, Rana's assessment beliefs were characterized by a personal and supportive approach, favoring impressionistic, teacher-based judgment over strict adherence to established assessment protocols and fixed rubric guidelines. These beliefs motivated Rana to foster positive student

interactions by exploring the best approaches to assessment in her writing classes, like peer and one-on-one verbal feedback.

These narratives collectively illustrated the sensitivity of belief systems to initial events or critical incidents, specifically the significant influence of teachers' learning experiences. The five teachers' personal experiences not only shaped their core beliefs about how teaching and assessing writing should be conducted but also facilitated harmony between their beliefs and their practices. The cases demonstrated that teacher cognition is not merely a collection of different belief systems but also includes the process of making teaching-related choices that are informed by complex, application-focused, individualized, and context-sensitive networks of knowledge, reasoning, and values.

Furthermore, other elements within the microsystem (classroom context), such as students' limited proficiency and motivation, as well as classroom technology, were recognized by Leena, Aya, and Rana as both obstacles and facilitators. Nevertheless, these aspects were not found to directly influence the alignment between the teachers' beliefs and their practices concerning L2 writing assessment. For instance, Leena pointed out how Microsoft Teams served as a beneficial tool in improving communication with her students. Conversely, Aya voiced her frustrations with the university's Internet service, which often wasted class time as she attempted to connect the smart board to the computer. Meanwhile, both Rana and Leena observed that the students' proficiency levels and motivation frequently diverted their instructional time to explaining fundamental grammatical rules rather than concentrating on teaching writing.

10.3 Summary of Cross-case Findings

Findings regarding the first research question unveiled a rich tapestry of beliefs among the five EFL teachers, characterized by heterogeneity and interactive interplay. These teachers

held various interacting core and peripheral beliefs that spanned multiple domains, such as the teaching and learning of L2 writing, assessment purposes and methods, evaluation criteria, and assessment processes and tools like feedback and rubrics.

Each teacher's beliefs were heterogeneous, generally unique, and reflected their individual systems of L2 writing assessment beliefs, particularly within the domain of beliefs about teaching and learning L2 writing. Leena adhered to a skill-based, technical approach to academic writing, viewing it as a collection of learnable techniques, whereas Aya concentrated on fostering the generation of ideas and the construction of content. Batool championed skill development through learning from errors independently. Conversely, Rana promoted students' active involvement and self-determination in selecting writing topics and managing deadlines, all within a supportive environment. Likewise, Noor endorsed a modern, learner-centered approach that incorporates multimedia tools to enhance reflective writing that resonates with the students' everyday experiences.

However, there was common ground among the teachers regarding certain beliefs related to writing assessment purposes and methods. They collectively recognized the formative and practical nature of writing assessments, advocating for direct writing tasks to cultivate student development and learning. Leena, for example, preferred extensive writing practices that can develop students' writing skills. Batool and Rana also supported direct writing assessment and stressed the importance of feedback for learning improvement. Aya aimed to assess individual progress within the framework of the writing course, and Noor believed in the role of assessment in advancing students' writing and English proficiency. Despite a shared preference for direct writing tasks, beliefs concerning self- and peer assessments were mixed: Batool and Rana

endorsed them as clear instructional tools for student learning, while Leena, Aya, and Noor perceived them as integral components of the revision and feedback cycles.

The teachers held varied beliefs about L2 writing evaluation criteria, procedures, and tools. Leena emphasized grammatical precision, while Aya and Noor gave precedence to expression clarity, considering it more important than strict grammatical conformity. Batool and Rana believed in evaluating students based on their understanding of course content and adherence to course objectives. Regarding assessment procedures, while all teachers valued feedback, their beliefs in practice differed: Leena was in favor of indirect written feedback; Aya leaned towards one-on-one verbal feedback; Batool found value in a blend of indirect written and verbal feedback; Noor championed feedback tailored to individual needs; and Rana criticized the program feedback policies, advocating for feedback to be given post-grade assignment. Regarding the use of rubrics, views were split: Aya and Leena considered them crucial for ensuring fairness in structured programs such as PYPs, while others viewed them as potentially constraining teacher discretion.

The teachers' beliefs about L2 writing assessment displayed varied natures and interactions. All teachers possessed core beliefs that acted as strong 'attractors' within their belief systems, influencing their peripheral beliefs about writing assessments, with Leena being an exception. Aya's beliefs, which valued individual progress and content clarity, informed her assessment beliefs in one-on-one verbal feedback. Batool's fundamental belief in the importance of formative assessment guided her advocacy for self-assessment and student independence. Noor focused on informal writing and the use of multimedia, highlighting the importance of real-life applicability and hence prioritized qualitative feedback over grammatical exactness. Rana's empathetic approach emphasized student psychological well-being and empathetic assessment

methods to encourage student self-assurance, thereby favoring teacher judgment. In contrast, Leena's beliefs, influenced by her experiences in PYPs, appeared to be more fluid, focusing on grammatical accuracy in line with the English II curriculum standards but without a clear core belief.

The investigation into how teachers' writing assessment beliefs relate to their actual practices (research questions 2 and 3) unveiled a nuanced picture. As indicated above, the narratives illustrated that teachers' beliefs about writing assessment were heterogenous; they involved a range of interacting core and peripheral beliefs. Similarly, teachers' practices were not isolated but consisted of interconnected activities, such as feedback methods and grading practices, mainly influenced by teachers' core beliefs and various contextual factors. Additionally, the analysis revealed the complexities of ecological systems, highlighting how different factors from different levels—microsystem, exosystem, and macrosystem—each play a role in influencing teachers' writing assessment beliefs and practices and their relationship.

Firstly, it was found that teachers' core beliefs about L2 writing assessment were influenced by their personal learning and teaching experiences (microsystem), which markedly impacted their practices in the classroom. There appeared to be a tendency among teachers to emulate the way they themselves were taught. These personal experiences were instrumental in shaping their beliefs and facilitating a harmonious relationship between their beliefs and practices. For example, Leena's preference for a skill-based approach in writing assessment, rooted in her own English learning experience, was manifested in her classroom methods, where she used practical resources like phrase banks and writing models. Aya's perception of L2 learners was influenced by her experience as an L2 learner herself, leading to her prioritization of individual students' progression and focusing on idea development over grammatical precision.

Batool's emphasis on learner autonomy, a belief stemming from her self-directed writing learning experience, was reflected in her classroom practices of encouraging student self-assessment and independent error correction. Similarly, Noor's belief in the importance of informal writing in developing academic writing, shaped by her own informal writing learning experience, was evident in her classroom practices of prioritizing clarity of ideas and logical flow over strict adherence to grammatical structures. Rana's assessment beliefs and practices, which focused on students' emotional well-being and impressionistic judgment, were found to result from her past educational experiences marked by anxiety. Her beliefs guided her toward cultivating a nurturing setting where personalized feedback took precedence.

Additionally, exosystem factors, such as strict curriculum requirements of the program, formal assessment protocols, and constrained teacher autonomy, stood out as prominent influencers. These elements often created tension between teachers' beliefs about assessments and their actual practices, particularly when it came to the selection of assessment methods and tools. On the other hand, macrosystem factors, such as national educational policies, insidiously influenced teachers' perspectives regarding English learners in KSA as lacking basic English knowledge. This was particularly evident in the cases of Leena and Batool, though these broader policies did not seem to directly affect their specific beliefs or practices related to L2 writing assessment.

CHAPTER 11: DISCUSSION AND CONCLUSIONS

This chapter summarizes the main findings of this study and relates them to previous studies on teacher cognition. The discussion is structured around the use of CT to explore the L2 writing assessment beliefs and practices of the five EFL teachers in the study and examine the interplay among these beliefs, practices, and ecological factors. Subsequently, the chapter delineates the study's limitations. Finally, it outlines the potential implications of the findings for teachers and policymakers in the EFL educational landscape and implications for future research in the domain of teachers' beliefs.

11.1 Summary and Discussion of the Findings

In this study, CT served as an analytical lens to explore the writing assessment belief systems, practices, and their interplay within the ecological systems of five EFL teachers: Leena, Aya, Batool, Noor, and Rana. The application of CT metaphorically enabled a comprehensive examination of the intricate interactions among the various CASs: belief systems, practices, and ecological systems. This approach aligns with previous research underscoring the utility of CT in investigating teacher cognition (Alzaanin, 2019; Feryok, 2010; Finch, 2010; Gao & Zhou, 2021; Kiss, 2012; Li, 2013; Yu et al., 2020; Zheng, 2015).

Larsen-Freeman and Cameron (2008) suggested that an effective CT application should detail the system's components, the influencing conditions, and their interconnectedness while illustrating the complex nature of these relationships. This study embraced this methodology, not only narrating the experiences of the five teachers but also demonstrating the efficacy of CT as a tool in unraveling the complex writing assessment realities of EFL teachers in the specified PYP.

A significant contribution of CT to this study was its facilitation of understanding and visualizing the complex interactions within EFL teachers' belief systems relating to L2 writing

assessment. It highlighted that beliefs, practices, and ecological factors do not exist in isolation or simple linear, consistent, or contradictory relationships. Instead, they exhibit varied levels of interaction, challenging the notion of dualistic or narrowly causal relationships. This finding echoed the long call in existing literature for a departure from simplistic, reductionist perspectives in analyzing the complex interplay between language teachers' beliefs and practices (Borg, 2019; Breen et al., 2001; Fives & Buehl, 2012; Li, 2012; Phipps & Borg, 2009; Tudor, 2003). Such an approach underscores the necessity to comprehend these relationships as multifaceted and complex.

Additionally, teachers' belief systems about L2 writing assessment were found to be heterogeneous and dependent on context. Despite teaching the same course within the same PYP setting, the five teachers maintained unique and varied belief systems about L2 writing assessment. These belief systems comprised diverse interacting core and peripheral beliefs across multiple facets of writing assessment, including writing instruction, assessment purposes, approaches, evaluation criteria, processes, and instruments. Some studies, although not explicitly focused on L2 writing assessment, have found that teacher EFL teachers' cognition is heterogeneous and encompasses various aspects of EFL instruction (Andrews, 2003; Gao & Zhou, 2021; Gatbonton, 2008; Johnston & Goettsch, 2000; Kiss, 2012; Li, 2013; Zheng, 2015). Other studies also found that teachers' beliefs or philosophies about L2 writing instruction and assessment consist of a range of beliefs that are distinct from each other and related to different aspects of L2 writing assessment (Alzaanin, 2019; Crusan et al., 2016; Lam, 2019; Mao & Crosthwaite, 2019; Yu et al., 2020).

The participating teachers articulated their beliefs with clarity, but their envisioned practices were nebulous; they struggled to define what they genuinely intended to implement.

Within the PYP context, when teachers described their beliefs in practice, some conveyed their actual assessment practices without being able to clearly link them to their stated beliefs, even when prompted to outline their ideal writing classes in any other context. This disconnect might stem from a deficiency in assessment literacy, coupled with the influence of their teaching experiences and deep immersion in the PYP framework, which shaped their understanding of writing assessment. Some research in Saudi higher education has also highlighted that EFL teachers' exclusion from curriculum development processes has contributed to a decline in their assessment literacy, underscoring the need for systematic professional development (Al-Seghayer, 2015; Al-Jarf, 2011; Ghalib & Al-Hattami, 2015; Obeid, 2017; Salami & Alharthi, 2022).

This may also suggest that the interplay between teachers' beliefs and practices is intricately reciprocal. The habitual assessment practices of teachers within the PYP context appeared to mold their writing assessment beliefs, especially their beliefs in practice. Even though the participating teachers critiqued the English II curriculum and assessment strategies, they failed to propose specific alternative writing assessment methods beyond general reference to PYP assessment processes and methods, such as rubrics, MCQs, or direct writing assessment tasks. The observation that practices can influence beliefs and that their relationship is not unidirectional aligns with scholars' view that practices can impact and inform teachers' beliefs (Basturkmen, 2012; Borg, 2018; Li, 2013).

In analyzing the content of the participating teachers' writing assessment beliefs, the study discovered significant variations in core beliefs about their definitions of good writing or effective writing instruction and assessment. These differences largely stem from the distinct learning experiences of each teacher, which act as influential microsystemic factors. Specifically,

teachers' personal histories—such as their previous education and training—shaped their core beliefs about the acquisition, teaching, and evaluation of L2 writing. These core beliefs, deeply rooted within their broader belief system, played a pivotal role in aligning their assessment practices with their beliefs and informing their peripheral assessment beliefs. In other words, when institutional constraints were minimal, such as in classroom assessments, teachers tended to implement writing assessment practices that aligned with their core beliefs.

Some scholars (e.g., Holt-Reynolds, 1992) have emphasized that teachers' beliefs about learning and teaching are molded by their learning experiences, whether positive or negative. For instance, emotionally charged incidents from their past can shape teachers' approaches to writing instruction and assessment. This notion was illustrated by Rana, who, due to a significant incident during her undergraduate studies in a Saudi PYP, placed considerable emphasis on students' psychological well-being and readiness during class writing instruction and assessment. Similarly, Kiss (2012), Zheng (2015), and Finch (2010) found that teachers form their educational beliefs through complex networks of past experiences, which are crucial for assimilating new knowledge and shaping teaching practices.

In the domain of assessment purposes, Brown (2004) recognized four primary assessment purposes: learning improvement, student accountability, school accountability, and irrelevance. In the current study, teachers emphasized one formative purpose, specifically improving student writing skills. Yet, their expressions of this belief varied. Some linked writing assessments to improving skills via specific methods like direct assessment tasks, while others associated it with the assessment process, such as the role of feedback. Aya and Noor explicitly stated the purpose of writing assessment: Aya focused on individualized learning, and Noor focused on language proficiency enhancement. This finding aligns with Brown et al. (2019), who found that there is a

‘global localism’ in which teachers’ assessment purpose beliefs tend to be consistent within cultures due to local contextual influences.

Likewise, the participating teachers agreed that assessment methods should prioritize formative and direct writing tasks over summative and indirect ones, such as MCQs in quizzes and exams, which the PYP assessment policies required. Research has shown that while language teachers generally value formative assessment methods, they often rely on summative assessment in practice due to various institutional constraints (Crusan et al., 2016; Harris & Brown, 2019; Karp & Woods, 2008; Lam, 2019; Wang et al., 2020). In the current study, teachers voiced a preference for formative assessment approaches; however, there was a noticeable emphasis on summative methods in their actual practices, driven by a focus on grades and test preparation—a requirement of the PYP, whether through its assessment policies or students’ focus on grades due to the program’s competitive environment. In Saudi PYPs, Alshakhi (2019) and Obeid (2017) observed that EFL writing teachers adopt the PYP-recommended assessment practices, such as utilizing criterion-referenced, often in response to the stringent PYP assessment policies.

The participating teachers held diverse beliefs concerning other aspects of writing assessment—including evaluation criteria, processes, and tools. For example, Leena emphasized the importance of grammatical accuracy, whereas Aya and Noor prioritized meaning; Rana and Batool advocated for a holistic evaluation of all errors following PYP curriculum requirements. Furthermore, while all teachers recognized the significance of feedback for students’ development, the forms and approaches of feedback, they preferred using, varied. Leena and Noor were proponents of written feedback, Aya and Rana advocated one-on-one verbal feedback, and Batool supported combining both approaches. Most teachers endorsed the idea

that indirect feedback could enable students to identify and amend their errors autonomously. However, Noor and Aya contended that direct correction was preferable to prevent students from losing grades for minor mechanical and grammatical mistakes. Notably, Noor and Aya focused their evaluation on content-related criteria. Their stance underscored the intricacy of their belief systems, which seemed to be self-organized to promote reasoned and logical decision-making. Specifically, since they did not emphasize spelling or grammar, they preferred to prevent grade reduction for what they considered minor errors. In contrast, they believed that providing corrections for content-related errors could hinder students' ability to develop their capacity to construct meaning in their writing.

The study findings highlighted that the interaction between teachers' beliefs and their actual practices is complex, characterized by a blend of harmony and tension. CT offers a lens that reveals how these relationships are influenced by various factors. For instance, a policy requiring written feedback on students' drafts revealed the simultaneous tension and harmony between teachers' beliefs and their practices. Teachers' practices often aligned with their beliefs through verbal and indirect feedback methods, yet they complied with policy demands for written feedback, sometimes contrary to their beliefs. This echoes prior research findings, indicating that while there is consistency in teachers' feedback approaches, significant misalignments due to contextual constraints are prevalent (Al-Bakri, 2016; Ferris, 2014; Lee, 2009; Mao & Crosthwaite, 2019). Research in Saudi universities further corroborates this, showing teachers' compliance with institutional guidelines on written feedback despite personal belief conflicts (Alkhatib, 2015; Alshahrani & Storch, 2014). Additionally, while beliefs about grading and rubric use varied—with some teachers favoring teacher-judgment—their practices

conformed to the PYP assessment policies that mandated the use of numerical grades and a specific rubric.

Such factors at the exosystem level, particularly institutional factors, exerted the most significant influence on the misalignment between teachers' beliefs and practices. Restrictive institutional factors—such as limited teacher autonomy, curricular limitations, inflexible assessment policies, students' low proficiency levels, and time restrictions — impeded the harmony between teachers' beliefs and practices. Research within the context of PYP in Saudi universities indicates that such constraints, compounded by national educational policies that marginalize teachers' input in curriculum and policymaking, pose substantial challenges for teachers (Alharbi, 2017; Alharbi & Albelihi, 2023; Alshakhi, 2019; Al-Seghayer, 2014; Obeid, 2017; Alsharani, 2015). Alharbi and Albelihi (2023) highlighted that EFL writing teachers in KSA universities face difficulties related to curricular limitations and students' low proficiency stemming from inadequate pre-university writing instruction. Alsharani (2015) identified additional challenges for English teachers in Saudi Arabia, such as students' low motivation, language anxiety, and a highly centralized administrative system that restricts teachers' pedagogical freedom and ability to innovate.

Meanwhile, elements at the macrosystem level, such as the national language education system or cultural perspectives, seemed to exert minimal influence on teachers' beliefs and practices regarding L2 writing assessment. Despite any concerns about ELT within the Saudi educational system, the participating teachers did not perceive a direct impact of these factors on their L2 writing assessment practices. This finding contrasts with other research indicating that macrosystem factors do influence teacher beliefs (e.g., Wang et al., 2020; Zheng, 2015). Specifically, Wang et al. (2020) discovered that, at the macro level, Chinese teachers' belief in

the educational values of AfL conflicted with the prevailing high-stakes testing culture. This conflict forced them to prioritize scores over providing solely comment-based feedback in student assessments. In this study, the minimal influence of macrosystem factors may be due to teachers' perceptions that ELT in higher education is distinct from ELT in general school education, as suggested by Leena and Batool. Teachers appeared to disregard the potential impact of national language education policies on their L2 writing assessment beliefs and practices, focusing instead on institutional factors more immediately relevant to their teaching contexts.

11.2 Limitations of the Study

This study, although comprehensive in its exploration of EFL teachers' L2 writing assessment beliefs and practices within a specific program at a Saudi university, is not without its limitations.

Firstly, this study focuses only on five teachers within a single program. Qualitative research methodologies, particularly case studies, prioritize depth over breadth, offering rich, detailed descriptions that highlight the uniqueness of the observed phenomena. I adhered to this qualitative tradition by providing thorough methodological information and dense descriptions of the settings, participants, and processes involved. Such comprehensive detail may help readers determine the potential transferability of the study findings to their contexts (Creswell & Poth, 2018; Lincoln & Guba, 1985).

Secondly, a few classroom observations were conducted, focusing solely on English II during the second trimester. According to the PYP curriculum structure, only one writing course is offered each trimester, which inherently limited the study to the only available course during the period when data were collected. This limitation, however, mirrors the linear progression of

the writing curriculum and the genuine framework of the program. In addition, the observation timeframe was strategically planned to align with the study's need to delve deeply into a representative sample of assessment practices at a critical point in the curriculum, which was focusing on in-depth insights into the formative assessment component of the curriculum (the paragraph writing tasks). As a result, each participant was observed twice, except for two teachers who were observed once due to their overlapping and busy schedules. Despite this, the duration of most teachers' observations extended to two hours or more, providing substantial depth and detail in each session to compensate for the fewer occurrences. This duration was believed to be significant enough to allow for a comprehensive understanding of the teaching and assessment practices within the periods observed. Furthermore, both constraints were mitigated by the study's transparent methodology and the rich context it provides, allowing for careful extrapolation to similar educational contexts.

Thirdly, as discussed in Chapter 4, cultural and practical challenges affected the study's ability to use video for recording TAPs and classroom observations and include teachers and students from the male section. Saudi cultural norms around video recording in educational contexts and gender segregation in the region required methodological adjustments. Practical issues, such as securing permission to include male teachers and obtaining female teachers' consent for video recording in a setting unaccustomed to such practices, limited the depth of data analysis. Despite these limitations, audio recordings complemented by field notes were utilized. The audio recorder was strategically placed to ensure sound clarity for follow-up interviews and data analysis. These mentioned limitations, while inherent in the study's design, also offer avenues for future research. The following section discusses future directions for teacher cognition research.

Fourthly, the study's qualitative methodology focused on examining teachers' beliefs through self-reported data (i.e., initial interviews, TAPs, and follow-up interviews for TAPs and observations), complemented by observations and document analysis to measure actual writing assessment practices over one trimester. A limitation encountered was the inability to conduct stimulated recall interviews instead of follow-up interviews, owing to the constrained timeframe of the study and the teachers' demanding, often overlapping schedules. This limitation was, however, partly mitigated by scheduling the follow-up interviews immediately after the TAP sessions and sometimes on the same day as the observations. Additionally, augmenting the audio-recorded TAPs and observations with field notes facilitated my reflection and the formation of questions that assisted teachers in recalling and discussing their practices during the follow-up interviews.

Finally, applying CT introduced certain challenges in data analysis. CT is a framework for understanding complex interactions that are shaped by context. It aims to account for the varied nature of each interaction, whether it be harmony, tension, or influence from external or internal sources. The task of capturing the multiple levels of interactions among three distinct complex systems (i.e., belief, practice, and ecological systems) was daunting due to the sheer volume of interactions, making inclusive discussion nearly impossible. Nevertheless, the study overcame this obstacle primarily in two ways. As specified in Chapter 4, it adhered to the guidelines set by Larsen-Freeman and Cameron (2008) for establishing analytical boundaries when employing CT, such as adopting an ecological perspective of context to streamline the categorization of contextual factors and examination of their interactions. Also, the analysis specifically focused on the interactions pertaining solely to L2 writing assessment, employing CT concepts metaphorically and using dichotomies like 'harmony' and 'tension' to better

understand the realities of the teachers' cases. Secondly, the final findings centered more on identifying patterns of interactions within and across cases rather than the intricacies of each individual case, aiming to derive implications that could aid in enhancing L2 writing assessment and developing professional development courses for teachers.

11.3 Implications

Exploring EFL teachers' beliefs and practices regarding writing assessment at a Saudi university reveals insights for policymakers and curriculum developers. The study uncovered varied beliefs about writing assessment and the impact of ecological factors on these beliefs and practices, illustrating the complexities of teaching and assessing writing. It highlights demands placed by assessment on EFL teachers and suggests implications for those in similar educational settings. The subsequent section will detail these implications.

11.3.1 Implications for Teachers

The findings of this study illuminate several important considerations for teachers. EFL teachers within the PYP were aware of their personal beliefs concerning writing assessment, yet they frequently faced challenges when attempting to translate these beliefs into practical assessment procedures. For instance, this dissonance was exemplified by their inclination to move beyond the use of rubrics despite lacking a definitive alternative to adopt.

The results underscore the importance of reflective practice among teachers, a process that involves pinpointing both core and peripheral beliefs and understanding how these impact their instructional and assessment practices. Furthermore, teachers need to cultivate assessment literacy, specifically concerning alternative approaches, and to scrutinize existing L2 writing assessment techniques critically. Reflective practices should also encompass comprehending the practicalities of varying assessment strategies and the ways in which they can be applied

effectively within the unique confines of their teaching context, considering the specific challenges and limitations they may encounter. This could involve integrating systematic reflection exercises into teacher meetings, enabling teachers to critically evaluate and verbalize the amendments they aspire to implement in their assessment routines. Possessing a coherent vision and comprehensive knowledge of L2 writing instruction and assessment methods can fortify teachers' advocacy for modifications within the PYP writing curriculum. For instance, by engaging with contemporary research on efficacious writing assessment strategies, teachers can refine their instructional practices and discover practical ways to align their beliefs with their contextual constraints.

Moreover, acknowledging the role of educational experiences in shaping teachers' core beliefs underscores the necessity for continuous professional development. This development is essential in narrowing the gap between theoretical knowledge and its real-world application in the writing classroom and within institutional frameworks. There is a need for specialized professional development programs focused on expanding teachers' knowledge and skills in L2 writing assessment. These programs should provide a comprehensive array of assessment methods and strategies to augment teachers' existing approaches. Furthermore, custom training initiatives could offer teachers valuable insights into their assessment beliefs and how these beliefs interact with their instructional methods. Recognizing the impact of contextual factors, such as institutional policies and personal experiences, is critical in understanding the formation and application of these beliefs and practices. This understanding has meaningful implications for language teacher education, including the integration of assessment literacy into pre-service programs, the development of strategies for contextualizing writing and assessment practices despite typical constraints, and the establishment of communities of practice.

11.3.2 Implications for Teacher Education

Building on the implications for teachers, the findings of this study emphasize several key considerations for teacher education, both for pre-service and in-service teachers. To effectively address the challenges teachers face in aligning their beliefs with practical assessment procedures, teacher education programs must prioritize the development of reflective practices and assessment literacy.

For pre-service teacher education, it is crucial to embed modules that focus on the importance of reflective practice. Trainee teachers need to be encouraged to identify and reflect on their core and peripheral beliefs about writing assessment through structured reflection exercises and peer discussions. This reflective process may help future teachers understand how their beliefs influence their instructional practices, enabling them to navigate the complexities of assessment in real classroom settings more effectively.

Enhancing assessment literacy should also be a significant component of pre-service programs. Comprehensive training on various assessment methods, including alternative approaches beyond traditional rubrics, should be provided. Workshops and courses should cover contemporary research on effective writing assessment strategies, teaching trainees to critically evaluate these methods and adapt them to their specific teaching contexts. Practical application opportunities, such as teaching practicums or simulated classroom activities, are essential for helping future teachers gain confidence in implementing diverse assessment techniques.

Pre-service teacher education programs must also prepare them to understand and navigate the influence of institutional policies and personal experiences on their beliefs and practices. Discussions on the impact of these factors and strategies for contextualizing writing

and assessment practices within different educational settings may equip future teachers with the skills needed to adapt effectively.

For in-service teacher education, continuous professional development should be a cornerstone. Specialized development programs focused on L2 writing assessment can bridge the gap between theoretical knowledge and practical application. These programs should offer in-depth training on various assessment methods and strategies, helping teachers expand their existing approaches and refine their practices. Custom training initiatives can provide valuable insights into the interaction between teachers' assessment beliefs and their instructional methods.

In addition, in-service teacher education programs should integrate systematic reflection exercises into regular teacher meetings. These exercises can enable teachers to critically evaluate and articulate the changes they wish to implement in their assessment routines. Reflective practices should encompass understanding the practicalities of different assessment strategies and how they can be effectively applied within the unique curricula constraints of their teaching context.

Establishing communities of practice within in-service teacher education programs can provide ongoing support and a platform for sharing best practices and challenges. These communities can facilitate the continuous exchange of ideas and experiences, helping teachers stay updated on contemporary research and innovative assessment strategies. Moreover, possessing a coherent vision and comprehensive knowledge of L2 writing instruction and assessment methods may empower in-service teachers to advocate for modifications within the PYP writing curriculum. Engaging with contemporary research on effective writing assessment strategies may enable teachers to refine their instructional practices and find practical ways to align their beliefs with their contextual constraints.

11.3.3 Implications for Policy Makers and Curriculum Developers

Recognizing the ongoing call to bridge the gap between educational policy and classroom practice, the evidence gathered from this study serves as a catalyst for policymakers and institutions that may value teachers' perspectives and be receptive to their voices. It offers valuable insights for those willing to align institutional goals with the nuanced realities of classroom teaching. While teachers' aspirations for their students' learning and language development often align with educational ideals, the practicalities and pressures of policy decisions can sometimes diverge from these objectives. By considering the voices of teachers, policymakers and curriculum developers can design more responsive and effective educational strategies that genuinely support the multifaceted aims of language teaching.

Given that assessments are most effective when crafted by teachers to resonate with the instructional context and meet the diverse needs of students, as posited by Calfee and Miller (2013), there is a critical necessity for policy and curriculum design processes to incorporate the voices of teachers actively. Such reforms should prioritize teacher autonomy and pedagogical integrity, moving away from rigid institutional frameworks to support the multifaceted aims of teaching.

In the realm of effective L2 writing instruction and assessment reform, it is recommended for policymakers and curriculum developers to understand the assessment beliefs of teachers fully. The study has highlighted challenges such as dense syllabi and constrained timelines, which prevented teachers from providing meaningful feedback and offering additional writing practice opportunities. This underscores the urgent need for curricula that consider teachers' insights and their concerns about students' varying levels and capabilities across different sections, like the unified and nursing sections. A more inclusive writing curriculum that involves

teachers in selecting topics, genres, and assessment methods could introduce a range of new assessment techniques, better addressing students' diverse linguistic abilities and learning goals and needs.

Furthermore, the implications of these tensions are not merely academic; they have real and profound effects on teachers' well-being and job satisfaction. The constant balancing act required to align personal assessment beliefs with institutional demands often leads to professional dissatisfaction and burnout, as Rana's case exemplifies. Consequently, there is a pressing need for more flexible, teacher-focused assessment policies.

Moreover, institutions and policymakers, who may acknowledge the pivotal role teachers play in the assessment process, can enhance teachers' understanding of assessment through professional development programs. Smith and Abouammoh (2013) noted that most universities in KSA typically lack specific policies or practices aimed at the professional development of faculty around student assessment. The PYPs in Saudi universities can cultivate teacher knowledge by organizing programs that address the core aims of teaching writing within PYPs and elucidate the objectives of the assessment methods employed. Additionally, these programs could encourage teachers to engage in action research and design formative assessments grounded in empirical data and practical experience. Such professional development efforts are crucial in equipping educators to develop assessments that genuinely meet student needs, ultimately improving L2 writing teaching effectiveness and student achievement.

To ensure that the insights and recommendations from this study are effectively communicated and utilized, a strategic plan for engagement with policymakers and curriculum developers can be proposed. The first step involves developing and distributing a white paper that summarizes this study's outcomes and recommendations. This document will serve as a

formal proposal to the administration within the specific PYP, advocating for necessary changes in policy and curriculum based on the study's evidence. For example, initial changes could include proposing new assessment methods, such as moving away from MCQs, which teachers believe do not help students develop their writing skills. More significant changes might involve including teachers in the Curriculum Development Committees to contribute to the development of writing course curricula and share insights about student needs in PYPs.

The second step could involve organizing a series of workshops and seminars at the institution where the findings of the study could be presented in detail. These events will not only share the results but also may initiate a dialogue between teachers, policymakers, and curriculum developers to discuss practical implications. This approach would be feasible as such institutions typically organize yearly symposiums and seminars to discuss and present research.

The ultimate step could involve planning to create a task force comprising representatives from both the teaching staff and the policymaking body. This group could be tasked with reviewing current policies and curriculum structures and proposing modifications to better align them with the study's findings. The task force may also monitor the implementation of these changes to ensure they effectively address the challenges identified.

11.3.4 Implications for Future Research

The limitations identified in this study pave the way for several recommendations for future research. Firstly, employing CT necessitates recognizing the diversity within CAS, such as teachers' belief systems. These belief systems comprise numerous interactive components, including knowledge, beliefs, beliefs in practice, and emotions. Moreover, these components vary in their centrality to the belief system, being either core or peripheral. A study dedicated to examining the complex interactions among these components and their prominence within the

system could lead to the development of more nuanced teacher professional development strategies. These strategies can consider teachers' distinct personalities and individualities, reflecting their specific modes of understanding and learning circumstances. For instance, future studies could investigate how emotions influence teachers' beliefs and their approaches to writing instruction and assessment, which could be particularly insightful for professional development purposes. This approach aligns with the principles of CT, acknowledging that teacher development is not one-size-fits-all but should be adaptive and responsive to individual variations.

Additionally, teachers' L2 writing assessment belief systems were found to be sensitive to learning experiences. Conducting a longitudinal study coupled with CT could yield rich findings about the long-term dynamics and the reasons for potential changes in teachers' beliefs and practices over time. Future longitudinal studies could also focus on developing interventions or professional development programs that enhance teacher L2 writing assessment literacy and bridge the gap between teachers' beliefs and institutional constraints. This can be achieved by allowing teachers to reflect on their context and the purpose of the assessment methods within it to deepen their understanding and knowledge. Subsequently, the study could measure changes in participating teachers' beliefs and practices following these interventions.

Given the localized context of this study, future research could expand to include multiple educational settings across different geographical locations. Further studies should also consider a more diverse and larger sample of teachers to capture a broader range of perspectives and experiences. This would aid in understanding the variability in teaching practices and belief systems across different teachers and contexts, thereby providing a more comprehensive view of

EFL teachers' L2 writing assessment belief systems and practices in varied cultural and institutional contexts.

An additional and significant area for future inquiry is the exploration of how teachers' beliefs and practices in writing assessments affect student learning. A deeper understanding of this impact necessitates examining the nuances of how teachers' beliefs translate into their assessment practices in the classroom and subsequently affect students' writing development. For instance, a study might investigate how teachers' feedback beliefs and practices enhance students' writing. This line of research could also explore the role of formative assessment in promoting metacognition among learners, enabling them to become self-regulated writers. Additionally, such research could delve into how disparities between teachers' beliefs and actual assessment practices create cognitive dissonance that can affect student outcomes. By drawing correlations between assessment practices and student learning, educational stakeholders can design assessment frameworks that reflect teachers' beliefs and are demonstrably beneficial to student learning.

Finally, integrating quantitative methods with the qualitative approach would permit a broader examination of patterns and trends in teacher beliefs, practices, and ecological factors in different educational settings within KSA or other contexts. This mixed-methods approach would bolster the qualitative findings and provide a more detailed and holistic understanding of the complexities involved in L2 writing assessment practices. It could also pinpoint the factors that significantly impact teachers' practices within their contexts, offering deeper insights into broader educational reform concerning L2 writing instruction and assessment.

References

- Adler-Kassner, L., & O'Neill, P. (2010). *Reframing writing assessment to improve teaching and learning*. Utah State University Press. <https://doi.org/10.2307/j.ctt4cgrrtq>
- Ahmed, A. (2018). Assessment of EFL writing in some Arab world university contexts: Issues and challenges. In A. Ahmed & H. Abouabdelkader (Eds.), *Assessing EFL writing in the 21st century Arab world: Revealing the unknown* (pp. 1–19). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-64104-1_1
- Ajzen, I. (1988). *Attitudes, personality and behavior*. London: Open University Press.
- Alamri, M. (2011). Higher education in Saudi Arabia. *Journal of Higher Education Theory and Practice*, 11(4), 88–91.
- Al-Bakri, S. (2016). Written corrective feedback: Teachers' beliefs, practices and challenges in an Omani context. *Arab Journal of Applied Linguistics*, 1(1), 44–73. <https://arjals.com/ajal/article/view/44>
- Alharbi, E. A. (2016). Higher education in Saudi Arabia: Challenges to achieving world-class recognition. *International Journal of Culture and History*, 2(4), 169–172. <https://doi.org/10.18178/ijch.2016.2.4.058>
- Alharbi, N. S. (2017). *An investigation into the academic writing: Difficulties of Saudi postgraduate students*. [Doctoral dissertation, University of Exeter]. E-theses Online Service (EThOS).
- Alharbi, M. A., & Albelihi, H. H. (2023). Transforming writing education: WAC faculty experiences and challenges in Saudi universities. *Asian Journal of Second and Foreign Language Education*, 8(1), 28. <https://doi.org/10.1186/s40862-023-00199-0>

- Al-Hazmi, S. (2006). English and Arabization. Friends or foes? The Saudi experience. In A. Hashim & N. Hassan (Eds.) *English in Southeast Asia: Prospects, perspectives, and possibilities* (pp. 109–122). University of Malaya Press.
- Al-Hazmi, S. (2017). Current issues in English language education in the Kingdom of Saudi Arabia. *Journal of Modern Languages*, 17(1), 129–150.
<https://ejournal.um.edu.my/index.php/JML/article/view/3720>
- Al-Jarf, R. (2011). Creating and sharing writing iRubrics. *Asian EFL Journal*, 51(1), 41–62.
- Alkhatib, N. (2015). *Written corrective feedback at a Saudi university: English language teachers' beliefs, students' preferences, and teachers' practices*. (Publication No. 74770910) [Doctoral dissertation, University of Essex]. Semantic Scholar.
- Alkhazim, M. A. (2003). Higher education in Saudi Arabia: Challenges, solutions, and opportunities missed. *Higher Education Policy*, 16(4), 479–486.
<https://doi.org/10.1057/palgrave.hep.8300035>
- Alsamaani, A. (2014). Evaluating classroom assessment techniques of novice Saudi EFL teachers. *Journal of Arabic and Human Sciences*, 7(2), 1–40.
<https://doi.org/10.12816/0009602>
- Al-Seghayer, K. (2005). Teaching English in the Kingdom of Saudi Arabia: Slowly but steadily changing. In G. Braine (Ed.), *Teaching English to the world: History, curriculum, and practice* (pp. 125–134). Routledge. <https://doi-org.ezproxy.library.yorku.ca/10.4324/9781410612861>
- Al-Seghayer, K. (2014). The four most common constraints affecting English teaching in Saudi Arabia. *International Journal of English Linguistics*, 4(5), 17–26.
<http://dx.doi.org/10.5539/ijel.v4n5p17>

- Al-Seghayer, K. (2015). Salient key features of actual English instructional practices in Saudi Arabia. *English Language Teaching*, 8(6), 89–99. <https://doi.org/10.5539/elt.v8n6p89>
- Alshahrani, A.S., & Storch, N. (2014). Investigating teachers' written corrective feedback practices in a Saudi EFL context: How do they align with their beliefs, institutional guidelines, and students' preferences? *Australian Review of Applied Linguistics*, 37(2), 101–122. <https://doi.org/10.1075/araal.37.2.02als>
- Alshakhi, A. (2019). Revisiting the writing assessment process at a Saudi English Language Institute: Problems and solutions. *English Language Teaching*, 12(1), 176–185. <https://doi.org/10.5539/elt.v12n1p176>
- Alsharif, D. T. (2019, November 13). The new universities law in Saudi Arabia. *Arab News*. <https://www.arabnews.com/node/1583251>
- Al-Shehri, S. (2017). A developmental paradigm for English language instruction at preparatory year programs. *Arab World English Journal*, 8(3), 432–447. <https://dx.doi.org/10.24093/awej/vol8no3.28>
- Alzaanin, E. I. (2019). An exploratory study of the interplay between EFL writing teacher cognition and pedagogical practices in the Palestinian University context. *Arab World English Journal*, 10(3), 113–132. <https://dx.doi.org/10.24093/awej/vol10no3.8>
- Anderson, R. A., Crabtree, B. F., Steele, D. J., & McDaniel, R. R. (2005). Case study research: The view from complexity science. *Qualitative health research*, 15(5), 669–685. <https://doi.org/10.1177/1049732305275208>
- Andrews, S. (2003). Teacher language awareness and the professional knowledge base of the L2 teacher. *Language Awareness*, 12(2), 81–95. <https://doi.org/10.1080/09658410308667068>

- Ansari, A. A. (2012). Teaching of English to Arab students: Problems and remedies. *Educational Research*, 3(6), 519–524. <http://www.interestjournals.org/ER>
- Ashraf, T. A. (2018). Teaching English as a foreign language in Saudi Arabia: struggles and strategies. *International Journal of English Language Education*, 6(1), 133-154. <https://doi.org/10.5296/ijele.v6i1.13148>
- Azis, A. (2015). Conceptions and practices of assessment: A case of teachers representing improvement conception. *Teflin Journal*, 26(2), 129–154. <https://doi.org/10.15639/teflinjournal.v26i2/129-154>
- Bamanger, E. M., & Gashan, A. K. (2014). In-service EFL teachers' beliefs about teaching reading strategies. *English Language Teaching*, 7(8), 14–22. <https://doi.org/10.5539/elt.v7n8p14>
- Barkaoui, K. (2011). Think-aloud protocols in research on essay rating: An empirical study of their veridicality and reactivity. *Language Testing*, 28(1), 51–75. <https://doi.org/10.1177/0265532210376379>
- Barrot, J. S. (2016). Examining the teaching beliefs and practices of experienced ESL teachers: A sociocognitive-transformative perspective. *3L: Language, Linguistics, Literature*, 22(1), 153–163. <https://doi.org/10.17576/3L-2016-2201-12>
- Basturkmen, H. (2012). Review of research into the correspondence between language teachers' stated beliefs and practices. *System*, 40(2), 282–295. <https://doi.org/10.1016/j.system.2012.05.001>
- Bazerman, C., & Prior, P. (2004). *What writing does and how it does it: An introduction to analyzing texts and textual practices*. Lawrence Erlbaum.

- Behizadeh, N. & Engelhard, G. (2011). Historical view of the influences of measurement and writing theories on the practice of writing assessment in the United States. *Assessing writing*, 16(3), 189–211. <https://doi.org/10.1016/j.asw.2011.03.001>
- Ben Hedia, N. (2020). How writing teachers' beliefs influence grading practices. In L. McCallum & C. Coombe (Eds.), *The assessment of L2 written English across the MENA Region* (pp. 143–170). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-53254-3_7
- Blatter, J. K. (2008). Case Study. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (pp. 68–71). SAGE. <https://doi.org/10.4135/9781412963909>
- Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology*, 39(3), 206–219. <https://doi.org/10.1016/j.cedpsych.2014.05.003>
- Bonner, S. M. (2016). Teachers' perceptions about assessment: Competing narratives. In G. T. Brown & L. R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment* (1st ed., pp. 21–39). Routledge. <https://doi.org/10.4324/9781315749136>
- Borg, M. (2005). A case study of the development in pedagogic thinking of a pre-service teacher. *TESL-EJ*, 9(2), 1–30. <https://tesl-ej.org/wordpress/issues/volume9/ej34/ej34a5/>
- Borg, S. (1998). Teachers' pedagogical systems and grammar teaching: A qualitative study. *TESOL Quarterly*, 32(1), 9–38. <https://doi.org/10.2307/3587900>
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81–109. <https://doi.org/10.1017/S0261444803001903>
- Borg, S. (2006). *Teacher cognition and language education: Research and practice*. Continuum.

- Borg, S. (2010). Language teacher research engagement. *Language Teaching*, 43(4), 391–429.
<https://doi.org/10.1017/S0261444810000170>
- Borg, S. (2011). The impact of in-service education on language teachers' beliefs. *System*, 39(3), 370–380. <https://doi.org/10.1016/j.system.2011.07.009>
- Borg, S. (2012). Current approaches to language teacher cognition research: A methodological analysis. In R. Barnard & A. Burns (Eds.), *Researching language teacher cognition and practice: International case studies* (pp. 11–29). Multilingual Matters.
- Borg, S. (2018). Teacher's beliefs and classroom practices. In P. Garret & G. M. Cots (Eds.), *The Routledge handbook of language awareness* (pp. 75–91). Routledge.
<https://doi.org/10.4324/9781315676494-5>
- Borg, S. (2019). Language teacher cognition: Perspectives and debates. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. 1149–1170). Springer International Publishing.
- Borg, S., & Alshumaimeri, Y. (2019). Language learner autonomy in a tertiary context: Teachers' beliefs and practices. *Language Teaching Research*, 23(1), 9–38.
<https://doi.org/10.1177/1362168817725759>
- Borg, S., & Sanchez, H. S. (2020). Cognition and good language teachers. In C. Griffiths & Z. Tajeddin (Eds.), *Lessons from good language teachers* (pp. 16–27). Cambridge University Press. <https://doi.org/10.1017/9781108774390.005>
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>

- Breen, M. P., Hird, B., Milton, M., Oliver, R., & Thwaite, A. (2001). Making sense of language teaching: Teachers' principles and classroom practices. *Applied linguistics*, 22(4), 470–501. <https://doi.org/10.1093/applin/22.4.470>
- Broad, B. (2003). *What we really value: Beyond rubrics in teaching and assessing writing*. Utah State University Press.
- Bronfenbrenner, U. (1979). *The ecology of human development*. Harvard University Press.
- Bronfenbrenner, U. (2005). *Making human beings human: Bioecological perspectives on human development*. SAGE.
- Brown, G. T. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, 11(3), 301–318. <https://doi.org/10.1080/0969594042000304609>
- Brown, G. T., Gebril, A., & Michaelides, M. P. (2019). Teachers' conceptions of assessment: A global phenomenon or a global localism. *Frontiers in Education*, 4(16), 1–13. <https://www.frontiersin.org/articles/10.3389/feduc.2019.00016>
- Brown, G. T., Harris, L. R., & Harnett, J. (2012). Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education*, 28(7), 968–978. <https://doi.org/10.1016/j.tate.2012.05.003>
- Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy*. Longman Inc.
- Brown, H. D. (2006). *Principles of language learning and teaching* (5th ed.). Pearson Longman.

- Buehl, M. M., & Beck, J. S. (2014). The relationship between teachers' beliefs and teachers' practices. In H. Fives & M. Gregoire Gill (Eds.), *Handbook of research on teacher's beliefs* (pp. 78–96). Routledge. <https://doi.org/10.4324/9780203108437-11>
- Bustamante, A.P., & Yilmaz, S. (2020). Exploring the essay rating judgements of English instructors in the Middle East. In L. McCallum & C. Coombe (Eds.), *The assessment of L2 written English across the MENA Region* (113–141). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-53254-3_6
- Burns, A., & Knox, J. S. (2011). Classrooms as complex adaptive systems: A relational model. *TESL–EJ*, 15(1), 1–25. <https://tesl-ej.org/wordpress/issues/volume15/ej57/ej57a1/>
- Burns, A., Freeman, D., & Edwards, E. (2015). Theorizing and studying the language-teaching mind: Mapping research on language teacher cognition. *Modern Language Journal*, 99(3), 585–601. <https://doi.org/10.1111/modl.12245>
- Burri, M., Chen, H., & Baker, A. (2017). Joint development of teacher cognition and identity through learning to teach L2 pronunciation. *The Modern Language Journal*, 101(1), 128–142. <https://doi.org/10.1111/modl.12388>
- Busch, D. (2010). Pre-service teacher beliefs about language learning: The second language acquisition course as an agent for change. *Language Teaching Research*, 14(3), 318–337. <https://doi.org/10.1177/1362168810365239>
- Calfee, R. C., & Miller, R. G. (2013). Best practices in writing assessment for instruction. In S. Graham, C. MacArthur, & J. Fitzgerald (Eds.), *Best practices in writing instruction* (2nd ed., pp. 351–380). Guilford.
- Cameron, B. (1995). Systems for philosophy. In K. Ellis, A. Gregory, B. R. Mears-Young, & G. Ragsdell (Eds.), *Critical issues in systems theory and practice* (pp. 227–232). Springer.

- Cameron, L. (1999). Operationalising 'metaphor' for applied linguistic research. In L. Cameron & G. Low (Eds.), *Researching and applying metaphor* (pp. 3–28). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524704.004>
- Cameron, L., & Larsen-Freeman, D. (2007). Preview article: Complex systems and applied linguistics. *International Journal of Applied Linguistics*, 17(2), 226–240. <https://doi.org/10.1111/j.1473-4192.2007.00148.x>
- Camp, R. (1993). Changing the model for the direct assessment of writing. In M. Williamson & B. Huot (Eds.), *Validating holistic scoring for writing assessment: Theoretical and empirical foundations*, (pp. 45-79). Hampton Press.
- Capra, F. (2013). Deep ecology: Educational possibilities for the twenty-first century. *North American Montessori Teachers' Association (NAMTA) Journal*, 38(1), 201–216.
- Charters, E. (2003). The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education*, 12(2), 68–82. <https://doi.org/10.26522/brocked.v12i2.38>
- Cheng, X., Zhang, L. J., & Yan, Q. (2021). Exploring teacher written feedback in EFL writing classrooms: Beliefs and practices in interaction. *Language Teaching Research*. LTR, 136216882110576–. <https://doi.org/10.1177/13621688211057665>
- Calderhead, J. (1996). Teachers: Beliefs and knowledge. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 709–725). Macmillan Library Reference: Prentice Hall International.
- Clandinin, D. J., & Connelly, F. M. (1988). Studying teachers' knowledge of classrooms: Collaborative research, ethics, and the negotiation of narrative. *The Journal of*

Educational Thought (JET)/Revue de la Pensée Educative, 22(2A), 269-282.

<https://www.jstor.org/stable/23768371>

Clark, C. M. & Leat, S. (1998). The use of unstructured observation in teacher assessment. In C. Tilstone (Ed.), *Observing teaching and learning: Principles and practice* (1st ed., pp. 71–87). David Fulton Publishers.

Clark, C. M. & Peterson, P. L. (1986). Teachers' thought process. In M. Wittrock (Ed.), *Handbook of research on teaching*, (3rd ed., pp. 255–296). Macmillan.

Clark, C. M., & Yinger, R. J. (1977). Research on teacher thinking. *Curriculum Inquiry*, 7(4), 279–394. <https://doi.org/10.2307/1179499>

Cochran-Smith, M., Ell, F., Ludlow, L., Grudnoff, L., & Aitken, G. (2014). The challenge and promise of complexity theory for teacher education research. *Teachers College Record*, 116(4), 1–38. <https://doi.org/10.1177/016146811411600407>

Cohen, L., Manion, L., & Morison, K. (2007). *Research methods in education* (6th ed.). Routledge.

Cohen, D. & Crabtree, B. (2006, July). *Qualitative Research Guidelines Project*. The Robert Wood Johnson Foundation.

Couper, G. (2017). Teacher cognition of pronunciation teaching: Teachers' concerns and issues. *Tesol Quarterly*, 51(4), 820–843. <https://doi.org/10.1002/tesq.354>

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). SAGE.

Creswell, J., & Poth, C. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.

- Crusan, D. (2010). *Assessment in the second language writing classroom*. The University of Michigan Press.
- Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing writing*, 28(1), 43–56. <https://doi.org/10.1016/j.asw.2016.03.001>
- Cumming, A. (2009). Language assessment in education: Tests, curricula, and teaching. *Annual Review of Applied Linguistics*, 29(1), 90–100. <https://doi.org/10.1017/S0267190509090084>
- Cumming, A., Kantor, R. & Powers, D. (2002) Decision making while rating ESL/EFL writing tasks: A descriptive framework. *Modern Language Journal*, 86(1), 67–96. <https://doi.org/10.1111/1540-4781.00137>
- Cundale, N. (2001). Do we practice what we preach? Stated beliefs about communicative language teaching and classroom questioning strategies. *Japan Association for Language Teaching*, 25(5), 4–9. <https://doi.org/10.1177/00336882177165>
- Davis, B. & Sumara, D. (1997). Cognition, complexity and teacher education. *Harvard Educational Review*. 67(1), 105–127. <https://doi.org/10.17763/haer.67.1.160w00j113t78042>
- Davis, B., & Sumara, D. (2005). Challenging images of knowing: Complexity science and educational research. *International Journal of Qualitative Studies in Education*, 18(3), 305–321. <https://doi.org/10.1080/09518390500082293>
- Davis, B., & Sumara, D. (2006). *Complexity and education: Inquiries into learning, teaching, and research*. Erlbaum.

- Davis, B., & Sumara, D. (2007). Complexity science and education: Reconceptualizing the teacher's role in learning. *Interchange*, 38(1), 53–67. <https://doi.org/10.1007/s10780-007-9012-5>
- Davis, D. S. & Neitzel, C. (2011). A self-regulated learning perspective on middle grades classroom assessment. *The journal of educational research*, 104(3), 202–215. <https://doi.org/10.1080/00220671003690148>
- Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21(3), 305–334. <https://doi.org/10.1191/0265532204lt286oa>
- de Bot, K. (2008). Introduction: Second language development as a dynamic process. *The Modern Language Journal*, 92(2), 166–178. <https://doi.org/10.1111/j.1540-4781.2008.00712.x>
- de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, 10(1), 7–21. <https://doi.org/10.1017/S1366728906002732>
- DeRemer, M. (1998). Writing assessment: Raters' elaboration of the rating task. *Assessing Writing*, 5(1), 7–29. [https://doi.org/10.1016/S1075-2935\(99\)80003-8](https://doi.org/10.1016/S1075-2935(99)80003-8)
- Donaghue, H. (2003). An instrument to elicit teachers' beliefs and assumptions. *ELT journal*, 57(4), 344–351. <https://doi-org.ezproxy.library.yorku.ca/10.1093/elt/57.4.344>
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

- Dörnyei, Z. (2014). Researching complex dynamic systems: 'Retrodictive qualitative modelling' in the language classroom. *Language Teaching*, 47(1), 80–91.
<https://doi.org/10.1017/S0261444811000516>
- Edwards, E. (2021). The ecological impact of action research on language teacher development: A review of the literature. *Educational Action Research*, 29(3), 396–413.
<https://doi.org/10.1080/09650792.2020.1718513>
- Elbaz, P. (1983). *Teacher thinking: A study of practical knowledge*. Croom Helm.
- Elliot, N. C. (2005). *On a scale: A social history of writing assessment in America*. Peter Lang.
- England, N. (2017). Developing an interpretation of collective beliefs in language teacher cognition research. *Tesol Quarterly*, 51(1), 229–238. <https://doi.org/10.1002/tesq.334>
- Erdosy, M. U. (2004). Exploring variability in judging writing ability in a second language: A study of four experienced raters of ESL compositions. *Educational Testing Service Research Report Series*, 2003(1), I–62. <https://doi.org/10.1002/j.2333-8504.2003.tb01909.x>
- Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. MIT Press.
- Ezza, E. Y. (2017). Criteria for assessing EFL writing at Majma'ah University. In S. Hidri, & C. Coombe (Eds.), *Evaluation in foreign language education in the Middle East and North Africa* (pp. 185–200). Springer International Publishing. https://doi.org/10.1007/978-3-319-43234-2_11
- Farrell, T. (2011). Exploring the professional role identities of experienced ESL teachers through reflective practice. *System (Linköping)*, 39(1), 54–62.
<https://doi.org/10.1016/j.system.2011.01.012>

- Feryok, A. (2010). Language teacher cognitions: Complex dynamic systems? *System (Linköping)*, 38(2), 272–279. <https://doi.org/10.1016/j.system.2010.02.001>
- Finch, A. (2010). Critical incidents and language learning: Sensitivity to initial conditions. *System*, 38(3), 422–431. <https://doi.org/10.1016/j.system.2010.05.004>
- Fives, H., & Buehl, M. (2012). Spring cleaning for the “messy” construct of teachers’ beliefs: What are they? Which have been examined? What can they tell us? In K.R. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook: Volume 2 Individual differences and cultural and contextual factors* (pp. 471–499). American Psychological Association. <https://doi.org/10.1037/13274-019>
- Flores, B. B. (2001). Bilingual education teachers' beliefs and their relation to self-reported practices. *Bilingual Research Journal*, 25(3), 275–299. <https://doi.org/10.1080/15235882.2001.10162795>
- Freeman, D. (2002). The hidden side of the work: Teacher knowledge and learning to teach. *Language Teaching*, 35(1), 1–13. <https://doi.org/10.1017/S0261444801001720>
- Freeman, D., & Johnson, K. E. (1998). Reconceptualizing the knowledge-base of language teacher education. *TESOL Quarterly*, 32(3), 397–417. <https://doi.org/10.2307/3588114>
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment*. Routledge.
- Fulmer, G. W., Lee, I. C., & Tan, K. H. (2015). Multi-level model of contextual factors and teachers’ assessment practices: An integrative review of research. *Assessment in Education: Principles, Policy & Practice*, 22(4), 475–494. <https://doi.org/10.1080/0969594X.2015.1017445>

- Gao, Y., & Zhou, Y. (2021). Exploring language teachers' beliefs about the medium of instruction and actual practices using complex dynamic system theory. *Frontiers in Education*, 6(1), Article 708031. <https://doi.org/10.3389/feduc.2021.708031>
- Gao, Q., and Liu, L. (2013). Investigating college English teacher beliefs in listening and their relationships with teaching practices. *Foreign Language World*, 155(02), 33–41.
- Gatbonton, E. (2008). Looking beyond teachers' classroom behaviour: Novice and experienced ESL teachers' pedagogical knowledge. *Language teaching research*, 12(2), 161–182. <https://doi.org/10.1177/1362168807086286>
- Ghalib, T. K., & Al-Hattami, A. A. (2015). Holistic versus analytic evaluation of EFL writing: A case study. *English Language Teaching*, 8(7), 225–236. <https://doi.org/10.5539/elt.v8n7p225>
- Gipps, C. (2008). Sociocultural perspectives on assessment. In G. Wells & G. Claxton (Eds.), *Learning for life in the 21st century: Sociocultural perspectives on the future of education* (pp. 73–83). Wiley-Blackwell.
- Gipps, C. V. (2012). *Beyond Testing: Towards a Theory of Educational Assessment*. Routledge.
- Gleick, J. (1987). *Chaos: Making a new science*. Penguin Books.
- Goertzel, B. (2006). *The Hidden Pattern: A Patternist Philosophy of Mind*. Brown Walker Press.
- Golombek, P., & Doran, M. (2014). Unifying cognition, emotion, and activity in language teacher professional development. *Teaching and Teacher Education*, 39(1), 102–111.
- Green, T. F. (1971). *The activities of teaching*. McGraw-Hill.
- Grossman, P. L., Wilson, S. M., & Shulman, L. S. (1989). Teachers of substance: Subject matter knowledge for teaching. In M. C. Reynolds (Ed.), *Knowledge base for the beginning teacher* (pp. 23-36). Pergamon.

- Guadu, Z. B., & Boersma, E. J. (2018). EFL Instructors' Beliefs and Practices of Formative Assessment in Teaching Writing. *Journal of Language Teaching and Research*, 9(1), 42–50. <https://doi.org/10.17507/jltr.0901.06>
- Guskey, T. R. (1986). Staff development and the process of teacher change. *Educational Researcher*, 15(5), 5–12. <https://doi.org/10.3102/0013189X015005005>
- Hamdan, A. (2013). An exploration into "private " higher education in Saudi Arabia: improving quality and accessibility? *The ACPET Journal for Private Higher Education*, 2(2), 33–44. <https://doi.org/10.4324/9780203796139-7>
- Hamp-Lyons, L. (2001). Fourth generation writing assessment. In T. Silva & P. K. Matsuda (Eds.), *On second language writing* (pp. 117–127). Lawrence Erlbaum.
- Hamp-Lyons, L. (2019). Reflecting on the past, embracing the future. *Assessing Writing*, 42 (Complete). <https://doi.org/10.1016/j.asw.2019.100423>
- Harlen, W. (2005). Trusting teachers' judgment: Research evidence of the reliability and validity of teachers' assessment used for summative purposes. *Research Papers in Education*, 20(1), 245–270. <https://doi.org/10.1080/02671520500193744>
- Harris, L. R., & Brown, G. T. (2009). The complexity of teachers' conceptions of assessment: Tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practice*, 16(3), 365–381. <https://doi.org/10.1080/09695940903319745>
- Hase, S., & Kenyon, C. (2013). The nature of learning. In S. Hase & C. Kenyon (Eds.), *Self-determined learning: Heutagogy in action* (pp. 19–35). Bloomsbury Academic.

- Haukås, Å. (2016). Teachers' beliefs about multilingualism and a multilingual pedagogical approach. *International Journal of Multilingualism*, 13(1), 1–18.
<https://doi.org/10.1080/14790718.2015.1041960>
- Hayes, J. R. (2000). A New Framework for Understanding Cognition and Affect in Writing. In J. R. Squire (Ed.), *Perspectives on writing: Research, theory, and practice* (pp.6-44). International Reading Association.
- Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford University Press.
- Herdina, P., & Jessner, U. (2002). *A dynamic model of multilingualism: Changing the psycholinguistic perspective*. Multilingual Matters Ltd.
- Heylighen, F. (2008). Complexity and self-organization. In M. J. Bates & M. N. Maack (Eds.), *Encyclopedia of Library and Information Sciences*. Taylor & Francis.
- Hilal, K. T. (2013). Between the fears and hopes for a different future for the nation-states: Scholarship programs in Saudi Arabia and United Arab Emirates from a public policy standpoint. *International Journal of Higher Education*, 2(2), 195–210.
<https://doi.org/10.5430/ijhe.v2n2p195>
- Hiver, P., Al-Hoorie, A. H., & Larsen-Freeman, D. (2021). Toward a transdisciplinary integration of research purposes and methods for complex dynamic systems theory: Beyond the quantitative-qualitative divide. *IRAL - International Review of Applied Linguistics in Language Teaching*, February. <https://doi.org/10.1515/iral-2021-0022>
- Hofman, B. H. & Seidel. (2015). Measuring teachers' beliefs: For what purposes? In H. M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 106–127). Routledge. <https://doi.org/10.4324/9780203108437-14>

- Hofstadler, N., Babic, S., Lämmerer, A., Mercer, S., & Oberdorfer, P. (2021). The ecology of CLIL teachers in Austria – An ecological perspective on CLIL teachers' wellbeing. *Innovation in Language Learning and Teaching*, 15(3), 218–232. <https://doi.org/10.1080/17501229.2020.1739050>
- Holland, H. (1998). *Emergence: From Chaos to Order*. Oxford University Press.
- Holliday, A. (2015). Qualitative research and analysis. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: A Practical resource* (pp. 1–25). Bloomsbury.
- Holliday, A. & Aboshiha, P. (2009). The denial of ideology in perceptions of 'nonnative speaker' teachers. *TESOL Quarterly*, 43(4), 669–689. [https://doi.org/10.1002/\(ISSN\)1545-7249](https://doi.org/10.1002/(ISSN)1545-7249)
- Holt-Reynolds, D. (1992). Personal history-based beliefs as relevant prior knowledge in course work. *American educational research journal*, 29(2), 325–349. <https://doi.org/10.3102/00028312029002325>
- Hoy, A. W., Davis, H., & Pape, S. J. (2006). Teacher Knowledge and Beliefs. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 715–737). Lawrence Erlbaum Associates Publishers.
- Humphry, S., & Hedsinger, S. (2019). Raters' perceptions of assessment criteria relevance. *Assessing Writing*, 41(July 2019), 1–13. <https://doi.org/10.1016/j.asw.2019.04.002>
- Huot, B. (1996). Toward a new theory of writing assessment. *College Composition and Communication*, 47(4), 549–566. <https://doi.org/10.2307/358601>
- Huot, B. (2002). Toward a new discourse of assessment for the college writing classroom. *College English*, 65(2), 163–180. <https://doi.org/10.2307/3250761>
- Huot, B., & O'Neill, P. (2009). *Assessing writing: A critical sourcebook*. Bedford/St. Martin's.

- Huot, B., O'Neill, P., & Moore, C. (2010). A Usable Past for Writing Assessment. *College English*, 72(5), 495–517. <http://www.jstor.org/stable/20749294>
- Hyland, K. (2003). *Second language writing*. Cambridge University Press.
- Hyland, K. (2004). *Genre and second language writing*. University of Michigan Press.
- Isaacs, T., Zara, C., Herbert, G., Coombs, S. J., & Smith, C. (2013). *Key concepts in educational assessment*. SAGE.
- Jackson, D. O., & Cho, M. (2018). Language teacher noticing: A socio-cognitive window on classroom realities. *Language Teaching Research*, 22(1), 29–46.
<https://doi.org/10.1177/1362168816663754>
- Jamal, S. (2021). *Saudi Arabia Education Report 2021*. Knight Frank Middle East Limited.
<https://www.knightfrank.com.sa/en/research/saudi-arabia-education-report-2021-7999.aspx>
- Jensen, D. (2008). Credibility. In L. M. Given (Ed.), *The SAGE encyclopedia of qualitative research methods* (pp. 138–139). SAGE. <https://doi.org/10.4135/9781412963909>
- Jessner, U. (2008). A DST model of multilingualism and the role of metalinguistic awareness. *The modern language journal*, 92(2), 270–283. <https://doi.org/10.1111/j.1540-4781.2008.00718.x>
- Johnson, K. E. (1992). The relationship between teachers' beliefs and practices during literacy instruction for non-native speakers of English. *Journal of Reading Behavior*, 24(1), 83–108. <https://doi.org/10.1080/10862969209547763>
- Johnson, K. E. (2006). The sociocultural turn and its challenges for second language teacher education. *TESOL Quarterly*, 40(1), 235–257. [https://doi.org/10.1002/\(ISSN\)1545-7249](https://doi.org/10.1002/(ISSN)1545-7249)

- Johnson, K. E. (2018). Studying language teacher cognition: Understanding and enacting theoretically consistent instructional practices. *Language Teaching Research*, 22(3), 259–263. <https://doi.org/10.1177/1362168818772197>
- Johnston, B., & Goettsch, K. (2000). In search of the knowledge base of language teaching: Explanations by experienced teachers. *Canadian Modern Language Review*, 56(3), 437–468. <https://doi.org/10.3138/cmlr.56.3.437>
- Jorgensen, R., Grootenboer, P., Niesche, R., & Lerman, S. (2010). Challenges for teacher education: The mismatch between beliefs and practice in remote indigenous contexts. *Asia-Pacific Journal of Teacher Education*, 38(2), 161–175. <https://doi.org/10.1080/13598661003677580>
- Junqueira, L., & Payant, C. (2015). “I just want to do it right, but it’s so hard”: A novice teacher’s written feedback beliefs and practices. *Journal of Second Language Writing*, 27(Mar), 19–36. <https://doi.org/10.1016/j.jslw.2014.11.001>
- Kabouha, R., & Elyas, T. (2015). Aligning teaching and assessment to course objectives: The case of preparatory year English program at King Abdulaziz University. *International Journal of Applied Linguistics and English Literature*, 4(5), 82–91. <http://dx.doi.org/10.7575/aiac.ijalel.v.4n.5p.82>
- Kagan, D. (1992). Implication of research on teacher belief. *Educational psychologist*, 27(1), 65–90. https://doi.org/10.1207/s15326985ep2701_6
- Kamberelis, G., & Dimitriadis, G. (2005). *On qualitative inquiry: Approaches to language and literacy research*. Teachers College Press. <https://doi-org.ezproxy.library.yorku.ca/10.1111/j.1467-9620.2005.00618.x>

- Kim, C., Kim, M. K., Lee, C., Spector, J. M., & DeMeester, K. (2013). Teacher beliefs and technology integration. *Teaching and teacher education*, 29(1), 76–85.
<https://doi.org/10.1016/j.tate.2012.08.005>
- Kimberly, J. (2017). Exploring Teachers' beliefs and their influence on grading practices. (Publication No. ED577639) [Doctoral dissertation, University of La Verne]. ERIC.
- Kiss, T. (2012). The complexity of teacher learning: Reflection as a complex dynamic system. *Journal of Interdisciplinary Research in Education (JIRE)*, 2(1), 17–35.
<https://doi.org/10.1080/13540602.2018.1533811>
- Kramersch, C. (2003). *Language Acquisition and language socialization: Ecological perspectives*. (2003). Bloomsbury Publishing.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Pergamon.
- Kubanyiova, M., & Feryok, A. (2015). Language Teacher Cognition in Applied Linguistics Research: Revisiting the Territory, Redrawing the Boundaries, Reclaiming the Relevance. *The Modern Language Journal (Boulder, Colo.)*, 99(3), 435–449.
<https://doi.org/10.1111/modl.12239>
- Kuusela, H., & Paul, P. (2000). A comparison of concurrent and retrospective verbal protocol analysis. *American journal of psychology*, 113(3), 387–404.
<https://doi.org/10.2307/1423365>
- Kvale, S. & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing* (2nd ed.). SAGE.
- Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong. *System*, 81(1), 78–89.
<https://doi.org/10.1016/j.system.2019.01.006>

- Lamie, J. M. (2004). Presenting a model of change. *Language teaching research*, 8(2), 115–142.
<https://doi.org/10.1191/1362168804lr137oa>
- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18(2), 141–165. <https://doi.org/10.1017/S0261444822000362>
- Larsen-Freeman, D. (2002). Language acquisition and language use from a chaos/perspective of complexity theory. In C. Kramsch (Ed.), *Language acquisition and language socialization* (pp. 33–46). Continuum.
- Larsen-Freeman, D. (2011). A complexity theory approach to second language development/acquisition. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 48–72). Routledge.
- Larsen-Freeman, D. (2012). Complex, dynamic systems: A new transdisciplinary theme for applied linguistics? *Language Teaching*, 45(2), 202–214.
<https://doi.org/10.1017/S0261444811000061>
- Larsen-Freeman, D. (2013). Complexity theory: A new way to think. *Revista Brasileira de Linguística Aplicada*, 13(2), 369–373. <https://doi.org/10.1590/S1984-63982013000200002>
- Larsen-Freeman, D. (2014). Complexity theory. In B. Vanpatten & J. Williams (Eds.), *Theories in Second Language: An introduction* (pp. 227–244). Routledge.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.
- Lawrence, G. (2018). The role of language teacher beliefs in an increasingly digitalized world. In B. Zou & M. Thomas (Eds.), *Integrating technology into contemporary language learning and teaching* (pp.140-160). IGI Global.

- Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal*, 63(1), 13–22. <https://doi.org/10.1093/elt/ccn010>
- Lee, I. (2011). Formative assessment in EFL writing: An exploratory case study. *Changing English: Studies in Culture and Education*, 18(1), 99–111.
<https://doi.org/10.1080/1358684X.2011.543516>
- Lee, I. (2018). Future Directions for Writing Teacher Cognition. *The TESOL Encyclopedia of English Language Teaching*, 1–7. <https://doi.org/10.1002/9781118784235.eelt0563.pub2>
- Lemke, J. L. & Sabelli, N. H. (2008). Complex systems and educational change: Towards a new research agenda. *Educational Philosophy and Theory*, 40(1), 118–129.
<https://doi.org/10.1111/j.1469-5812.2007.00401.x>
- Li, J. & Barnard, R. (2011). Academic tutors' beliefs about and practices of giving feedback on students' written assignments: A New Zealand case study. *Assessing Writing*, 16(2), 137–148. <https://doi.org/10.1016/j.asw.2011.02.004>
- Li, L. (2012). Belief construction and development: Two tales of non-native English speaking student teachers in a TESOL programme. *Novitas-ROYAL (Research in Youth and Language)*, 6(1), 33–58. <file:///Users/WidAltalhi/Desktop/Li.pdf>
- Li, L. (2013). The complexity of language teachers' beliefs and practice: One EFL teacher's theories. *Language Learning Journal*, 41(2), 175–191.
<https://doi.org/10.1080/09571736.2013.790132>
- Li, L. (2017). *Social interaction and teacher cognition*. Edinburgh University Press.
- Li, L. (2020). *Language teacher cognition: A sociocultural perspective*. Palgrave Macmillan.
<https://doi.org/10.1057/978-1-137-51134-8>

- Li, L. & Walsh, S. (2011). 'Seeing is believing': looking at EFL teachers' beliefs through classroom interaction. *Classroom Discourse*, 2(1), 39–57.
<https://doi.org/10.1080/19463014.2011.562657>
- Lim, S. (2016). Learning to Teach Intelligible Pronunciation for ASEAN English as a Lingua Franca: A Sociocultural Investigation of Cambodian Pre-service Teacher Cognition and Practice. *RELC Journal*, 47(3), 313–329. <https://doi.org/10.1177/0033688216631176>
- Lim, C. P., & Chai, C. S. (2008). Teachers' pedagogical beliefs and their planning and conduct of computer-mediated classroom lessons. *British journal of educational technology*, 39(5), 807–828. <https://doi.org/10.1111/j.1467-8535.2007.00774.x>
- Lincoln, Y., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.
- Liu, Y., & Fisher, L. (2006). The development patterns of modern foreign language student teachers' conceptions of self and their explanations about change: Three cases. *Teacher Development*, 10(3), 343–360. <https://doi.org/10.1080/13664530600922203>
- Liviero, S. (2017). Grammar teaching in secondary school foreign language learning in England: Teachers' reported beliefs and observed practices. *The Language Learning Journal*, 45(1), 26–50. <https://doi.org/10.1080/09571736.2016.1263677>
- Lucas, C. (2004, March). *Attractors Everywhere - Order from Chaos*. CALResCo.
<http://www.calresco.org/attract.htm>
- Lumley, T. (2002). Assessment criteria in a large-scale writing test: what do they really mean to the raters? *Language Testing*, 19(3), 246–276.
<https://doi.org/10.1191/0265532202lt230oa>
- Mackey, A & Gass, S. (2015). *Second language research: Methodology and design* (2nd ed.). Routledge.

- Mahboob, A. & Elyas, T. (2014). English in the Kingdom of Saudi Arabia. *World Englishes*, 33(1), 128–142. <https://doi.org/10.1111/weng.12073>
- Maher, F. & Al-Khasawneh, F. M. (2010). Writing for academic purposes: Problems faced by Arab postgraduate students of the College of Business, UUM. *English for Specific Purposes World*, 9(28), 1–23. http://esp-world.info/Articles_28/WRITING.pdf
- Mao, S. S., & Crosthwaite, P. (2019). Investigating written corrective feedback:(Mis)alignment of teachers' beliefs and practice. *Journal of Second Language Writing*, 45(Complete), 46–60. <https://doi.org/10.1016/j.jslw.2019.05.004>
- Martin, S. D., & Dismuke, S. (2018). Investigating differences in teacher practices through a complexity theory lens: The influence of teacher education. *Journal of Teacher Education*, 69(1), 22–39. <https://doi.org/10.1177/0022487117702573>
- Matlin, M. (2009). *Cognition*. John Wiley & Sons.
- Matsuda, P. K. (2003). Second language writing in the twentieth century: A situated historical perspective. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 15–34). Cambridge University Press.
- Mattheoudakis, M. (2007). Tracking changes in pre-service EFL teacher beliefs in Greece: A longitudinal study. *Teaching and teacher education*, 23(8), 1272–1288. <https://doi.org/10.1016/j.tate.2006.06.001>
- McKay, S. L. (2006). *Researching second language classrooms*. Routledge.
- Merriam, S. B & Tisdell, E. J. (2016). *Qualitative Research: A guide to design and implementation* (4th ed.). Wiley.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education* (2nd ed.). Jossey-Bass Publishers.

- Messick, S. (1996). Validity and washback in language testing. *Language testing*, 13(3), 241–256. <https://doi.org/10.1177/026553229601300302>
- Min, H. T. (2013). A case study of an EFL writing teacher's belief and practice about written feedback. *System*, 41(3) 625–638. <https://doi.org/10.1016/j.system.2013.07.018>
- Ministry of Education. (n.d.). *About the ministry*. Kingdom of Saudi Arabia: Ministry of Education.
<https://www.moe.gov.sa/en/TheMinistry/AboutMinistry/Pages/Education.aspx>.
- Ministry of Finance. (2019). *Budget Statement Fiscal Year 2019*.
<https://www.mof.gov.sa/en/financialreport/budget2019/Documents/Budget Statement 2019.pdf>.
- Miri, M., Alibakhshi, G., & Mostafaei-Alaei, M. (2017). Reshaping teacher cognition about L1 use through critical ELT teacher education. *Critical Inquiry in Language Studies*, 14(1), 58–98. <https://doi.org/10.1080/15427587.2016.1238286>
- Mitchell, M. (2009). *Complexity: A guided tour*. Oxford University Press.
- Mitchell, S. D. (2003). *Biological complexity and integrative pluralism*. Cambridge University Press.
- Mohammad, T., & Hazarika, Z. (2016). Difficulties of learning EFL in KSA: Writing skills in context. *International Journal of English Linguistics*, 6(3), 105–117.
<https://doi.org/10.5539/ijel.v6n3p105>
- Montgomery, J. L., & Baker, W. (2007). Teacher-written feedback: Student perceptions, teacher self-assessment, and actual teacher performance. *Journal of Second Language Writing*, 16(2), 82–99. <https://doi.org/10.1016/j.jslw.2007.04.002>

- Mori, R. (2011). Teacher cognition in corrective feedback in Japan. *System*, 39(4), 451–467.
<https://doi.org/10.1016/j.system.2011.10.014>
- Morowitz, H. J., & Singer J. L. (1995). *The mind, the brain, and complex adaptive systems*.
Addison-Wesley.
- Morrison, K. (2008). Educational philosophy and the challenge of complexity theory.
Educational Philosophy and Theory, 40(1), 19–34. <https://doi.org/10.1111/j.1469-5812.2007.00394.x>
- Munby, H. (1984). A qualitative approach to the study of a teacher's beliefs. *Journal of Research in Science Teaching*, 21(1), 27–38. <https://doi.org/10.1002/tea.3660210104>
- Murphy, P. K., & Mason, L. (2006). Changing knowledge and beliefs. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology*, (pp. 305–324). Lawrence Erlbaum Associates Publishers.
- Murray, H. (2003). *Tracing the development of language awareness: An exploratory study of language teachers in training*. [Unpublished doctoral dissertation, University of Bern].
- Neff-Lippman, J. (2012). Assessing writing. In I. L. Clark (Ed.), *Concepts in composition: Theory and practice in the teaching of writing*, (2nd ed., pp. 145–167). Routledge.
- Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of curriculum studies*, 19(4), 317–328. <https://doi.org/10.1080/0022027870190403>
- Ng, E. K. J., & Farrell, T. S. C. (2003). Do teachers' beliefs of grammar teaching match their classroom practices? A Singapore Case Study. In D. Deterding, A. Brown & E. L. Low (Eds.), *English in Singapore: Research on grammar*, (pp. 128–137). McGraw Hill.

- Ngo, X. M. (2018). A sociocultural perspective on second language writing teacher cognition: A Vietnamese teacher's narrative. *System*, 78(1), 79–90.
<https://doi.org/10.1016/j.system.2018.08.002>
- Nisbet, S., & Warren, E. (1999). The effects of a diagnostic assessment system on the teaching of mathematics in the primary school. In O. Zaslavesky (Ed.), *Proceedings of the 23rd Conference of the International Group for the Psychology of Mathematics Education*, (pp. 1134–1157). ERIC.
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1), 1–13. <https://doi.org/10.1177/1609406917733847>
- Obeid, R. (2017). Second language writing and assessment: Voices from within the Saudi EFL context. *English Language Teaching*, 10(6), 174–181.
<https://doi.org/10.5539/elt.v10n6p174>
- Oliver, Serovich, J. M., & Mason, T. L. (2005). Constraints and Opportunities with Interview Transcription: Towards Reflection in Qualitative Research. *Social Forces*, 84(2), 1273–1289. <https://doi.org/10.1353/sof.2006.0023>
- Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, 81(3), 376–407. <https://doi.org/10.3102/0034654311413609>
- Opre, D. (2015). Teachers' conceptions of assessment. *Procedia-Social and Behavioral Sciences*, 209(1), 229–233. <https://doi.org/10.1016/j.sbspro.2015.11.222>
- Oshima, A., & Hogue, A. (2020). *Longman Academic Writing Series: Paragraphs to essays* (4th ed., Vol. 3). Pearson Education.

Oxford University Press. (n. d.). System. *In Encyclopedia.*

<https://www.encyclopedia.com/humanities/dictionaries-thesauruses-pictures-and-press-releases/system-0>

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct.

Review of educational research, 62(3), 307–332.

<https://doi.org/10.3102/00346543062003307>

Patton, M. Q. (2002). Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative Social Work*, 1(3), 261–283.

<https://doi.org/10.1177/1473325002001003636>

Peacock, M. (2001). Pre-service ESL teachers' beliefs about second language learning: A longitudinal study. *System*, 29(1), 177–195. [https://doi.org/10.1016/S0346-](https://doi.org/10.1016/S0346-251X(01)00010-0)

[251X\(01\)00010-0](https://doi.org/10.1016/S0346-251X(01)00010-0)

Pennington, M. C., & Urmston, A. (1998). The teaching orientation of graduating students on a BATESL course in Hong Kong: A comparison with first-year students. *Hong Kong Journal of Applied Linguistics*, 3(1), 17–46.

Phakiti, A. & Paltridge, B. (2015). Approaches and methods in applied linguistics research. In B.

Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: A Practical resource* (pp. 1–25). Bloomsbury.

Phipps, S., & Borg, S. (2009). Exploring tensions between teachers' grammar teaching beliefs and practices. *System (Linköping)*, 37(3), 380–390.

<https://doi.org/10.1016/j.system.2009.03.002>

Prior, L. (2008). Researching documents. In S. Hesse-Biber & P. Leavy (Eds.), *Handbook of emergent methods* (pp.111–126). Guilford.

- Punch, K. F., & Oancea, A. (2014). *Introduction to Research Methods in Education* (2nd ed.). SAGE.
- QS World University Rankings. (2021). <https://www.topuniversities.com/university-rankings/world-university-rankings/2021>.
- Rapley, T. (2018). *Doing conversation, discourse and document analysis* (2nd ed.). SAGE.
- Rashidi, N., & Moghadam, M. (2015). The discrepancy between teachers' belief and practice, from the sociocultural perspective. *Studies in English language teaching*, 3(3), 252–274. <https://doi.org/10.22158/selt.v3n3p252>
- Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, 21(3), 249–258. <https://doi.org/10.1191/0265532204lt283ed>
- Richards, K. (2009). Trends in qualitative research in language teaching since 2000. *Language Teaching*, 42(2), 147–180. <https://doi.org/10.1017/S0261444808005612>
- Richards, K. (2011). Case Study. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 207–220). Routledge. <https://doi-org.ezproxy.library.yorku.ca/10.4324/9780203836507>
- Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula (Ed.), *Handbook of Research on Teacher Education* (2nd ed., pp. 102–19). Macmillan.
- Rogers, W., Cheng, L., & Hu, H. (2007). ESL/EFL instructors' beliefs about assessment and evaluation. *Comparative and International Education*, 36(1), 39–61. <https://doi.org/10.5206/cie-eci.v36i1.9088>
- Sakyi, A. (2000). Validation of holistic scoring for ESL writing assessment: A study of how raters evaluate ESL compositions on a holistic scale. In A. J. Kunnan (Ed.), *Fairness and validation in language assessment* (pp. 130–153). Cambridge University Press.

- Salami, F., & Alharthi, R. M. (2022). Improving language assessment literacy for in-service Saudi EFL teachers. *Arab World English Journal (AWEJ)*, 13(3), 536–554.
<https://doi.org/10.24093/awej/vol13no3.35>
- Saldana, J. (2013). *The Coding Manual for Qualitative Researchers* (2nd ed.). SAGE.
- Sampson, R. J. (2016). *Complexity in Classroom Foreign Language Learning Motivation A Practitioner Perspective from Japan*. Multilingual Matters.
- Sasaki, T. (2003). Recipient orientation in verbal report protocols: Methodological issues in concurrent think-aloud. *Second Language Studies*, 22(1), 1–54.
<https://doi.org/10.1017/S0261444803001749>
- Scott, D., & Usher, R. (2010). *Researching education: Data, methods and theory in educational enquiry*. Bloomsbury Publishing.
- Scott, R., & Rodgers, B. (1995). Changing teachers' conceptions of teaching writing: A collaborative study. *Foreign Language Annals*, 28(2), 234–246.
<https://doi.org/10.1111/j.1944-9720.1995.tb00788.x>
- Seedhouse, P. (2004). Transcription Conventions. *Language Learning*, 54(S1), 267–270.
<https://doi.org/10.1111/j.1467-9922.2004.00275.x>
- Seibert, A. (2022). Transparency with second language and multilingual transcription. *TESOL Quarterly*, 56(2), 499–524. <https://doi.org/10.1002/tesq.3063>
- Shavelson, R. J., & Borko, H. (1979). Research on teachers' decisions in planning instruction. *Educational Horizons*, 57(4), 183–189. <http://www.jstor.org/stable/42924342>
- Shavelson, R. J., Cadwell, J., & Izu, T. (1977). Teachers' sensitivity to the reliability of information in making pedagogical decisions. *American Educational Research Journal*, 14(2), 83–97. <https://doi.org/10.3102/00028312014002083>

- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Skar, G. B., & Jølle, L. J. (2017). Teachers as raters: Investigation of a long term writing assessment program. *L1-Educational Studies in Language and Literature*, 17(Open Issue), 1–30. <https://doi.org/10.17239/L1ESLL-2017.17.01.06>
- Skott, J. (2014) The promises, problems and prospects of research on teachers' beliefs. In H. Fives & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs*, (pp. 13–30). Routledge.
- Smith, D. (2000). Rater judgments in the direct assessment of competency-based second language writing ability. In G. Brindley (Ed.), *Studies in immigrant English language assessment*, (pp. 159–189). Macquarie University.
- Smith, L. B., & Thelen, E. (Eds.). (1993). *A dynamic systems approach to development: Applications*. The MIT Press.
- Smith, L., & Abouammoh, A. (2013). *Higher education in Saudi Arabia*. Springer.
- Smith, W. L. (1992). The importance of teacher knowledge in college composition placement testing. In R. J. Hayes (Ed.), *Reading empirical research studies: The rhetoric of research* (pp. 289–316). Norwood.
- Soliman, M. S., & Alenazi, M. M. (2017). Primary teachers' beliefs and knowledge about self-regulated learning in the Kingdom of Saudi Arabia. *International Journal of Educational Sciences*, 18(3), 1–15. <https://doi.org/10.1080/09751122.2017.1346573>
- Speer, N. (2005). Issues of methods and theory in the study of mathematics teachers' professed and attributed beliefs. *Educational Studies in Mathematics*, 58(3), 361–391. <https://doi.org/10.1007/s10649-005-2745-0>

- Springsteen, S. M. (2014). *Examining student motivation in Saudi Arabia*. [Unpublished master's thesis, School for International Training (SIT) Graduate Institute].
- Stake, R. E. (2013). *Multiple Case Study Analysis*. Guilford Publications.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *Canadian Modern Language Review*, 50(1), 158–164.
- Tajeddin, Z., & Aryaeian, N. (2017). A collaboration-mediated exploration of nonnative L2 teachers' cognition of language teaching methodology. *Australian Journal of Teacher Education*, 42(6), 81–99. <https://doi.org/10.14221/ajte.2017v42n6.6>
- Taras, M. (2005). Assessment—summative and formative—some theoretical reflections. *British journal of educational studies*, 53(4), 466–478. <https://doi.org/10.1111/j.1467-8527.2005.00307.x>
- Thelen, E. & Smith, L. (1996). *A dynamic systems approach to the development of cognition and action*. MIT Press.
- Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127–146). Macmillan.
- Thoonen, E.J., Slegers, P.C., Peetsma, T.D., & Oort, F.J. (2011). Can teachers motivate students to learn? *Educational Studies*, 37(3), 345–360. <https://doi.org/10.1080/03055698.2010.507008>
- Tillema, H. H. (2000). Belief change towards self-directed learning in student teachers: Immersion in practice or reflection on action. *Teaching & Teacher Education*, 16(1), 575–591. [https://doi.org/10.1016/S0742-051X\(00\)00016-0](https://doi.org/10.1016/S0742-051X(00)00016-0)

- Troudi, S., Coombe, C., & Al-Hamly, M. (2009). EFL teachers' views of English language assessment in higher education in the United Arab Emirates and Kuwait. *TESOL Quarterly*, 43(3), 546-555. <http://dx.doi.org/10.1002/j.1545-7249.2009.tb00252.x>
- Tudor, I. (2003). Learning to live with complexity: towards an ecological perspective on language teaching. *System (Linköping)*, 31(1), 1–12. [https://doi.org/10.1016/S0346-251X\(02\)00070-2](https://doi.org/10.1016/S0346-251X(02)00070-2)
- Urmston, A. (2003). Learning to teach English in Hong Kong: The opinions of teachers in training. *Language and Education*, 17(2), 112–137. <https://doi.org/10.1080/09500780308666843>
- van der Leeuw, S. & Folke, C. (2021). The social dynamics of basins of attraction. *Ecology and Society*, 26(1), 33. <https://doi.org/10.5751/ES-12289-260133>
- van Geert, P. (2007). Dynamic systems in second language learning: Some general methodological reflections. *Bilingualism: Language and Cognition*, 10(1), 47–49.
- van Lier, L. (2004). *The ecology and semiotics of language learning: A sociocultural Perspective*. Kluwer Academic Publishers. <https://ebookcentral.proquest.com/lib/york/detail.action?docID=3036411>
- van Lier, L. (2005). Case study. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 195–208). Routledge. <https://doi-org.ezproxy.library.yorku.ca/10.4324/9781410612700>
- van Nes, F., Abma, T., Jonsson, H., & Deeg, D. (2010). Language differences in qualitative research: Is meaning lost in translation? *European Journal of Ageing*, 7(4), 313–316. <https://doi.org/10.1007/s10433-010-0168-y>

- van Orden, G. C. (2002). Nonlinear dynamics and psycholinguistics. *Ecological Psychology*, 14(1), 1–4. https://doi.org/10.1207/S15326969ECO1401&2double_1
- Waldrop, M. (1992). *Complexity: The emerging science at the edge of order and chaos*. New York: Simon and Schuster.
- Wang, L., Lee, I., & Park, M. (2020). Chinese university EFL teachers' beliefs and practices of classroom writing assessment. *Studies in Educational Evaluation*, 66(1), Article 100890. <https://doi.org/10.1016/j.stueduc.2020.100890>
- Weigle, S. C. (1994). Effects of training on raters of ESL compositions. *Language Testing*, 11(1), 197–223. <https://doi.org/10.1177/026553229401100206>
- Weigle, S. C. (1999). Investigating rater/prompt interactions in writing assessment: Quantitative and qualitative approaches. *Assessing Writing*, 6(1), 145–178. [https://doi.org/10.1016/S1075-2935\(00\)00010-6](https://doi.org/10.1016/S1075-2935(00)00010-6)
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.
- Weigle, S. C. (2007). Teaching writing teachers about assessment. *Journal of Second Language Writing*, 16(3), 194–209. <https://doi.org/10.1016/j.jslw.2007.07.004>
- White, E. M. (1995). An apologia for the timed impromptu essay test. *College Composition and Communication*, 46(1), 30–45. <https://doi.org/10.2307/358868>
- White, E. M. (2001). The opening of the modern era of writing assessment: A narrative. *College English*, 63(3), 306–320. <https://doi.org/10.2307/378995>
- Williams, J. D. (2003). *Preparing to teach writing: Research, theory, and practice* (3rd ed.). Erlbaum.
- Woods, D. (1996). *Teacher cognition in language teaching*. Cambridge University Press.

- Wyatt, M., & Ager, E. (2017). Teachers' cognitions regarding continuing professional development. *ELT Journal*, 71(2), 171–185. <https://doi.org/10.1093/elt/ccw059>
- Yancey, K. B. (1999). Looking back as we look forward: Historicizing writing assessment. *College Composition and Communication*, 50(3), 483–503. <https://doi.org/10.2307/358862>
- Yancey, K. B., & Huot, B. (Eds.). (1997). *Assessing writing across the curriculum: Diverse methods and practices*. Ablex.
- Yin, M. (2010). Understanding classroom language assessment through teacher thinking research. *Language Assessment Quarterly*, 7(2), 175–194. <https://doi.org/10.1080/15434300903447736>
- Yin, R. K. (2014). *Case study research: Design and methods*. SAGE.
- Yu, S. (2013). EFL teachers' beliefs and practices regarding peer feedback in L2 writing classrooms. *Polyglossia*, 24(1), 74–79.
- Yu, S., Xu, H., Jiang, L., and Chan, I. K. I. (2020). Understanding Macau novice secondary teachers' beliefs and practices of EFL writing instruction: A complexity theory perspective. *Journal of Second Language Writing*, 48(1), Article 100728. <https://doi.org/10.1016/j.jslw.2020.100728>
- Zhang, W. (1999). *The rhetorical patterns found in Chinese EFL student writers' examination essays in English and the influence of these patterns on rater response*. [Unpublished Doctoral dissertation, Hong Kong Polytechnic University].
- Zheng, H. (2015). *Teacher beliefs as a complex system: English language teachers in China*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-23009-2>

Zhu, Y., & Shu, D. (2017). Implementing foreign language curriculum innovation in a Chinese secondary school: An ethnographic study on teacher cognition and classroom practices. *System*, 66(1), 100–112. <https://doi.org/10.1016/j.system.2017.03.006>

Appendices

Appendix A: York University Ethics Approval



OFFICE OF
RESEARCH
ETHICS (ORE)
3rd Floor, Kaneff
Tower

4700 Keele St.
Toronto ON
Canada M3J 1P3
Tel 416 736 5914
www.research.yorku.ca

Certificate #:	STU 2022-087
Approval Period:	08/11/22-08/11/23

ETHICS FULL APPROVAL

To: **Wid Altalhi**
Graduate Student of Linguistics & Applied Linguistics
altalhw@yorku.ca

From: Alison M. Collins-Mrakas, Director, Research Ethics
(on behalf of You-ta Chuang, Chair, Human Participants Review Committee)

Date: Tuesday, September 13, 2022

Title: **Multiple-Case Studies of the Complexity of EFL Teachers' Writing Assessment Beliefs and Practices in a Preparatory Year Program at a Saudi University**

Risk Level: Minimal Risk More than Minimal Risk

Level of Review: Delegated Review Full Committee Review

I am writing to inform you that this research project, "**Multiple-Case Studies of the Complexity of EFL Teachers' Writing Assessment Beliefs and Practices in a Preparatory Year Program at a Saudi University**" has received ethics review and approval by the Human Participants Review Sub-Committee, York University's Ethics Review Board and conforms to the standards of the Canadian Tri-Council Research Ethics guidelines.

Note that approval is granted for one year. Ongoing research – research that extends beyond one year – must be renewed prior to the expiry date.

Any changes to the approved protocol must be reviewed and approved through the amendment process by submission of an amendment application to the HPRC prior to its implementation.

Any adverse or unanticipated events in the research should be reported to the Office of Research ethics ore@yorku.ca as soon as possible.

For further information on researcher responsibilities as it pertains to this approved research ethics protocol, please refer to the attached document, "**RESEARCH ETHICS: PROCEDURES to ENSURE ONGOING COMPLIANCE**".

Should you have any questions, please feel free to contact me at: 416-736-5914 or via email at: acollins@yorku.ca

Yours sincerely,

Alison M. Collins-Mrakas M.Sc., LLM
Director, Office of Research Ethics

Appendix B: Informed Consent Form for Participants

Study Name: Multiple-Case Studies of the Complexity of EFL Teachers' Writing Assessment Beliefs and Practices in a Preparatory Year Program at a Saudi University

Researcher Name: Wid Altalhi (Principal investigator)
Doctoral Candidate

Email address: altalhiw@yorku.ca

Linguistics and Applied Linguistics, York University,
4700 Keele Street, Toronto, ON, Canada M3J 1P3

Purpose of the Study: The study aims to investigate Saudi EFL teachers' writing assessment beliefs, writing assessment practices, and their contexts.

This study will be presented to York University as fulfillment for the degree of Doctor of Philosophy. Findings from the study will be submitted for publication to several peer-reviewed journals after the dissertation defense (Winter 2024), such as *Language Assessment Quarterly* and *Language Testing*. The findings will be presented at relevant conferences, such as Language Testing Research Colloquium (LTRC), American Association for Applied Linguistics (AAAL), Teaching English as a Second Language (TESOL) International Convention, and TESOL Arabia.

What You Will Be Asked to Do in the Research: First, you will be kindly asked to answer interview questions about your writing assessment beliefs, practices, and contexts. The interview will take about 60-90 minutes and be audio recorded. You can choose to do this interview in one or two meetings. Second, you will be kindly asked to allow the researcher to video or audio record and observe two of your writing classes. Third, you will be kindly asked to do two think-aloud sessions while marking three of your students' writing samples. The sessions will also be video or audio recorded. The estimated time for each session will be 40 minutes. Finally, after completing the observations and think-aloud sessions, you will be kindly asked to answer follow-up interview questions. The estimated time for the follow-up interview will be 30-45 minutes. You will be given an appreciation gift for your time (a \$20 gift card from Jarir bookstore).

Risks and Discomforts: The researcher does not foresee any risks or discomfort from your participation in the research.

Benefits of the Research and Benefits to You: I hope this research will enable you to be more aware of and reflect on your own writing assessment beliefs, practices, and context. The study will provide essential insights into the complexity of teacher cognition and its relationship with their practices and contextual factors. The study findings will have implications for developing curricula and teacher professional development courses and informing policy and reforms in Saudi educational institutes.

Voluntary Participation and Withdrawal from the Study: Your participation in the study is completely voluntary, and you may choose to stop participating at any time. Your decision not to volunteer, to stop participating, or to refuse to answer interview questions will not influence the

nature of the ongoing relationship you may have with the researcher or the nature of your relationship with York University, either now or in the future. If you decide to stop participating, you may withdraw without penalty, financial or otherwise, and you will still receive the promised inducement, an appreciation gift for your time. Should you wish to withdraw after the study, you will also have the option to withdraw your data up until the analysis is complete. If you withdraw from the study, all associated data collected will be immediately destroyed wherever possible.

Confidentiality: all information you supply during the interviews' audio recordings, observation, and think-aloud sessions 'video recordings will be confidential, and your name will not appear in any report or publication of the research. Only you, the participating teacher, will be video recorded, not your students. Your data and recordings (audio/video), will be saved in a password-protected file and safely stored only in the researcher's laptop and on a flash memory drive that is held in a locked cabinet, not a cloud-based service. Only the researcher will have access to the information and cabinet. The data will be stored and saved till the successful defense of the research in Fall 2023 or Winter 2024. After that, the data files will be deleted from the researcher's laptop and flash memory drives on 4/30/2024. Confidentiality will be provided to the fullest extent possible by law.

Questions About the Research? If you have questions about the research in general or about your role in the study, please feel free to contact me at my email altalhiw@yorku.ca or phone numbers [+1\(647\)671-7267](tel:+16476717267) and [+966\(54\)177-7840](tel:+966541777840) or my Graduate Advisor, Dr. Khaled Barkaoui at barkaoui@yorku.ca. You may also contact the Graduate Program in *Linguistics and Applied Linguistics* at redge@yorku.ca.

This research has received ethics review and approval by the Delegated Ethics Review Committee, which is delegated authority to review research ethics protocols by the Human Participants Review Sub-Committee, York University's Ethics Review Board, and conforms to the standards of the Canadian Tri-Council Research Ethics guidelines. If you have any questions about this process or about your rights as a participant in the study, please contact the Sr. Manager & Policy Advisor for the Office of Research Ethics, 5th Floor, Kaneff Tower, York University (telephone [416-736-5914](tel:4167365914) or email ore@yorku.ca).

Legal Rights and Signatures:

I, _____ <<insert participant's name>>, consent to participate in the study, *Multiple-Case Studies of the Complexity of EFL Teachers' Writing Assessment Beliefs and Practices in a Preparatory Year Program at a Saudi University*, conducted by *Wid Altalhi*. I have understood the nature of this project and wish to participate. I am not waiving any of my legal rights by signing this form. My signature below indicates my consent.

Participant Signature _____ **Date** _____

Principal Investigator Signature _____ **Date** _____

Additional Consent

1. Audio Recording

I, __<<insert participant's name>>, consent to the audio recording of my interview(s).

2. Video Recording

I, _____<<insert participant's name>>, consent to the use of images of me (including photographs, video, and other moving images), my environment, and property in the following ways (please check all that apply):

In academic articles N Y

In print, digital, and slide form N Y

In academic presentations N Y

In media N Y

In thesis materials N Y

OR

I, _____<<insert participant's name>>, consent to the audio recording of my observations and think-aloud sessions to be used only by the researcher for research data analysis purposes.

3. Final Consent

I, _____<<insert participant's name>>, consent that all the information I reported in the interviews, observations, and think-aloud sessions will be either audio or video recorded based on my previous choices.

Appendix C: Initial Interview Questions Guidelines

Date of the Interview: _____

Place of the Interview: _____

Length of the Interview: _____

First Section: Teacher's Background Questions

1. Highest educational qualification _____
2. Years of experience teaching English _____
3. Years of experience in PYP _____
4. How did you learn English in the past? And English writing?
5. You were an English learner before you became a teacher. So, how did you learn English? And writing?
6. Have you taught English in other places? If yes, where and at what grade levels?
7. What English courses have you taught?
8. How many years have you been teaching writing?
9. Did you get professional training in teaching, testing, and assessing writing? (Adapted from Ben Hedia, 2020). If yes:
 - a. What was the training program specifically about?
 - b. What did you learn from this professional training?
 - c. How did this program contribute to your beliefs and knowledge about second language writing and writing assessment?
10. How many students are enrolled in your current writing classes?
 - a. How would you describe the level of your students in writing?
11. Have you ever been a coordinator of any writing course in PYP? If yes, can you tell me about your experience?
12. What are your responsibilities in your writing classes as a teacher?
 - a. What is your role in the writing assessment in your courses?

Second Section: General Writing Assessment Beliefs & Practices Beliefs

13. How would you define *writing assessment*? Provide examples.
14. In your opinion, what are the purposes of writing assessments? (Adapted from Rea-Dickins, 2001)
15. How do you think writing assessment influence students' learning?
16. What kind of assessments do you believe are most effective for measuring students' performance in writing? (e.g., formative assessment, summative assessment) explain more.
17. How do you feel about writing assessments?

18. Do you believe those feelings would change for any reason?

Practices

19. How do you currently assess your students' writing? Why?

20. What essential parts do you focus on when assessing your students' writing? Why?

21. What are some challenges you face in assessing your students' writing? (Adapted from Troudi et al., 2009) please explain how you overcome them.

22. What coping strategies do you utilize when faced with some of the challenges you mentioned?

23. What type of writing assessment activities, tools, and strategies do you use in the classroom? Please describe your process of applying these strategies. Give examples.

Third Section: Specific Writing Assessment Beliefs & Practices

Tasks

24. In your opinion, what kind of writing tasks are the best to assess students' writing skills? Why?

a. How does explaining the task to your students help their performance?

25. How do you ensure that your students are prepared for the writing task?

26. How and what techniques, strategies, and materials do you use to explain the writing task to your students?

27. What would you change in the writing tasks and assignments in the course you are teaching? Why?

Giving Written Corrective Feedback

28. What comes into your mind when you hear the word *corrective feedback*?

29. To what extent do you believe *written feedback* is necessary? Explain.

30. Please describe how you give *feedback* to your students.

a. Are there any specific ways you like to follow? (e.g., written feedback, verbal feedback-direct, indirect, peer assessment, self-assessment, other)

31. What usually do you focus on when giving written feedback? (e.g., language, Content, Organization...etc.) (Adapted from Ben Hedia, 2020)

32. Do you correct/highlight errors? What type of errors? What are the most serious errors? (Ben Hedia, 2020)

33. In your opinion, what are the best ways to give feedback to students? Written or verbal feedback.

a. How do you think this will help your students?

Rating Students' Written Texts

34. To what extent do you think giving specific grades to your students is necessary?

35. How do you score or rate your students' writing? Explain.
36. How many times do you read the essay? Why? (Ben Hedia, 2020)
37. Which items do you focus most on while grading: content, organization, language, cohesion, punctuation, and spelling? Why? (Ben Hedia, 2020)
 - a. What do you believe is the most important among these rating criteria?
38. What challenges do you face when grading your students' papers?
 - a. How would you deal with a student unsatisfied with their score?
39. In your opinion, what are the best assessment tools and criteria for rating and grading students' writing?

Fourth Section: Teachers' Contexts

40. Please tell me how the following factors have contributed to your beliefs, knowledge, and conceptions about writing assessments.
 - a. schooling
 - b. teaching experiences
 - c. students' proficiency level
 - d. PYP culture and context
41. What do you think of the current writing assessment practices in your institution?
42. How would you describe an ideal writing assessment practice and context? Please elaborate.
 - a. How is that different from what you do now?
 - e. Do you have any suggestions about how students' writing should be assessed?
 - f. What would you change in the rating tools and criteria in the current course? Why?
43. Are you satisfied with your role in writing assessments in your current context? If no. What do you feel your role should be?

Appendix D: Informed Consent Form for Students

Study Name: Multiple-Case Studies of the Complexity of EFL Teachers' Writing Assessment Beliefs and Practices in a Preparatory Year Program at a Saudi University

Researcher Name: Wid Altalhi (Principal investigator)

Doctoral Candidate

Email address: altalhiw@yorku.ca

Linguistics and Applied Linguistics, York University,
4700 Keele Street, Toronto, ON, Canada M3J 1P3

I am a doctoral student at York University, and I am conducting a study about *teachers' writing assessment beliefs, writing assessment practices, and context*. Your teacher has agreed to participate in the study. Your teacher will be the focus of the study. As part of my study, I will observe two of your writing classes, video record the teacher, audio record the classroom, including your voices, and discuss the data acquired from the class with your teacher.

I would like to ask you to consent for the class and your voice to be audio recorded and discussed by the researcher and the teacher. The data will be used for research purposes only. All names used in the study, including your teacher's, will be anonymous. If I use the recordings for purposes other than research, I will ask for your consent. Please feel free to ask me any questions and express any concerns. York University's rules require I have a paper record of your consent. Please check the appropriate box below, complete the rest of the form, and return it to me.

Your cooperation is highly appreciated. If you have any questions or concerns, please contact me at altalhiw@yorku.ca.

Student Consent Form

Please choose the suitable choice for you, write your name, and then sign and date below.

1. I, _____, consent to be audio recorded by *Wid Altalhi* in relation to this study, and I consent to such data being analyzed and used for study purposes. I understand that anonymity will be preserved if extracts are included in the study.
2. I, _____, do not consent to be audio recorded and discussed by *Wid Altalhi* in relation to this study.

Signature _____ Date _____

Appendix E: Classroom Observation Guidelines

(Adapted from Cohen et al., 2007)

Date and time of the observation: _____

Duration of the observation: _____

Pseudonym of the teacher: _____

1. The Physical Environment

- a. Classroom layout
 - i. Students/ teacher arrangement
 - ii. Board position and what is on the board?
- b. Duration of the class:
- c. Student number:

2. Teaching

- a. What assignments and lessons are the teacher teaching?
 - i. writing topics and genres
 - ii. writing tasks

- b. How is the teacher facilitating students' learning?
 - i. What materials and tools are the teacher using to deliver the assignment?
 - ii. What strategy is the teacher using to elicit writing performance and introduce assignments?
 - (1) classroom activities and exercises

- c. How is the teacher monitoring students' comprehension of the assignment?
 - i. How are students responding to teachers' monitoring?
 - ii. What appears to help or hinder students' understanding? (Easen, 1985, as cited in Clark & Leat, 1998)

- d. What is the role of the teacher?

- e. What is the role of the students?

3. Interactions

- a. Description of the teacher and students' interactions during the lesson?
- b. What discussions are taking place?
 - i. Important discussions
 - ii. Side discussions

- c. Types of teacher's interaction and practices (Clark & Leat, 1998)

- i. greeting
- ii. giving information
- iii. making suggestions
- iv. requesting
- v. answering questions
- vi. reprimanding
- vii. praising
- viii. using non-verbal communication
- ix. differentiating
- x. counseling

4. Writing Assessment

- a. Does the teacher mention any writing assessment focus for the assignment? What are they?

Additional Notes

Appendix F: A Sample of an Observation Field Note

class started at 1:30 → computer problem 1 of 3

Classroom Observation Guidelines
(Adapted from Cohen et al., 2007)

Date and time of the observation: hour of 50 mins
Duration of the observation: 27 Dec → 11:10
Pseudonym of the teacher: Aya


1. The Physical Environment

a. Classroom layout

- Students/ teacher arrangement
- Board position and what is on the board?

b. Duration of the class:
c. Student number: 21

① The writing task instruction → the teacher reading the instruction word by word
② writing paragraphs topics → Compare & Contrast



2. Teaching

a. What assignments and lessons are the teacher teaching?

- writing topics and genres → introducing the first task
- writing tasks
 - Revision of the compare & contrast paragraph structure.
 - Three process of writing the task (outline-draft- Final draft)

Asking Qs ← b. How is the teacher facilitating students' learning?

- What materials and tools are the teacher using to deliver the assignment? small board presenting the task
- What strategy is the teacher using to elicit writing performance and introduce assignments?
 - classroom activities and exercises
Brainstorming & outlining for the first task (50 mins). The teacher gave them 10 mins then she started circulating around the class, monitoring the students writing outline.

Asking & Answering Qs
Asking for examples ← c. How is the teacher monitoring students' comprehension of the assignment?

- How are students responding to teachers' monitoring?
- What appears to help or to hinder students' understanding? (Easen, 1985, as cited in Clark & Leat, 1998)
The students asked the teacher that they didn't finish the chapter. The teacher said it was ok because ^{task} it will be easy and what left doesn't affect the task.

Note: ① The teacher emphasized many times that the outline & the 1st draft would be graded
② The teacher read the instruction & then asked if the students understand.

2 of 3

- d. What is the role of the teacher? *The main provider of the info and knowledge*
 e. What is the role of the students? *used L1 sometimes*

asking Qs & receiving info

3. Interactions

- a. Description of the teacher and students' interactions during the lesson?
 b. What discussions are taking place? → *Outline / compare & contrast*
 i. Important discussions → *all about task completion*
 ii. Side discussions

↓

- *Zodiac signs*
- *not finishing the chapter*

c. Types of teacher's interaction and practices (Clark & Leat, 1998)

- i. greeting
- ii. giving information → *all the time*
- iii. making suggestions → *for topic sentences*
- iv. requesting → *focus on topic sentences*
- v. answering questions → *not the time*
- vi. reprimanding
- vii. praising → *couple of times. she gave positive words (minimal)*
- viii. using non-verbal communication → *not much just interaction*
- ix. differentiating
- x. counseling

* *she is suggest by giving examples of good topic sentences*

← * *Q2 why did you game the students " " " " " ?*

* *why did you circulate around?*

4. Writing Assessment

- a. Does the teacher mention any writing assessment focus for the assignment? What are they?

- ① what part of the writing process will be graded.
- ② monitoring student's writing of Topic Sentences.
- ③ Asking many Qs.

5. Additional notes

① She started the class with revision of the cogate paragraph structure.

② Facing difficulty with the cogate-

③ The class focused on teacher-instruction

④ The teacher used worksheet for the students to use in their ^{distributed} outlining.

Q: What was the work sheet? did you prepare it? why?

Q: You read the instruction to the students, when the first time they have been introduced to instruction? who prepared them?

Q: who choose the topics of the paragraph?

Q: Before class, you expressed that you were annoyed by course coordinator for changing stuff all the time? what did you mean?

Q1
Q4

Appendix G: Think-aloud Protocols Instructions and Training Guidelines

Date and time of the TAP session: _____

Duration of the TAP session: _____

Pseudonym of the teacher: _____

Instructions about Think-aloud Protocols (TAPs)

What is a Think-aloud Protocol?

TAP requires you to say everything you think while doing a task (rating a student's paper). The purpose of the TAP is to know and understand your thinking and decision-making processes while you are *rating* and *giving written feedback* to your students' writing.

What do you need to do?

You need to rate and give written feedback to three papers of your students' essays in each assignment (2 & 3) using the same rating criteria you always use in the same way. When you correct the papers, remember to say whatever comes into your mind. Treat this session as if you are talking to yourself. Think of yourself as being alone.

What do I need to do?

I will video or audio record the session. I will sit without saying anything to interrupt your thoughts and talking except to remind you to continue talking if you stop talking for more than a few seconds.

Training Tasks to Practice TAP

Before we start, I would like you to practice talking aloud while thinking. The following mathematical warm-up tasks will help you better understand TAP and familiarize you with the procedure.

The Training Task

Without a calculator, please work out the following mathematical problems while thinking loudly.

1. $1998 + 1895$
2. $(3 + 2) \times (6 - 4)$
3. $[(3 + 2) \times (6 - 4) + 2] \times 4$
4. What is the value of x in: $4x - 3 = 3(3x + 4)$?

Appendix H: The Descriptions and Instructions of the English II Two Writing Tasks

Assignment 1: Compare/Contrast Paragraph (word count: 200-250)

Assignment 1 Objective

Students are expected to produce a well-structured paragraph (comparing & contrasting two elements) making use of pre-writing and a multi-step writing process. Students should produce an appropriate topic sentence, well-developed supporting points, and a clear conclusion using a variety of grammatical structures, academic language, discourse markers, and contextual devices appropriate to the genre.

Week 4 – Session 1 – brainstorming & outline [50 min in class]

Week 4\5 – First Draft [at home]

Week 5 – submission of Final Draft (Students should get min 3 days to submit their final draft after the feedback is given to them)

Students are not allowed to use dictionaries, electronic translators, phones, or any other digital devices during the assignment. They are also not allowed to use class material or notes.

Session 1 [In class- 50 min session]

- Students will be presented with the four topics. They will choose one topic, brainstorm ideas, and prepare a rough outline.
- This stage will not be marked as part of the final assignment, but it must be completed before the student can move on to the next stage.

Session 2 [At home]

- Students will use the outline they have prepared to write a first draft at home.
- The first draft should reach the word count and must contain all the content they intend to include in their final draft. The students will not be permitted to add content to their final draft that is not present in their first draft.
- This stage will not be marked as part of the final assignment, but it must be completed before the student can move on to the next stage.
- On the date specified by the teacher, all material from sessions 1 and 2 should be submitted to the teacher.

Feedback during Office Hours or extra hours

- Teachers will provide written feedback to the students, preferably during office hours.
- Students are permitted to keep their papers after the feedback session.

Session 3 [At home]

- The students will write their final draft based on material from sessions 1 and 2.
- Students are not allowed to make substantial changes to the content of their first draft. The content of the first draft and final draft may differ by only approximately 10%. Students can make minor changes based on the feedback they receive. Improving sentence structure, correcting grammar or vocabulary errors, changing the order of information, adding cohesive devices, etc., is permitted.

- Teachers are not required to give additional feedback on the final draft. However, teachers can use their own discretion.
- On the date specified by the teacher, the final draft should be uploaded on the assignment created on Blackboard.
- All the materials from sessions 1, 2, and 3 should be submitted to the teacher to make sure that the details have not been changed.

Please follow the late submission and plagiarism policy in such cases.

Teachers will not give the graded rubric to the students. However, they can show it to the one who wants to discuss it during office hours.

Assignment 2: Definition Paragraph (word count: 200-250)

Assignment 2 Objective

Students are expected to produce a well-structured definition paragraph making use of pre-writing and a multi-step writing process. Students should produce an appropriate topic sentence, well-developed supporting points, and a clear conclusion using a variety of grammatical structures, academic language, discourse markers, and contextual devices appropriate to the genre.

Week 7/8 – Session 1 – brainstorming & outline [50 min **in class**]

Week 7/8 – Session 2 - First Draft [at home]

Week 9 – submission of Final Draft (Students should get min 3 days to submit their final draft after the feedback is given to them)

Students are not allowed to use dictionaries, electronic translators, phones, or any other digital devices during the assignment. They are also not allowed to use class material or notes.

Session 1 [In class- 50 min session]

- Students will be presented with the four topics. They will choose one topic, brainstorm ideas, and prepare a rough outline.
- This stage will not be marked as part of the final assignment, but it must be completed before the student can move on to the next stage.

Session 2 [At home]

- Students will use the outline they have prepared to write a first draft at home.
- The first draft should reach the word count and must contain all the content they intend to include in their final draft. The students will not be permitted to add content to their final draft that is not present in their first draft.
- This stage will not be marked as part of the final assignment, but it must be completed before the student can move on to the next stage.
- On the date specified by the teacher, all material from sessions 1 and 2 should be submitted to the teacher.

Feedback during Office Hours or extra hours

- Teachers will provide written feedback to the students, preferably during office hours.
- Students are permitted to keep their papers after the feedback session.

Session 3 [At home]

- The students will write their final draft based on material from sessions 1 and 2.
- Students are not allowed to make substantial changes to the content of their first draft. The content of the first draft and final draft may differ by only approximately 10%. Students can make minor changes based on the feedback they receive. Improving sentence structure, correcting grammar or vocabulary errors, changing the order of information, adding cohesive devices, etc., is permitted.
- Teachers are not required to give additional feedback on the final draft. However, teachers can use their own discretion.
- On the date specified by the teacher, the final draft should be uploaded on the assignment created on Blackboard.
- All the materials from sessions 1, 2, and 3 should be submitted to the teacher to make sure that the details have not been changed.

Please follow the late submission and plagiarism policy in such cases.

Teachers will not give the graded rubric to the students. However, they can show it to the one who wants to discuss it during office hours.

Appendix I: Follow-up Interview Questions Guidelines

Post-TAP Questions

1. What did you think when you said that, and why?
2. What did you mean by saying or doing that?
3. What do you want to comment on further on your verbalization?
4. What were you mainly focusing on while rating and giving feedback on this essay?

Please explain why.

5. What is the best thing that the student did in this paper? What is the worst? And explain why.
6. How do you think your written feedback will improve your student's writing? Why.

Post-observation Questions

1. What did you think when this practice took place?
2. Why did you use these strategies/materials to explain and introduce the writing task to your students

Appendix J: An Example of MAXQDA Coding

The screenshot displays the MAXQDA Plus 2022 software interface. The main window shows a document titled "Document Browser: 1_II (303 Paragraphs)" with a zoom level of 130%. The document content is a text excerpt discussing writing practice and time constraints. The text is annotated with various codes and segments.

Document Browser: 1_II (303 Paragraphs)

it Writing Learning & Teaching > The Key is Practice

158 [But the] key factor in any writing course is (.) practice (.)

159 Um

160 If the students don't practice, they will not be able to write. And for this trimester and last semesters, even if we have like 18 teaching weeks or, yeah, 15, we didn't have time to, um, (.) to, to practice or to ask the students to practice because we have other, other assessments. It's not only about the writing tasks, we have editing and revising tasks, although I don't feel that this is, uh, necessary because the students, when they write a paragraph, they already edit and revise. So we don't need the, uh, editing and revising tasks

161 you mean as separate [tasks.]

162 [Yeah, we] didn't/ we don't have it this trimester, but we used to have, uh, this kind of assessment. So I think that we need more time, this is another challenge. We need more time to, uh, give the students time to practice and ask them to practice in class. Because when you ask students to practice outside the class, they will not do it. They have, like, other courses, other concerns, other/ they will not. So this is the second challenge. Also, teachers' authority is so important, like, in any, in any course, but even in, in writing, uh, specifically. Because, for example, you notice that this group of students has specific, uh, issues. So you need as a teacher to focus on this, those issues. Yeah. You don't need to give them like other, other things. So, like, based on these issues, you can design your own teaching materials and give these materials to the students. But we don't have this kind of authority. Also {Laughter} There are a lot of challenges. Another challenge is the textbook used in this writing course. It doesn't reflect real-life language. It doesn't have authenticity. It doesn't have, / I think that we need to provide the students with real-life language, with authentic language.

Code System

- Evidence of Interactions B, P, & C 44
- Description of PYP Context 41
- Teaching Background (Experience) 60
- Learning Background (Experience) 100
- Feelings about teaching & assessing writing 75
- Beliefs About Writing Learning & Teaching 133
- Beliefs About Writing Assessment 261
- Professed Writing Assessment Practices 134
- Actual Teaching & Assessment Practices 325
- Contextual Factors 223
- Challenges 22

Document Browser: 1_II (303 Paragraphs) - Code System

- ..Time Constraints
- ..Perspective about Stude
- ..Limited Teacher Authorit
- Evidence of Interactions B
- Beliefs About Writing Le
- ..Authentic Language

Document Browser: 1_II (303 Paragraphs) - Comments

- The teacher connected the time constraints because the students wont practice outside the classroom so they nee more time to practice in class
- The teacher talked about the need for teacher to assess her students' need and create materials based on those needs
- Teacher believes that students must have authentic language to learn writing.