MAJELIS PENDIDIKAN TINGGI MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH PROF. DR. HAMKA
Jl. Limau II, Kebayoran Baru, Jakarta 12130 Telp. (021) 7208177, 7222886, Fax. (021) 7261226, 7256620
Website : www.uhamka.ac.id; E-mail : info@uhamka.ac.id, uhamka1997@yahoo.co.id

KEPUTUSAN REKTOR
UNIVERSITAS MUHAMMADIYAH PROF. DR. HAMKA
Nomor: 0719/R/KM/2024

Tentang
PENGANGKATAN PANITIA SIDANG TESIS (ONLINE)
PROGRAM STUDI PENDIDIKAN BAHASA INGGRIS
SEKOLAH PASCASARJANA UNIVERSITAS MUHAMMADIYAH PROF. DR. HAMKA
SEMESTER GENAP TAHUN AKADEMIK 2023/2024

Bismillahirrahmanirrahim,
REKTOR UNIVERSITAS MUHAMMADIYAH PROF. DR. HAMKA:

Menimbang : a. Bahwa mahasiswa Program Studi Pendidikan Bahasa Inggris Sekolah Pascasarjana UHAMKA yang telah menyelesaikan ujian semua mata kuliah dan penyusunan tesisnya yang berbobot 4 (empat) sks, dipandang perlu untuk dilaksanakan Sidang Tesis.

b. Bahwa untuk kelancaran sidang tesis sebagaimana dimaksud konsideran a, maka dipandang perlu mengangkat Panitia Sidang Tesis dengan Surat Keputusan Rektor.

Mengingat :
1. Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tanggal 8 Juli 2003, tentang Sistem Pendidikan Nasional;
2. Undang-Undang Republik Indonesia Nomor 12 Tahun 2012 tanggal 10 Agustus 2012, tentang Pendidikan Tinggi;
3. Undang-Undang Republik Indonesia Nomor 14 Tahun 2005 tanggal 30 Desember 2005, tentang Guru dan Dosen;
4. Peraturan Pemerintah Republik Indonesia Nomor 4 Tahun 2014 tanggal 30 Januari 2014, tentang Penyelenggaraan Pendidikan Tinggi dan Pengelolaan Perguruan Tinggi;
5. Peraturan Presiden Republik Indoensia Nomor 8 Tahun 2012 tanggal 17 Januari 2012, tentang Kerangka Kualifikasi Nasional Indonesia;
6. Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 3 Tahun 2020 tanggal 24 Januari 2020, tentang Standar Nasional Pendidikan Tinggi ;
7. Keputusan Direktur Jenderal Pendidikan Tinggi Depdikbud Republik Indonesia Nomor 138/DIKTI/Kep/1997 tanggal 30 Mei 1997, tentang Perubahan Bentuk Institut Keguruan dan Ilmu Pendidikan (IKIP) Muhammadiyah Jakarta menjadi Universitas Muhammadiyah Prof. DR. HAMKA;
8. Keputusan Direktur Jenderal Pendidikan Tinggi Depdikbud Republik Indonesia Nomor 155/D/T/2007 tanggal 29 Januari 2007, tentang Izin Pembukaan Program Studi Pendidikan Bahasa Inggris Program Magister Pada Universitas Muhammadiyah Prof. DR. HAMKA di Jakarta yang diselenggarakan oleh Persyarikatan Muhammadiyah;
9. Peraturan Pimpinan Pusat Muhammadiyah Nomor 01/PRN/I.O/B/2012 tentang Majelis Pendidikan Tinggi dan Pedoman Pimpinan Pusat Muhammadiyah Nomor 02/PED/I.O/B/2012 tentang Perguruan Tinggi Muhammadiyah;
10. Ketentuan Majelis Pendidikan Tinggi Pimpinan Pusat Muhammadiyah Nomor 178/KET/I.3/D/2012 tentang Penjabaran Pedoman Pimpinan Pusat Muhammadiyah Nomor 02/PED/I.0/B/2012 tentang Perguruan Tinggi Muhammadiyah;

11. Keputusan Pimpinan Pusat Muhammadiyah Nomor 66/KEP/I.0/D/2023 tanggal 24 Januari 2023, tentang Penetapan Rektor UHAMKA Masa Jabatan 2023-2027;
12. Statuta Universitas Muhammadiyah Prof. DR. HAMKA Tahun 2023;
13. Keputusan Rektor Universitas Muhammadiyah Prof. DR. HAMKA tentang Perbaikan Surat Keputusan Rektor Nomor 530/A.31.01/2012 tentang Pengubahan Nama Program Pascasarjana menjadi Sekolah Pascasarjana Universitas Muhammadiyah Prof. DR. HAMKA;
14. Keputusan Rektor Universitas Muhammadiyah Prof. DR. HAMKA Nomor 515/A.01.01/2023 tanggal 30 Mei 2023 tentang Pengangkatan Direktur Sekolah Pascasarjana UHAMKA Masa Jabatan 2023-2027

Memperhatikan : Kurikulum Operasional Program Studi Pendidikan Bahasa Inggris Sekolah Pascasarjana UHAMKA;

MEMUTUSKAN

Menetapkan :

Pertama : Mengangkat Panitia Sidang Tesis Pendidikan Bahasa Inggris Sekolah Pascasarjana Universitas Muhammadiyah Prof. DR. HAMKA Semester Genap Tahun Akademik 2023/2024 sebagaimana tercantum dalam lampiran 1 Keputusan ini.

Kedua : Apabila salah seorang di antara Panitia Penguji tidak dapat melaksanakan tugas karena sakit atau karena hal lainnya, maka ditunjuk penguji pengganti oleh Direktur.

Ketiga : Menetapkan peserta Ujian Sidang Tesis Program Studi Pendidikan Bahasa Inggris sebagaimana tercantum pada lampiran 2 lajur 4, dengan judul tesis sebagaimana tersebut pada lajur 5 Keputusan ini.

Keempat : Ujian sidang tesis dilaksanakan secara lisan oleh penguji pada hari dan tanggal sebagaimana tercantum dalam lampiran Surat Keputusan ini.

Kelima : Pelaksanaan Sidang Tesis diketuai oleh Direktur, diuji oleh dua orang penguji dan dua orang pembimbing sebagai anggota tim penguji tesis dari masing-masing mahasiswa yang mengikuti sidang tesis.

Keenam : Semua biaya yang berkaitan dengan sidang tesis ini dibebankan kepada anggaran Sekolah Pascasarjana UHAMKA yang diatur khusus untuk kepentingan tersebut.

Ketujuh : Pengumuman lulus atau tidak lulus disampaikan oleh Direktur kepada peserta ujian tesis berdasarkan hasil rapat Panitia Sidang Tesis pada hari pelaksanaan ujian, setelah keseluruhan peserta selesai mengikuti Sidang Tesis.

Kedelapan : Keputusan ini berlaku sejak tanggal ditetapkan sampai dengan berakhirnya pelaksanaan Sidang Tesis

Kesembilan : Surat Keputusan ini disampaikan kepada pihak-pihak yang terkait untuk dilaksanakan sebagaimana mestinya.

Kesepuluh : Apabila dalam keputusan ini terdapat kekeliruan, maka akan diperbaiki sebagaimana mestinya.

Ditetapkan di : Jakarta
Pada tanggal : 29 Dzulhijjah 1445 H
6 Juli 2024 M

Rektor,

Prof. Dr. H. Gunawan Suryoputro, M.Hum.

Tembusan:

Yth.
1. Direktur SPs
2. Kepala Biro Akademik
3. Kaprodi Pendidikan Bahasa Inggris SPs
4. Dosen dan Mahasiswa yang bersangkutan
   Universitas Muhammadiyah Prof. DR. HAMKA

Lampiran 1 Keputusan Rektor UHAMKA
Nomor : 0719 /R/KM/2024
Tanggal : 29 Dzulhijjah 1445 H / 6 Juli 2024 M

PANITIA SIDANG TESIS
PROGRAM STUDI PENDIDIKAN BAHASA INGGRIS
SEKOLAH PASCASARJANA UNIVERSITAS MUHAMMADIYAH PROF. DR. HAMKA
SEMESTER GENAP TAHUN AKADEMIK 2023/2024

Penanggung Jawab : Rektor
Prof. Dr. H. Gunawan Suryoputro, M.Hum.

Ketua : Direktur Sekolah Pascasarjana
Prof. Dr. H. Ade Hikmat, M.Pd.

Sekretaris : Ketua Program Studi Pendidikan Bahasa Inggris
Siti Zulaiha, M.AL, Ph.D.

Anggota Penguji :
1. Siti Zulaiha, M.AL, Ph.D.
2. Prof. Herri Mulyono, Ph.D.
3. Dr. Syaadiah Arifin, M.Pd.
4. Hamzah Puadi Ilyas, Ph.D.
5. Dr. Akhmad Haqiqi Ma'mun, M.Pd.

Pelaksana Teknis :
1. Sekretaris SPs I, Prof. Dr. H. Abd. Rahman. A. Ghani, M.Pd.
2. Sekretaris SPs II, Dr. Ihsana El Khuluqo, M.Pd.
3. Kepala Tata Usaha, Taufan Maulana Yusuf, S.Ak.
4. Kasubag Akademik, Nurlaelah, S.KM.
5. Kasubag Keuangan, Ahmad Rizky, A.Md.
6. Kasubag Umum, Agus Purlianto, A.Md.
7. Staf Sekolah Pascasarjana
Universitas Muhammadiyah Prof. DR. HAMKA

Rektor,

Prof. Dr. H. Gunawan Suryoputro, M.Hum.

Lampiran 2 Keputusan Rektor UHAMKA
Nomor : 079 /R/KM/2024
Tanggal : 9 Muharram 1446 H
15 Juli 2024 M

DAFTAR NAMA PESERTA, PEMBIMBING DAN PENGUJI SIDANG TESIS DARING (ONLINE)
PROGRAM STUDI PENDIDIKAN BAHASA INGGRIS
SEKOLAH PASCASARJANA UNIVERSITAS MUHAMMADIYAH PROF. DR. HAMKA
SEMESTER GENAP TAHUN AKADEMIK 2023/2024

Hari, Tanggal : Rabu, 17 Juli 2024
Tempat : ZOOM

NO	WAKTU	NIM	NAMA	JUDUL TESIS	PEMBIMBING / PENGUJI	PENGUJI
(1)	(2)	(3)	(4)	(5)	(6)	(7)
08.00 – 08.30		PEMBUKAAN DAN PENGARAHAN UJIAN TESIS				
1.	07.00-08.00	2009067010	SITTI AISYAH RATNA FURI	Digital Storytelling Project for Developing Students' Writing Skill	1. Dr. Syaadiah Arifin, M.Pd. 2. Hamzah Puadi Ilyas, Ph.D.	1. Siti Zulaiha, M.AL., Ph.D. 2. Dr. Akhmad Haqiqi Ma'mun, M.Pd.
2.	08.00-09.00	2109067006	SHAFA RAMADHANYA JAMIL	The Application of Semiotic Technology Approach in Developing Digital Story Telling for EFL Reading Materials	1. Prof. Herri Mulyono, Ph.D. 2. Siti Zulaiha, M.AL., Ph.D.	1. Dr. Syaadiah Arifin, M.Pd. 2. Hamzah Puadi Ilyas, Ph.D.
3.	09.00-10.00	2009067001	FACHRY ALI WIBOWO	Language Assessment: Indonesian Postgraduate Students Conception	1. Siti Zulaiha, M.AL., Ph.D. 2. Dr. Syaadiah Arifin, M.Pd.	1. Hamzah Puadi Ilyas, Ph.D. 2. Dr. Akhmad Haqiqi Ma'mun, M.Pd.

Rektor,

Prof. Dr. H. Gunawan Suryoputro, M.Hum.

LANGUAGE ASSESSMENT: INDONESIAN ENGLISH TEACHERS' CONCEPTION

THESIS

Submitted to Fulfill Requirement for Thesis Writing of
the Master of Education Degree in English

By

FACHRY ALI WIBOWO
NIM 2009067001

DEPARTMENT OF ENGLISH EDUCATION
GRADUATE SCHOOL
UNIVERSITY OF MUHAMMADIYAH PROF. DR. HAMKA
2024

ABSTRAK

Fachry Ali Wibowo, Language Assessment: Indonesian English Teachers' Conception. Thesis. Master of English Education, Post Graduate School University of Muhammadiyah Prof. DR. HAMKA. July 2024.

Peneliti memilih 10 guru bahasa Inggris yang secara khusus mengambil jurusan Pendidikan Bahasa Inggris di sebuah Universitas Negeri di Banten. Peneliti bertujuan untuk memperluas penelitian ini dengan pandangan yang beragam, memanfaatkan keahlian Universitas dalam pembelajaran jarak jauh. Karena populasi siswa yang besar, perbedaan budaya dan latar belakang mungkin memberikan tantangan tambahan selain subjek utama.

Penelitian ini menggunakan pendekatan survei deskriptif kualitatif. Secara garis besar, langkah-langkah penelitian survei dibagi menjadi 4 tahap, yaitu persiapan, pengumpulan data, analisis data, dan penarikan kesimpulan. Data yang diperoleh dari penelitian ini adalah data kualitatif. Data kualitatif diperoleh dari perspektif responden dalam mengisi pertanyaan terbuka. Teknik pengumpulan data yang digunakan adalah wawancara dengan responden. Wawancara adalah bentuk pengumpulan data yang paling sering digunakan dalam penelitian kualitatif (Indrawami, 2022).

Peserta penelitian akan terdiri dari kelompok guru bahasa Inggris yang tidak terpisah yang saat ini merupakan mahasiswa pascasarjana dari Departemen Pendidikan Bahasa Inggris. Para mahasiswa ini berasal dari berbagai provinsi di Indonesia dan sedang mengejar gelar magister dalam penilaian bahasa. Sekitar setengah dari peserta adalah guru bahasa yang mengejar gelar magister mereka secara paruh waktu. Mayoritas dari mereka yang tidak mengajar saat ini memiliki pengalaman sebelumnya sebagai guru bahasa sebelum memulai gelar magister mereka.

Penelitian ini telah memberikan wawasan tentang masalah konsep penilaian guru. Kontribusi utama dari penelitian ini adalah memberikan model untuk memahami konsep penilaian sebagai proses yang kompleks. Konteks dengan elemen pendidikan, kontekstual, atau budaya yang serupa mungkin mengungkapkan persepsi yang serupa. Lebih penting lagi, kerangka konseptual mendorong peneliti untuk terus mencari model yang sesuai yang cocok dengan konteks tertentu daripada mengikuti model internasional tertentu.

ABSTRACT

Fachry Ali Wibowo, Language Assessment: Indonesian English Teachers' Conception. Thesis. Master of English Education, Post Graduate School University of Muhammadiyah Prof. DR. HAMKA. July 2024.

The researcher selected 10 English teachers who are specifically majoring in Department of English Education at State University in Banten. The researcher aims to extend this research with diverse views, leveraging the University's expertise in distance learning. Due to their large student population, cultural and background disparities may provide additional challenges alongside the primary subjects.

This study used a qualitative descriptive survey research approach. Broadly speaking, survey research steps are divided into 4 stages, namely preparation, data collection, data analysis, and drawing conclusions. The data obtained from this research is qualitative data. Qualitative data obtained from the perspective of respondents in filling out open questions. The data collection technique used was to conduct interviews with respondents. Interviews are the most frequently used form of data collection in qualitative research (Indrawami, 2022).

The research participants will consist of an undivided group of English teachers who are currently postgraduate students from the English Education Department. These students hail from various provinces in Indonesia and are pursuing a master's degree in language assessment. Approximately half of the participants are language teachers pursuing their master's degree on a part-time basis. The majority of those not currently teaching had prior experience as language teachers before embarking on their master's degrees.

This study has provided insight into the issue of teachers' conceptions of assessment. The major contribution of this study is providing a model to understand conceptions of assessment as a complex process. Contexts with similar educational, contextual, or cultural elements might reveal similar perceptions. More importantly, the conceptual framework encourages researchers to continue searching for a suitable model that fits a specific context rather than following a particular international model.

THESIS ADVISOR APPROVAL

LANGUAGE ASSESSMENT: INDONESIAN POSTGRADUATE STUDENTS' CONCEPTION

THESIS

By

FACHRY ALI WIBOWO
NIM 2009067001

Approved defended on the thesis examination

Advisors	Signature	Date
Siti Zulaiha, M.AL., Ph.D.	……..	……..
Dr. Syaadiah Arifin M.Pd.	……..	……..

Jakarta, ……..
Chair of English Education Department
Graduate School
University of Muhammadiyah Prof. DR. HAMKA,

Siti Zulaiha, M.AL., Ph.D.

THESIS COMMITTEE APPROVAL

LANGUAGE ASSESSMENT: INDONESIAN POSTGRADUATE STUDENTS' CONCEPTION

THESIS

By

FACHRY ALI WIBOWO

NIM 2009067001

The Thesis Committee and Oral Defense Committee have approved this Thesis as partial fulfillment of the requirement of the Master of Education Degree in English

Thesis Committee	Signature	Date
Prof. Dr. H. Ade Hikmat, M.Pd. (Chair)	……..	……..
Siti Zulaiha, M.AL., Ph.D. (Secretary)	……..	……..
Siti Zulaiha, M.AL., Ph.D. (Thesis Advisor 1)	……..	……..
Dr. Syaadiah Arifin M.Pd. (Thesis Advisor 2)	……..	……..
..... (Oral Defense Committee 1)	……..	……..
..... (Oral Defense Committee 2)	……..	……..

Jakarta, ……..

Director of Graduate School

University of Muhammadiyah Prof. DR. HAMKA,

Prof. Dr. H. Ade Hikmat, M.Pd.

TABLE OF CONTENTS

TITLE .....
ABSTRAK i
ABSTRACT ii
THESIS ADVISOR APPROVAL iii
THESIS COMMITTEE APPROVAL iv
ACKNOWLEDGEMENT v
TABLE OF CONTENTS vi
LIST OF TABLES vii
LIST OF FIGURES viii
LIST OF APPENDICES iv
CHAPTER I: INTRODUCTION 1
A. Background to the Problems 1
B. Identification of Problems 2
C. Limitation of the Problems 3
D. Research Questions 3
E. Objectives of the Research 3
F. Significance of the Research 3
CHAPTER II: THEORETICAL FRAMEWORK 5
A. Review of Previous Research 5
B. Theoretical Framework 7
CHAPTER III: RESEARCH METHODS 17
A. Research Setting 17
B. Research Design 17
C. Research Participants 18
D. Data Collecting Techniques 18
E. Data Analysis 18
CHAPTER IV: FINDING AND DISCUSSION 20
A. Finding 20
B. Discussion 40
CHAPTER V: CONCLUSION AND SUGGESTION 42
A. Conclusion 42
B. Suggestion 42
REFERENCES 43
APPENDICES 49

CHAPTER I

INTRODUCTION

A. Background of the Problems

Assessment plays a crucial role in the process of learning and teaching English, involving the analysis of student performance data collected through various methods (Brown & Hirschfeld, 2008, p. 4). According to Xu & Zammit (2020), language instructors rely on assessment for multiple purposes. Primarily, teachers use assessments to gauge students' understanding of the subject and adjust their teaching strategies and materials accordingly. Therefore, assessment is essential in the learning-teaching process, helping both learners and teachers confirm the occurrence of learning (Zolfaghari & Ahmadi, 2016). Additionally, assessments can be used to determine student placement, as educators need to assess language proficiency at certain stages. Assessments also diagnose specific difficulties in students' language learning. While some forms of assessment carry significant consequences and influence critical decisions, others may have a minimal impact on the student's academic path and are considered low stakes. Regardless of their purpose, assessments play a significant role in teaching and learning.

According to Douali et al. (2022), analyzing the assessment process within an educational setting allows for evaluating the overall quality of the system. In a poor educational system, assessment is often neglected and undervalued, primarily associated with scores and tests by most learners. For some, "assessment" might bring up negative emotions like pain, fear, punishment, and competition. In such environments, teachers may only assess students for post-course evaluations or when necessary.

For many years, summative assessment (SA) for accountability has been the dominant practice in classroom assessment. However, recent trends show a shift (Hargreaves, Earl, & Schmidt, 2002) towards using assessment to enhance learning (Guskey, 2003). This shift is particularly noticeable in countries with low-stakes examination policies. Conversely, in regions where SA practices are historically and culturally ingrained (I.-C. Choi, 2008), both teachers and the public often continue to uphold traditional SA methods (Earl, 2003). This indicates a discrepancy between high-stakes

and low-stakes assessment environments, suggesting that assessment practices are influenced by cultural beliefs and traditions. Therefore, it is important to investigate teachers' conceptions of assessment across different contexts and cultures.

Additionally, there is limited research on assessment in Asian contexts where exams are deeply embedded in teaching and learning cultures. This study explores how assessment functions in Indonesia and demonstrates how teachers' conceptions influence their practices, using qualitative methods to allow participants to express their views.

However, empirical studies indicate that many teachers still rely heavily on exams to assess students, using results to determine course success or failure and sometimes as a disciplinary measure (López, 2008). Research also reveals that students often lack sufficient training in assessment knowledge and skills. For instance, in Colombia, López and Bernal (2009) found that only about 25% of undergraduate programs offer a course on assessment. Similarly, in Iran, administrative, institutional, and professional challenges hinder the implementation of new English curriculum and assessment reforms (Razavipour & Rezagah, 2018). Language assessment is crucial for enhancing classroom instruction quality, making language assessment literacy essential for modern language teachers.

B. Identification of the Problems

After broadly discussing the issues in the previous section, the researcher proceeded in this section to specifically identify the problems as follows:

1. The participants need better conceptions of language assessment.
2. Less objective and reliable assessment processes can cause poor conception of language assessment.
3. Assessment is only a means for threatening students to comprehend the materials.
4. The participants need more motivation in language assessment.

C. Limitation of the Problems

This research focuses on exploring how language teachers conceive of assessment. Furthermore, this research seeks insight from individuals with diverse backgrounds and characteristics.

D. Research Questions

In more precise terms, the subsequent research questions are explored:

1. What are the conceptions of assessment held by Indonesian English teachers?
2. What factors do teachers perceive contribute to their understanding and use of assessment in student learning?

E. Objectives of the Research

Based on the formulation of the problems above, the study, therefore, aims to:

1. Find out the conceptions of assessment held by Indonesian English teachers.
2. Find out the factors that teachers perceive which contribute to their understanding and use of assessment in student learning.

F. Significance of the Research

The findings of the study will contribute not only theoretically but also practically to language assessment.

a. Theoretical Contributions

The research findings may be beneficial for supporting the theories on language assessment.

b. Practical Contributions

Assessment can influence learner behavior positively or negatively, depending on how it is implemented and the learner's perception of it (Gronlund, 1997). Believing in the value of information from adaptive learning can foster positive learning behaviors. Conversely, if learners see no relevance in it, they may develop negative behaviors. Therefore, learner behavior before, during, and after assessment depends on their understanding and perception of it (Segers et al., 2006).

It is crucial to understand students' perceptions of assessment (Muñoz, 2021). Researching the importance of assessment to learners is essential, as findings can offer valuable insights to improve teaching strategies and activities. This research is significant because its outcomes will have important implications for educators, educational institutions, parents, and policymakers in education.

The issue of teachers' conceptions of assessment, especially in an eastern context and at the secondary school level, has not been fully explored (Brown, 2008). This research aims to fill this gap by investigating Indonesian English teachers' conceptions of assessment in student learning. It identifies the teachers' beliefs and explores why and how they hold these conceptions.

CHAPTER II

THEORETICAL FRAMEWORK

A. Review of Previous Research

Assessment literacy is now seen as crucial to teaching (Popham, 2014; Stiggins, 2014). Yet, globally, teachers' assessment literacy remains underdeveloped, with many language teachers engaging in assessment activities with little or no training. Studies in several European countries (Berry et al., 2019; Vogt & Tsagari, 2014) show that most language teachers receive minimal formal training in assessment. For instance, Tsagari and Vogt (2017) found that foreign language teachers across various European countries perceived a need for more training in language testing and assessment.

Alternative assessment methods are not commonly used in everyday practices. To make up for the lack of formal training, teachers often rely on mentors, colleagues, and published materials, which, according to Tsagari and Vogt, may hinder the use of research-based assessment knowledge. These findings align with previous studies in the European assessment context (Hasselgreen et al., 2004; Kvasova & Kavytska, 2014).

Lam (2019) examined English language teachers' knowledge and practices of classroom-based writing assessment in Hong Kong, finding that their perceived language assessment literacy was moderate and not fully mastered (p. 85). Some teachers had a limited understanding of formative assessment concepts. Similar to previous studies (e.g., Tsagari & Vogt, 2017; Vogt & Tsagari, 2014), participants reported needing training in alternative assessment methods.

Research suggests that teachers' conceptions of assessment vary by context (Barnes et al., 2015; Bonner, 2016). Brown, Gebril, and Michaelides (2019) categorized assessment contexts into low-stakes (e.g., New Zealand, Queensland, Cyprus, and Catalonia) and high-stakes (e.g., Hong Kong, China, Iran, Egypt, India, and Ecuador). Teachers in low-stakes contexts see assessment as a tool for improvement and accountability, while those in high-stakes contexts view it as crucial

for student and school accountability, linking student accountability to improvement (Brown, Hui, Yu, & Kennedy, 2011; Gebril & Brown, 2014).

Razavipour and Rezagah (2018) find that teachers need clearer guidance on assessment content and objectives. They also struggle to align their assessment methods with communicative language teaching principles. These findings highlight the need to improve language teachers' assessment literacy, helping them understand and implement language assessment reforms effectively.

In a study by Hakim (2015) on the assessment literacy of 30 English language instructors at a language center in Saudi Arabia, each participant had a unique understanding of assessment tools. Importantly, their knowledge was minimally influenced by teaching experience, and despite their understanding, they saw a need to improve practical application. Tong and Adamson (2015) surveyed and interviewed 45 secondary school students in Hong Kong about school-based assessment. They found that most students had negative views on the assessment process and were dissatisfied with the feedback from teachers, though they agreed that feedback is crucial for learning.

These findings underscore the need for further research and improvement in language assessment methods in Indonesia. Proficiency in language assessment is essential for teachers to apply appropriate methods (Nurdiana, 2021). Language assessment literacy involves comprehensive knowledge and skills in creating, executing, and analyzing language tests (Hidri, 2020). Teachers need this literacy to evaluate students' language skills accurately and ensure fair and valid exams (Hidri, 2020).

Overall, teachers’ conceptions of assessment are complex and vary with assessment policies and educational levels. Primary school teachers tend to focus on assessment for improvement, while secondary school teachers view it as a means of accountability. The demand for using assessment to improve instruction while maintaining high-stakes assessments creates tension between summative and formative assessments, which teachers must learn to balance.

This literature review shows a gap in research on Indonesian language teachers' conceptions of assessment. Additional qualitative research in this area could provide new insights.

B. Theoretical Framework

1. Assessment in English language teaching

The terms evaluation, assessment, and test are sometimes used interchangeably in educational assessment literature, but they denote different activities (Khan, 2018). Assessment refers to all methods for gathering information about student knowledge, skills, and abilities, including observation and self-assessment (Purpura, 2016, p. 191). It includes both formal assessments (tests) and informal assessments (e.g., portfolios). A test specifically refers to formal assessments that measure a student's knowledge or ability to determine what they know or how well they have mastered the curriculum (Mathew & Poehner, 2013).

Classroom assessment serves various purposes, such as screening, diagnostic, record-keeping, feedback, certification, selection, motivation, control, and placement (Gipps & Stobart, 1993; Buhagiar, 2007). Screening identifies students who need additional assistance. Diagnostic assessment, conducted before instruction, identifies students' strengths and weaknesses, focusing more on weaknesses for remediation purposes (Gitsaki & Robby, 2018). Record-keeping generates assessment scores to decide on a student's progression from one class to another. Feedback aims to enhance the learning and teaching process. Placement assesses a student's current language ability to stream them into appropriate groups, facilitating targeted instruction and accelerating language acquisition (Gitsaki & Robby, 2018).

These assessment purposes can be categorized under summative and formative assessments. Summative assessments, often conducted at the end of a unit, semester, or school year, measure student achievement and are used to make high-stakes decisions, such as retention and promotion (Cizek, 2010).

2. Principles of language assessment and testing

The principles of language assessment are applicable to all types of assessments (Brown, 2004). However, the level of concern for these principles varies significantly depending on the

type of assessment. For example, high-stakes language assessments place a greater emphasis on reliability and validity compared to classroom-based assessments (Rossi & Brunfaut, 2020; Cumming, 2012). Bachman and Palmer (1996) introduced a framework for evaluating language assessments called the test usefulness model. They argued that "the most important consideration in designing and developing a language test is the use for which it is intended, making the test's usefulness its most important quality" (p. 17). The test usefulness model includes six qualities: reliability, validity, authenticity, interactiveness, impact, and practicality.

Bachman and Palmer also proposed three principles for applying the test usefulness model in the development and use of language tests. The first principle emphasizes maximizing overall test usefulness rather than focusing on individual test qualities. The second principle states that the six test qualities should not be evaluated separately but should be assessed based on their combined effect on the test's overall usefulness. The final principle involves determining the test's usefulness and finding an appropriate balance among the test qualities for each assessment context.

2.1. Reliability

In the twentieth century, the test evaluation process primarily emphasized two assessment principles: validity and reliability (Kunnan, 2005). Reliability refers to the consistency or stability of test scores (Fulcher, 2013) and can be understood through the classical true score measurement perspective. This model posits that a student's observed test scores are composed of the true score, reflecting the student's language ability, and the error score, which is influenced by factors unrelated to the student's language ability (Bachman, 1990, pp. 166-167). This approach to reliability involves dividing a test into two halves and measuring the consistency between them.

Brown (2004) identifies four sources of test score unreliability. The first source is the student, with factors such as illness, fatigue, lack of motivation, anxiety, and effective test-taking strategies contributing to unreliability. The second source is the rater (Brown, 2004), with rater reliability divided into intra-rater reliability (consistency of scores given by the same rater on different occasions) and inter-rater reliability (consistency of scores given by different raters for the same student performance) (Rossi & Brunfaut, 2020).

The third source of unreliability is the test administration conditions (Brown, 2004), where factors like noise, photocopying variations, lighting, temperature, and furniture can affect score consistency. The final source is the test itself (Brown, 2004), as the testing method can influence student scores (Bachman, 1990; Shohamy, 1997). Shohamy (1997) argued that factors such as item type (e.g., multiple-choice, open-ended questions), text genre (expository or narrative), and testing tasks (reporting or interviewing) impact students' scores.

2.2. Validity

Currently, validity is viewed as the most critical principle in language assessment and testing (Cumming, 2012). Traditionally, validity was seen as a feature of the test itself, meaning a test was considered valid if it measured what it was intended to measure (Chapelle, 1999, 2013). Validity was also understood as a tripartite concept, including content, criterion, and construct validity. Content validity, often assessed by experts, pertains to how well assessment items represent the skills and knowledge of the substantive domain or how well the assessment tasks reflect real-world scenarios (Li, 2016, p. 808). Criterion validity measures how well test scores correlate with a relevant criterion and is divided into concurrent and predictive validity (Li, 2016).

Construct validity, the third type, concerns "whether the theoretical claims relating to the concept are warranted and whether the hypothesized/postulated relationships are corroborated by empirical evidence" (Li, 2016, p. 808). The separation of validity into these three types led to the belief that each type of validity was specific to a particular test type: content validity for achievement tests, criterion validity for aptitude tests, and construct validity for personality tests (Newton & Shaw, 2014; Shepard, 1993). Messick (1989) argued for an integrated theory of validity, proposing that construct validity should encompass both content and criterion validity.

2.3. Authenticity

When the concept of authenticity first appeared in applied linguistics in the 1960s, it referred to using materials not specifically created for non-native speakers (Fulcher, 2000; Lewkowicz,

2000). Authentic materials are said to prepare learners for real-world use by closely mimicking the language used outside the classroom (Gan, 2012). Bachman (1990) identifies two perspectives on authenticity: the real-life approach and the interactional/ability approach. The real-life approach measures authenticity by how well student performance on a test reflects actual language use in specific situations. In contrast, the interactional/ability approach focuses on the interaction between the student and the assessment task.

Brown (2004) offers guidelines for teachers to create more authentic assessments. He suggests using natural language in assessment items, contextualizing these items, selecting topics that interest and are relevant to students, and designing tasks that mirror real-world activities. Brown also advocates for direct (performance) assessments to help teachers develop assessments that simulate real-life scenarios.

2.4. Interactiveness

Interactiveness is linked to the interactional/ability approach to authenticity (Bachman, 1990). It refers to "the extent and type of involvement of the test taker's characteristics in accomplishing a test task" (Bachman & Palmer, 1996, p. 25). The student attributes considered include language ability (comprising language knowledge and strategic competence, or metacognitive strategies), topical knowledge, and schemata. An assessment task's interactiveness is evaluated based on how well it engages these aspects of the student's abilities.

Like authenticity, interactiveness can vary significantly (Spence-Brown, 2001). In summary, both authenticity and interactiveness can differ widely in language assessment tasks. According to Bachman and Palmer (1996), some assessment tasks may be highly authentic but have low interactiveness, and vice versa. Additionally, some tasks may be low or high in both authenticity and interactiveness.

2.5. Impact

The terms impact (Bachman & Palmer, 1996), washback (Wall & Alderson, 1993), backwash (Hughes, 2003), and consequential validity (Messick, 1989) refer to various effects of testing on teaching and learning. Bachman and Palmer (1996) view washback as a component of impact. Impact is a broad term encompassing the effects of testing on individuals, policies, practices, classrooms, schools, educational systems, or society (Wall, 1997, p. 291), while washback specifically refers to the influence of testing on teaching and learning.

Washback is a multifaceted phenomenon with five dimensions: specificity, intensity, length, intentionality, and value (Watanabe, 1997, as cited in Watanabe, 2004). Specificity can be general or specific, with general washback related to any test-induced effect (e.g., encouraging students to study harder) and specific washback tied to a particular test type (e.g., emphasizing speaking skills in language tests). Intensity ranges from strong, where a test dominates classroom activities, to weak, where it only affects certain events or individuals.

The length of washback can be short-term, lasting only during the preparation period, or long-term, continuing even after the test is administered. Intentionality distinguishes between intended effects (intended washback) and unintended consequences on teaching and learning. Finally, washback value can be positive, fostering a favorable attitude towards the test, or negative, leading to practices like teaching to the test and reducing instructional time. Positive washback aligns with intended washback, whereas negative washback is associated with unintended consequences.

2.6. Practicality

Though often overlooked, practicality is vital for an effective assessment system (Brunfaut, 2014; Jin, 2018). Bachman and Palmer (1996) emphasized that while other test qualities like reliability, construct validity, authenticity, interactiveness, and washback are typically prioritized, practicality is no less important. In models for developing useful language assessments, such as Green's (2013b), practicality forms the foundation. An impractical assessment, regardless of its other qualities, cannot endure long-term.

Brown (2004) offers guidelines to help teachers ensure the practicality of classroom-based assessments. He suggests organizing administrative details before the assessment, ensuring students can complete the test within the allocated time, avoiding procedural issues during administration, keeping assessment costs within budget, making scoring feasible within the teacher's timeframe, and pre-determining methods for reporting results.

The principles of language assessment and testing discussed in this chapter apply to both high-stakes and classroom-based assessments. However, the emphasis on these principles varies depending on the stakes involved: high-stakes assessments prioritize them more than classroom-based ones. Nonetheless, teachers are urged to achieve acceptable levels of reliability, construct validity, authenticity, and interactiveness in their classroom assessments. They should also aim for practical assessments and promote positive washback by providing feedback that enhances student learning.

3. Classroom-based language assessment

The term classroom-based assessment is frequently used interchangeably with teacher-based assessment, school-based assessment, alternative assessments, formative assessment, and assessment for learning (Davison & Leung, 2009). However, classroom-based assessment is a broader concept, encompassing both formative and summative assessments (Hill & McNamara, 2012). Despite their differences, these terms all refer to teacher-mediated, context-specific, and classroom-embedded assessment activities (Davison & Leung, 2009, p. 395). Classroom-based assessment involves “any reflection by teachers (and/or learners) on the qualities of a learner’s (or group of learners’) work and the use of that information by teachers (and/or learners) for teaching, learning (feedback), reporting, management, or socialization purposes” (Hill & McNamara, 2012, p. 396).

This definition shows that classroom-based assessment includes both summative and formative functions (Boraie, 2018; Mathew & Poehner, 2013). Some experts argue that it should prioritize formative over summative assessment (Green, 2018; Migliacci, 2018). Despite these

recommendations, classroom-based assessment is still largely dominated by summative assessment (Buhagiar, 2007). Rea-Dickins (2004, p. 249) noted that teachers often prioritize “formal” and “procedural” aspects of assessment and overlook observation-driven approaches. Classroom-based assessment should not only include planned formative assessments but also what Hill (2017, p. 3) calls “the less visible types of assessment which occur spontaneously, in real-time, during routine classroom interactions.”

4. Conceptions of assessment

Research into educational assessment has also examined teacher conceptions of assessment (e.g., Lam, 2019; Sultana, 2019), which are crucial for assessment literacy (Xu & Brown, 2016). These conceptions influence how teachers comprehend, interpret, and apply assessment knowledge (Barnes et al., 2015; Fives & Buehl, 2012). Thus, the effectiveness of professional development programs and the successful implementation of innovative assessment policies depend on teachers’ conceptions of assessment (Brown, 2008).

In educational assessment literature, terms like beliefs, conceptions, and views are often used interchangeably (Pajares, 1992). Thompson (1992, p. 30) defines conceptions as "a more general mental structure, encompassing beliefs, meanings, concepts, propositions, rules, mental images, preferences, and the like." The term conception is used in this study because it includes both knowledge and beliefs (Opre, 2015), allowing researchers to treat knowledge and beliefs as a unified construct. This provides a crucial framework for analyzing teachers' "overall perception and awareness of assessment" (Barnes et al., 2015, p. 285).

Brown (2004; 2008b) proposed a conceptual framework of assessment conceptions, comprising three major purposes and one counter-purpose: improvement, school accountability, student accountability, and irrelevance. The improvement conception posits that assessment should enhance classroom instruction (Black, Harrison, Lee, Marshall, & Wiliam, 2003; Black & Wiliam, 1998) and is linked to formative and diagnostic assessments (Barnes, Fives, & Dacey, 2015).

The school accountability conception holds that assessment should evaluate the effectiveness of schools and teachers (Butterfield, Williams, & Marrs, 1999), and how well resources are utilized (Brown, 2004). Schools and teachers can be rewarded or penalized based on their success or failure to meet government standards (Nichols & Harris, 2016; Opre, 2015). The key idea here is that schools and teachers must demonstrate high-quality instruction and continuous improvement (Brown, Lake et al., 2011, p. 211).

The student accountability conception suggests that assessment should hold students responsible for their learning by grading their work, comparing it to performance standards, reporting grades to parents and other stakeholders, and awarding certificates based on achievement (Harris & Brown, 2009; Segers & Tillema, 2011). While some teachers see high-stakes assessments as motivating, others believe they negatively impact students (Brown, Lake, et al., 2011).

The irrelevance conception views external summative assessments as disconnected from the learning and teaching process (Brown, 2004). Assessment is seen as irrelevant if it detracts from teaching and learning, is perceived as unfair, invalid, unreliable, or if it is conducted but not used meaningfully (Harris, Irving, & Peterson, 2008, p. 3). It is also deemed irrelevant if done merely to comply with regulations (Harris & Brown, 2009).

Remesal (2011) proposed a continuum model of assessment conceptions, ranging from the pedagogical (improvement) conception to the societal conception (teacher accountability and certification). Mixed conceptions of assessment are placed in between these two ends.

Although assessment conceptions are discussed individually, teachers often hold multiple, sometimes conflicting, conceptions simultaneously (Barnes et al., 2017; Fives & Buehl, 2012). This multiplicity arises because educational assessment serves multiple concurrent purposes (Brown, Lake, et al., 2011).

5. Overview of education and assessment systems in Indonesia

Assessment measures student learning outcomes (Rokhim et al., 2021). Rosidah et al. (2021) state that assessment is integral to the learning process as it determines the quality of educational activities. It aims to evaluate and monitor learning processes, progress, and continuous improvement and is used by the government for educational policy-making (Wilson, 2018). According to Minister of Education and Culture Regulation Number 23 of 2016, educational assessment standards define the criteria for assessing student achievements in primary and secondary education, including scope, objectives, benefits, principles, mechanisms, procedures, and instruments (Permendikbud, 2016). An objective assessment requires tools like tests to accurately measure learning outcomes (Arifin, 2011).

In 2019, the Minister of Education and Culture announced that the National Examination (UN) would be replaced by the National Assessment (AN) in 2021 as part of the Freedom of Learning Program. This new assessment aims to shift the focus from evaluating student achievements to assessing and mapping the education system's inputs, processes, and results. The government hopes this policy will encourage educational improvements in subsequent years (Ministry of Education and Culture, 2019; Nurjanah, 2021).

Assessment is also a key component of curriculum implementation, used to measure and assess competency levels. Indonesia has undergone over ten curriculum changes since its independence, from the 1947 Learning Plan to the 2020 Merdeka Curriculum. Under Minister Nadiem Makarim, the Merdeka Curriculum reflects the country’s evolving needs. However, frequent curriculum changes, three in less than ten years, present challenges for students, teachers, and other stakeholders (Ministry of Education and Culture, 2019; Nurjanah, 2021; Rosidah et al., 2021). The Merdeka Curriculum aims to create a positive learning environment for teachers, students, and parents (Nasution, 2021).

Since the 2021–2022 school year, the Merdeka Curriculum has been implemented in Penggerak Schools, bringing significant changes to the learning process. Research by Surasih et

al. (2022) indicates that this curriculum helps develop students who are noble, independent, participatory, unique, and innovative.

However, teachers face challenges in implementing assessments under the Merdeka Curriculum. Cristy (2017) found that these issues stem from inadequate socialization and training for teachers. Syaifuddin (2016) identified factors such as uneven distribution of training, lack of focus on assessment in training materials, large student numbers, and limited evaluation time as contributing to difficulties in applying assessments.

CHAPTER III
RESEARCH METHODOLOGY

A. Research Setting

The researcher selected 10 English teachers who are specifically majoring in Department of English Education at State University in Banten. The researcher aims to extend this research with diverse views, leveraging the University's expertise in distance learning. Due to their large student population, cultural and background disparities may provide additional challenges alongside the primary subjects.

B. Research Design

This study utilized a qualitative descriptive survey research approach. Survey research involves a group of individuals responding to questions through interviews, questionnaires, or tests (Fraenkel et al., 2012). The survey research process generally consists of four stages: preparation, data collection, data analysis, and drawing conclusions.

The data gathered in this study is qualitative, derived from respondents' perspectives through open-ended questions. The data collection method used was interviews with respondents. Interviews are a common and powerful method for collecting data in qualitative research (Indrawami, 2022). Fontana and Frey (2000) describe interviews as “one of the most common and powerful ways in which we try to understand our fellow human beings” (p. 645). The interviews lasted between 20 to 30 minutes, with all interviews recorded with the interviewees' permission. The interview protocol included several open-ended questions aimed at exploring participants’ training experiences in language assessment and their conceptions of assessment. Open-ended questions are advantageous because they allow respondents to express their personal experiences in their own words (Christensen, 2016). To ensure reliability, the same interview protocol was consistently used with all interviewees (Cohen, Manion, & Morrison, 2002).

C. Research Participants

The research participants will consist of an undivided group of English teachers who are currently postgraduate students from the English Education Department. These students hail from various provinces in Indonesia and are pursuing a master's degree in language assessment. Approximately half of the participants are language teachers pursuing their master's degree on a part-time basis. The majority of those not currently teaching had prior experience as language teachers before embarking on their master's degrees.

D. Data Collecting Techniques

Data collection was conducted in March 2024 using in-depth interviews. These interviews were held via Zoom at times and locations convenient for the participants, with each interview scheduled on a different day. During the interviews, field notes were taken to capture key points made by the participants. After transcription, each transcript was sent back to the respective participant for member checking, allowing them to review their responses before and after translation into English. This process was intended to prevent any misunderstandings or misinterpretations of their words (Mertens, 2005). All participants reviewed and agreed to the transcriptions without requesting changes. Additionally, a fluent Bahasa-English speaker conducted member checking to ensure the accuracy of the quotes used in the findings.

E. Data analysis

The qualitative data were analyzed using the deductive thematic analysis approach (Creswell & Clark, 2017), following Braun and Clarke's (2006, p. 87) guidelines. The data were transcribed verbatim and carefully re-read multiple times. Coding involved identifying text segments that aligned with predefined codes such as reporting, compliance, extrinsically motivating students, school accountability, student accountability, and irrelevance. This process included highlighting and annotating data extracts.

Once coded, the extracts were organized into major themes, such as improvement and school accountability. The study utilized conceptual categories of assessment conceptions outlined by Harris and Brown (2009) and Brown (2004; 2008b). For instance, during transcript review, instances where teachers discussed assessment providing data for school administrators and subsequently to the Ministry of Education were coded under "reporting." This code was then categorized under the broader theme of school accountability. The researcher then evaluated how well the coded extracts supported each theme. Finally, excerpts were selected to exemplify the findings of the study.

CHAPTER IV

FINDING AND DISCUSSION

A. Finding

1. Teachers' Conceptions of assessment

Participants had teaching experience ranging from 2 to 15 years. The analysis categorized data according to the frameworks outlined in Brown (2004; 2008b) and Harris and Brown (2009). Six distinct categories of assessment conceptions emerged from the data: reporting, motivating students through external rewards, enhancing teaching and learning, holding students accountable, compliance, and perceptions of assessment as irrelevant.

1.1.Structure of teachers' conceptions of assessment

Initially, Brown's (2004; 2008b) model of assessment conceptions was used to analyze teachers' conceptions, but it did not align well with the data. Therefore, Remesal's (2011) framework, which better suited the data, was adopted while maintaining the primary categories from Brown (2004; 2008b). Remesal's continuum model places the pedagogical conception (focused on improvement) at one end and the societal conception (emphasizing teacher accountability and certification) at the other, with mixed conceptions in between. This framework closely mirrored the conceptions observed among participants in this study.

Prior to categorizing teachers, with assistance from a second coder, the researcher identified each teacher's overall conception of assessment and then placed them within the four categories from Brown (2004; 2008b): improvement, student accountability, school accountability, and irrelevant. Teachers whose beliefs clearly fit into one category were classified as holding a pure conception of assessment. Those whose beliefs spanned multiple categories were categorized as holding mixed conceptions. While some participants expressed pure conceptions (such as improvement or student accountability), others articulated mixed conceptions (combining school and student accountability, improvement and student accountability, or including aspects of irrelevance).

1.1.1. Improvement conception

Two out of ten (2/10) teachers held a pure improvement conception of assessment, emphasizing its role in enhancing learning and teaching.

…Assessment is crucial for gauging students' progress, development, and identifying their educational needs. It serves as a diagnostic tool that helps teachers understand where students currently stand and how to guide them towards achieving their learning goals. Simply put, assessment checks students' comprehension of classroom content. Without assessment, it would be difficult to gauge the effectiveness of teaching (P10).

This teacher strongly advocates for assessment's ability to enhance instructional quality and expresses concern about assessments primarily used for high-stakes decisions regarding students. There appears to be a discrepancy between the teacher's personal beliefs about assessment and the prevailing assessment practices within their institution. The first teacher (P10) also views summative assessment as a critical dimension that provides feedback on the quality of learning and teaching.

…Additionally, I believe that summative assessment is important for both teachers and students. It determines who passes and who fails, but it also offers insights into how well students have learned throughout the course or program, as well as how effective the teaching has been. This feedback is valuable for improving teaching and learning (P10).

The teacher's perspectives on assessment appear to align with the current secondary school assessment policy, which encourages the use of both summative and formative assessments to support effective learning and teaching practices.

1.1.2. Mixed accountability and improvement conception

Out of ten teachers, five exhibit mixed conceptions of assessment, where they perceive assessment as serving both accountability functions for students and schools, as well as

facilitating improvements in teaching and learning. However, their perspectives are not evenly balanced. Three teachers prioritize the accountability aspects, while the remaining two emphasize the formative functions. Examples from each category illustrate these viewpoints.

The first example pertains to a 38-year-old teacher with 13 years of teaching experience. His beliefs about assessment lean more towards school and student accountability rather than improvement. He emphasizes the roles of reporting and making high-stakes decisions about students.

…Tests are used to report student performance and determine who passes and who fails. A score of 9 means a student fails, while 10 means a pass. Testing is essential to gauge how much students have learned. It's the only way to assess their learning outcomes and determine their academic standing (P1).

The teacher also places significant importance on enhancing students' test scores, which he views as a measure of his effectiveness to parents and the school administration.

…At the end of each term, we must report test results to the headmaster and parents. The headmaster reports to the Ministry of Education to show that we've met our targets. Test results demonstrate our accountability as teachers for student learning. We're expected to achieve these goals and show good results. Student grades also influence the headmaster's assessment of our performance at the end of the year (P1).

In contrast, the second case involves a 39-year-old teacher with 10 years of experience who emphasizes the improvement function of assessment. Unlike the previous teacher, he believes assessment primarily tracks students' learning progress towards educational goals and informs teaching strategies.

…Assessment allows teachers to evaluate how well students have understood the material taught. If students haven't grasped the content, I adjust my teaching approach. Assessment helps me decide whether to move forward or review the topic again. For example, when

teaching the present simple, I assess students through exercises to gauge their understanding. This is the main role of assessment—it guides instructional decisions (P5).

While the teacher recognizes the role of examinations in assessment processes, he expresses reservations about their impact. He believes examinations can pressure teachers to rush through lessons to cover exam content, detracting from deeper learning experiences for students.

…I also assess my students at the end of units, chapters, or semesters as required. These tests determine who passes and who doesn't. They provide feedback on school and teacher performance to stakeholders. However, sometimes exams force us to rush through lessons to finish the curriculum by year-end. This rush limits opportunities for assessment, feedback, and ensuring student comprehension. This focus on completing the curriculum for exams rather than on learning outcomes contributes to students graduating without sufficient English proficiency (P5).

1.1.3. Mixed student accountability and irrelevant conception

Three out of ten (3/10) teachers appeared to perceive assessment as simultaneously holding students accountable while also considering it irrelevant. All three participants leaned more towards emphasizing student accountability rather than seeing assessment as irrelevant. Two detailed examples were provided to illustrate these perspectives.

In the first case, a 34-year-old teacher with 9 years of experience views assessment primarily as a tool for categorizing students based on their performance. However, he feels conflicted about its purpose, especially when pressured by school administrators to adjust test results to meet Ministry of Education pass rate expectations.

Assessment is used to categorize students, but this categorization is influenced by school administrators... In public schools, in a class of fifty students, teachers must ensure that more than 90% of students pass so that our superiors can demonstrate their effectiveness. Therefore,

assessments can sometimes seem paradoxical or purposeless because if students receive low grades, teachers are required to adjust them, or they will be perceived as ineffective... When I first started teaching, I remember presenting test results that were lower than what the principal wanted, so he asked me to adjust them. Initially, I resisted, but then I had a meeting where I was told that if I wanted to keep my job, I had to adjust the grades... So now we assess because we have to, not because it serves any meaningful purpose (P7).

Similarly, in the second instance involving a 42-year-old teacher with 18 years of teaching experience, assessment is seen as a tool for making crucial decisions about students' academic outcomes. However, faced with unrealistic pass rate expectations set by educational authorities, the teacher has resorted to conducting formal assessments merely to comply with school or Ministry of Education directives.

Assessment determines who passes and who doesn't, but in our school, I know these students cannot communicate in English. In a class of 50 students, perhaps only 20 can achieve good results. This means fewer than 50% of students achieve high marks. Even primary school teachers are pressured to produce results desired by the principal, who follows instructions from the Ministry of Education. Showing these results to the principal would lead to problems. I cannot report the true results. Therefore, there is no real reason to conduct assessments. I give out papers, but I do not even check them. If I showed the true results to the principal, I would face consequences. I must produce results, not tests. As a teacher, I conduct assessments because it is mandatory, but I observe the classroom and decide which students deserve to pass based on their performance. Quiet students may fail because they do not demonstrate their abilities (P3).

This teacher's beliefs about his students' academic capabilities influence his assessment practices. He conducts formal assessments to adhere to school policies, but they do not serve a meaningful purpose. Grades are assigned based on informal observations, such as participation and effort, enabling most students to pass and allowing the teacher to achieve the required pass rate.

1.1.4. Student accountability

Two out of ten (2/10) participants expressed a clear conception of assessment centered on student accountability. They view assessment primarily as a tool for making critical decisions about students' academic progress and qualifications.

Assessment is used to measure students' learning. For instance, when focusing on reading, I must assess their reading abilities, and similarly for grammar, their performance must be assessed at a certain level. I believe the purpose of assessment is to determine who passes and who fails based on their scores. A score of 10 signifies a pass, while 9 indicates a fail (P6).

The teacher's beliefs about assessment appear to have been influenced by both his experiences as a student and his teaching career. At 37 years old, it can be inferred that he encountered traditional assessment practices during his own schooling before recent changes in elementary and secondary education policies. He also perceives assessment as instrumental in selecting and placing students in classes or educational levels that align with their academic capabilities.

When I taught third grade, some students demonstrated higher abilities than others. Assessment helped me decide whether a student should advance to the next grade level or not. It determines their suitability for different classes based on their unique abilities. Keeping a student back would not be beneficial; it would hinder their progress. Assessment plays a crucial role in facilitating these decisions (P6).

1.2. The factors which contribute to teachers' understanding and use of assessment in learning

The process of analyzing core concepts included coding, consolidating codes, and labeling and refining themes. The interview data showed that teachers categorized their perceptions of assessment into two main types: internal and external assessments. Their interpretations of these assessment types are detailed, along with relevant quotes, in the subsequent section. To facilitate

coding, the analysis began with a case study focused on improvement, followed by one on irrelevance, and concluded with a case on accountability.

1.2.1. Improvement Teachers

P1, P2, P3, and P4 exemplify teachers who prioritize the role of assessment in enhancing teaching and learning. Their survey responses underscored their strong support for using assessment to aid student development, promote learning accuracy, and fulfill accountability objectives, while expressing uncertainty about its potential irrelevance. These participants exhibited a consistent pattern in their views.

The teachers' assessment values can be categorized into two main types: internal and external assessments. These categories also served as the basis for organizing themes related to irrelevance and accountability. Internal assessment within the context of enhancing teaching and learning encompasses themes such as readiness for change, fostering educational values, authentic use of formative assessment practices, grading to demonstrate achievement, and teacher autonomy.

1.2.1.1.Internal assessment

IM teachers preferred internal assessment methods that allowed them to adapt their teaching approaches, instill values, and employ diverse assessment strategies. Their views on assessment reflected openness to change, a belief in its role in promoting positive values, and a commitment to grading students. IM teachers viewed assessment primarily as a tool to inform and improve their teaching practices. They often reflected on students' assessment results to gauge the effectiveness of their instruction:

"…[low scores] sometimes make me disappointed; I thought I'd taught them well." (P1)

"I question myself, is my teaching effective? Do my students understand the lesson?" (P3)

These teachers were concerned about their teaching effectiveness and showed readiness to adjust their methods accordingly. They emphasized formative assessment as a means to continuously inform their teaching practices. They actively engaged students in the assessment process, seeking their feedback to enhance the learning experience:

"Every semester I ask my students to comment on my teaching…do they like the strategies I use, how they want the learning process run." (P2)

IM teachers also saw internal assessment as a way to build stronger connections with students, taking responsibility for motivating and supporting those who struggled academically:

"I try to motivate them, I give them feedback, I ask what causes such unsatisfactory results, what is the problem, which part is hard and so on…So I assist them to realize reasons behind their failure." (P4)

They found that assessments, such as tests and assignments, not only motivated students but also served as effective feedback mechanisms to make learning engaging and meaningful. IM teachers believed in using assessment to promote positive values among their students and emphasized the importance of formative assessment aligned with their constructivist teaching approach.

IM teachers structured their teaching and assessment around the four language skills—reading, writing, listening, and speaking—using a modified scoring grid to track student progress across these skills:

"…I measure speaking, listening, reading, and writing. So, students are scored based on these skills." (Lisa, p.4, 2012)

They viewed grading as integral to assessing student learning outcomes and argued that it accurately measured students' comprehension and proficiency:

"…to measure students' proficiency, how well they comprehend the lesson." (Intan, p.1, 2012)

IM teachers believed that grading provided a clear indication of students' achievement levels and emphasized its importance in communicating learning outcomes effectively:

"…it is a sign, whether students master the teaching materials or not, if they get 10 (ten), it means they understand the lesson well." (P3)

Overall, IM teachers saw assessment as inseparable from grading, viewing it as a vital tool for assessing and communicating student achievement within their educational context.

1.2.1.2.External assessment

IM teachers acknowledged the importance of exam-based external assessments in certifying students' learning, evaluating teachers' effectiveness, and ensuring school accountability. However, they also expressed concerns about the impact of external exams on teacher autonomy and equitable treatment of students, as well as doubts about the reliability of these assessments. These conflicting views were evident in their discussions on conceptions of assessment, teacher autonomy, and the credibility of external assessment practices.

IM teachers believed that external assessments could hold teachers accountable by reflecting the quality of their teaching:

"It tells me whether I am able to transfer knowledge to my students." (P1)

This perspective indicated a belief in more traditional, behaviorist views of learning, where student success is measured by their ability to reproduce teachers' knowledge. IM teachers viewed assessment as a powerful tool to communicate the standard and effectiveness of both student learning and teaching:

“...they [parents] look at the result. When it is good, it means the teachers are qualified.” (P2)

Additionally, IM teachers agreed with the government's use of external assessments to evaluate school quality. They understood that these high-stakes exams served to rank schools at regional, provincial, and national levels:

“To determine the quality of students, the school, the region, the province, through assessment we can measure the percentage of quality improvement.” (P2)
“The students’ proficiency indicates the quality of a school.” (P4)

However, IM teachers expressed reservations about recent changes where regional governments played a dominant role in creating semester tests, excluding input from rural teachers. This shift diminished teachers' trust in the fairness and accuracy of the assessments:

“...in developing the semester test... the department [of education] did not invite teachers from this [rural] area.” (P4)

IM teachers argued that they were best suited to understand their students' abilities and should have a significant role in designing and administering assessments. They believed that local teachers could create more relevant and reliable tests compared to assessments imposed by regional authorities:

“I think it would be better if the department returned the making of tests to teachers.” (P1)

Moreover, IM teachers expressed concerns about the credibility of external assessments, citing instances where they suspected unfair practices such as leaked answers:

“I have witnessed suspicious practices; sure, I did not make up this story, two students had identical answers with the key, including the words and commas. I know one student well; his competence is not at that level.” (P2)

They questioned the integrity of the assessment system and felt that localized assessments conducted by subject-specific teachers would be fairer and more trustworthy:

“I am sure, it’s impossible [for students to answer all questions correctly] …there must be a conspiracy.” (P3)

In summary, IM teachers recognized the role of external assessments in certification and accountability but raised significant concerns about their impact on teacher autonomy, equity, and credibility. They advocated for more localized assessment practices that involve teachers directly in the assessment process to ensure fairness and reliability.

1.2.2. Irrelevance group (IR)

P5, P6, and P7 are classified as holding conceptions of assessment that deem it irrelevant. Their responses exhibited inconsistencies across the three types of assessment conceptions. These inconsistencies suggest a belief that assessment lacks significance, emphasizing their concerns about its inaccuracies, their perceived lack of understanding regarding assessment outcomes, and their view that assessment minimally or negatively affects teaching practices.

1.2.2.1.Internal assessment

IR teachers' belief in internal assessment encompassed their views on adapting teaching methods, instilling values, employing authentic summative assessment practices, and utilizing grading as a motivational tool. They viewed assessment as instrumental in adjusting their teaching approaches, both during classroom interactions and after conducting internal tests. They saw internal assessment as a means to evaluate the effectiveness of their teaching:

"Assessment helps me evaluate the quality of my teaching." (P5)

Their approach involved monitoring student performance during tasks and responses to instructions:

"When I notice limited participation from students, I interpret it as the lesson being too challenging, prompting me to switch to easier, more approachable materials." (P6)

IR teachers believed that conducting regular class tests provided insight into whether they needed to revisit lessons or offer remedial classes:

"It tells me whether I should re-teach the same lesson or offer additional support sessions." (P7)

However, these adjustments were often superficial, lacking in-depth analysis of students' low scores or critical reflection on teaching methods:

"I re-teach the same material using the same strategies." (P5)

This approach suggested a reluctance among IR teachers to challenge themselves with more effective teaching strategies or engaging classroom activities. They seemed resigned to minimal impact of assessment on teaching outcomes, reflecting pessimism about students' potential for improvement:

"Even after repeating lessons and offering retests, their scores remain unchanged." (P6)
"We know our students' competency levels; retesting or re-teaching won't lead to improvement." (P7)

IR teachers believed that internal assessment could foster positive learning attitudes such as discipline and confidence:

"Students' willingness to participate and engage, whether their answers are correct or not, signals their desire to learn." (P5)

Despite their belief in the motivational aspect of assessment, IR teachers expressed skepticism about its effectiveness in instilling values:

"Assessment should teach fairness, but it doesn't." (P7)

IR teachers held contradictory views on external assessment, often citing concerns about its credibility:

"Cheating during exams undermines the values we try to teach." (P6)

They observed that students' behavior during external exams, such as reliance on answer keys, undermined the integrity of the assessment process:

"They only study for the first day of the exam; once they have the answer key, they stop studying." (P5)

This skepticism extended to the broader impact of assessments on character education and their preference for traditional assessment practices focused on preparing students for exams:

"Why assess those skills like listening and speaking when they aren't reported or tested?" (P7)

Authenticity in assessment for IR teachers meant aligning teaching with skills tested in exams, primarily reading and writing, while relying on observational methods to gauge student competence:

"Observation gives us genuine insights into students' abilities without the need for formal tests." (P6)

Their preference for traditional assessment methods included essay formats and scoring practices that emphasized achievement and recognition among students:

"Students like to show off their scores; it's a source of pride for them." (P7)

In grading practices, IR teachers focused primarily on circling incorrect answers without providing additional feedback, viewing grading as a pivotal tool for accountability and student motivation:

"Grading is essential for accountability and encourages students to engage in learning."

External assessment

IR teachers held conflicting views on external assessments. While they recognized the importance of external examinations for assessment purposes, they also criticized them for lacking credibility and being intimidating. This category encompasses themes of conflicting beliefs, teacher autonomy, and the reliability of external assessment.

IR teachers acknowledged the role of external assessment in demonstrating accountability and evaluating educational quality. However, they voiced concerns about the process and implementation of these assessments:

"The process is deceptive." (P5)

They argued that assessments did not accurately reflect a school's quality of education:

"It's artificial...the Department of Education claims assessments showcase educational quality in the region, but everyone knows scores will be inflated...every student must pass." (P7)

IR teachers observed that schools sometimes manipulated assessment results to protect their reputation:

"To maintain their good image...if a school has a reputation to uphold, low student scores will attract public scrutiny." (P6)

Despite their disagreement with such practices, IR teachers felt compelled to comply with the system:

"We're obligated to do it...the school principal directs us, and they in turn are directed by higher authorities." (P7)

IR teachers criticized the Department of Education for prioritizing regional reputation and striving for recognition as successful:

"The Department is always eager to be seen as successful." (P6)

This environment led to complex feelings among teachers about assessment, diminishing its credibility and eroding teachers' autonomy. They reported systemic interference in regional and national testing processes:

"Teachers are instructed by the school principal to assist students before exams." (P7)

Feeling powerless, IR teachers described marking students' work only to see their efforts disregarded:

"All our efforts are disregarded. We must follow the directives of the school principal." (P5)

IR teachers felt their autonomy undermined, leading them to question the meaningfulness of assessment:

"I doubt my abilities as a teacher. I feel inadequate." (P6)

This perception affected their relationships with students, who they felt perceived them differently due to the pressured manipulation of scores:

"I feel students mock me, expecting assistance all the time." (P6)

IR teachers contended that inconsistent policies and practices within assessments created a sense of conspiracy within the school community. They described pressure to present favorable scores to maintain the school's reputation:

"We're pressured to inflate scores...we're not allowed to report true scores." (P6)

Despite these challenges, IR teachers relied on their own judgment to determine appropriate adjustments to scores for student reports based on authentic assessments and daily observations:

"We use our own judgment to decide on the appropriate final scores for student reports." (P6)

In summary, IR teachers questioned the validity of standards used in external assessments. They expressed discomfort with policies governing regional and national exams and noted disparities between educational qualities in urban and rural areas:

"A 9 in my context is like a 6 in a city school." (P6)

This disparity underscored their concerns about the fairness and consistency of external assessments across different school settings:

"When I compare assessments here with those in the city where my daughter studies, they are vastly different." (P5)

Accountability group (AC)

P8, P9, P10 embody accountability perspectives on assessment. These teachers prioritize accountability as their primary viewpoint, followed by improvement perspectives, and they tend to reject irrelevance conceptions. Teachers embracing accountability views believe that assessment serves as a reliable tool to establish the accountability of schools or countries in fulfilling educational responsibilities. They argue that assessment effectively measures the capacity of teachers and schools to enhance the quality of teaching and learning. Similar to teachers in the

improvement and irrelevance categories, those in the accountability group also grapple with conflicting beliefs regarding internal and external assessment, which give rise to the central themes within their case.

Internal assessment

The internal assessment category encompasses various themes including openness to change, value development, varied perspectives on assessment, and teaching resources and grading practices.

Similar to their peers in the IM and IR groups, AC teachers utilize assessment data to adapt their teaching. They demonstrate flexibility and are open to seizing spontaneous teaching opportunities:
"I welcome and apply a sudden bright idea that comes." (P8)

AC teachers employ diverse teaching resources beyond textbooks, customizing their instructional materials to support student learning even if they are not directly tied to external exams:
"While I base my teaching journal on the textbook, students' activities come from many sources." (P8)

These teachers value internal assessment for its dual benefit to both teachers and students, fostering improved student-teacher relationships. They personalize their approach to cater to individual student needs, recognizing and addressing difficulties:
"I usually ask students to write down challenging materials or identify those needing re-teaching, mainly interacting and assisting them on a personal level." (P9)

AC teachers also emphasize values such as discipline and creativity, celebrating instances where students' creative outputs exceed expectations:
"Once I tasked my students with writing a letter, and their creativity in designing and presenting ideas was impressive." (P10)

They advocate that assessment serves as a motivational tool for students:
"Assessment motivates students to study; it sparks their enthusiasm." (P8)

However, like their counterparts in the IM and IR groups, AC teachers question whether students can apply these values effectively in external assessment settings, expressing concerns over fairness and integrity:
"Students' fairness in [external] examinations is questionable; some resort to cheating." (P10)

In alignment with their belief in accountability through assessment, AC teachers value both formative and summative assessment practices. They encourage student autonomy and engagement through performance-based assessments and peer evaluations within classroom settings:
"I ask students to compose stories and develop outlines based on their ideas." (P8)

Despite their emphasis on internal assessment practices, AC teachers tend to conform to standardized assessment formats used in external examinations such as essays and multiple-choice questions:
"We prioritize formats like essays and short answers that are also prevalent in external exams." (P9)

They acknowledge the importance of observation as a reliable measure of student performance in internal assessments:
"We evaluate their actual performance through class interactions." (P8)

However, AC teachers perceive limitations in their ability to implement certain assessment types due to inadequate resources or training, such as the absence of language laboratories or insufficient familiarity with available equipment:
"We never received training on how to use the laboratory equipment available in other schools." (P8)

Overall, AC teachers exhibit a nuanced understanding of authentic assessment, blending internal practices with compliance to external assessment standards. They emphasize the impact of grading

practices on student motivation and learning outcomes, underscoring the complex interplay between assessment beliefs and educational practices in their community.

External assessment

AC teachers, like their counterparts in other categories, held conflicting views regarding external assessment and expressed concerns about teachers' autonomy and the credibility of tests.

The belief in accountability within assessment was evident in AC teachers' emphasis on practices that measured students' proficiency:

"We aim to achieve specific goals in the curriculum or students' achievement levels, and only assessment can provide us with that information." (P9)

AC teachers viewed assessment as essential for gauging learning outcomes. One teacher queried, "How can we determine a student's capabilities without assessing them?" (P10)

For AC teachers, assessment served as a tool to evaluate students' understanding and proficiency levels:

"I can gauge how well my students grasp the material; it serves as a benchmark to assess their proficiency." (P8)

Many AC teachers believed that a school's reputation was closely tied to its students' performance scores. They regarded high-stakes exams as crucial because they influenced the school's standing or ranking:

"...the government receives information or reports on which region or province performed the best this year." (P8)

AC teachers highlighted negative impacts of external assessments, whether conducted regionally or nationally, on their professional autonomy. They critiqued various aspects, including the Education Department's perceived lack of trust in teachers' ability to design tests and expressed feelings of intimidation. Specifically, they resented the regional Education Department's administration of semester exams, which they felt undermined teachers' professional authority:

"I feel like the Education Department questions our capabilities...they don't trust us to manage our own exams." (P10)

This sentiment was exacerbated by the perception that teachers were excluded from the development of semester tests, which were supposed to be teacher-led:
"Many teachers question why they are excluded from developing semester tests." (P9)

AC teachers challenged regional policies regarding the management of exams, asserting that they deviated from the national government's intended assessment practices:
"I believe they [the regional Education Department] are not following the rules...in my understanding, 'government assessment' refers to assessments conducted by the Ministry of Education at the national level, not by regional Education Departments." (P10)

AC teachers expressed doubts about the credibility of external assessments, citing concerns about processes and outcomes that hinted at collusion within the school community:
"I am uncertain about the credibility of our external assessments." (P8)
Teachers were suspicious that students couldn't pass exams without external assistance:
"Everyone knows that students can't pass exams without teacher assistance." (P9)

This uncertainty led AC teachers to question the fairness and integrity of assessment practices, despite recognizing the accountability purposes of external exams. They were troubled by instances where students achieved high scores despite difficulties in learning:
"I feel conflicted...students complain about tough lessons but still manage to score 100 on exams." (P10)

AC teachers raised concerns about external tutoring institutions providing answer keys, suggesting a complex perception of external assessments. Their skepticism about assessment practices consistently underscored issues of autonomy, fairness, and community influence, which they believed compromised the integrity of assessments.

AC case summary

AC teachers strongly advocated for assessment practices that accurately reflect teaching and learning. They enthusiastically utilized assessment results to enhance their teaching effectiveness and were proponents of internal assessment. However, they held varying perspectives on external assessment. While recognizing its intended purpose to evaluate the effectiveness of students, teachers, and schools, AC teachers were disillusioned by how external assessment was implemented in their educational context. This dissatisfaction extended to its impact on students, schools, and local educational authorities, as well as other external institutions.

These teachers believed that unfair examination practices undermined students' motivation to learn. They argued that students' focus on achieving high scores without valuing the learning process compromised the positive aspects of assessment. AC teachers contended that principals' efforts to uphold the school's reputation sometimes led to unfair assessment practices, thereby muddling the intended function of assessment as a tool for accountability.

B. **Discussion**

The results also indicate that approximately half of the participants expressed clear improvement-oriented conceptions or mixed conceptions leaning towards improvement. These findings are promising as previous research has shown that teachers' beliefs about assessment significantly shape their assessment practices (Barnes et al., 2017; Xu & Brown, 2016). Moreover, they align with current secondary school assessment guidelines, which encourage educators to prioritize formative assessment over summative assessment.

Furthermore, the assessment policy itself exacerbates the tensions between summative and formative assessments. Initiatives promoting formative assessment can encourage teachers to adopt these practices (Kim, 2019). However, as noted by Brown and Remesal (2017), successful implementation of formative assessment in high-stakes assessment environments hinges on reducing the dominance of high-stakes assessments. Merely adding a new "soft" policy of formative assessment alongside existing "hard" policies of high-stakes assessments is insufficient. While the current policy emphasizes formative assessment, schools and teachers continue to be

evaluated based on students' performance in summative assessments. In such contexts, teachers often prioritize summative assessments due to the pressures of meeting accountability requirements (Brown & Gao, 2015). Addressing this issue is crucial to effectively implementing formative assessment initiatives.

Additionally, the findings underscore the shortcomings of using test results to evaluate teacher and school effectiveness. Concerns about potential consequences for failing to improve students' achievement sometimes lead school administrators to instruct teachers to inflate student test scores. This practice can distort information about teachers' quality and students' achievement levels (Morgan, 2016; Rose, 2015). Another undermining strategy is the neglect of recommended assessment and grading practices. Some educators rely solely on informal assessments and non-academic factors like participation and effort, rather than assessing academic performance against established standards. While this approach may inflate success rates, it can also provide misleading insights into students' actual achievements. Previous studies have highlighted that while teachers consider various factors in their assessment and grading practices, academic performance remains paramount (McMillan, 2001; McMillan, Myran, Workman, 2002).

CHAPTER V

CONCLUSION AND SUGGESTION

A. Conclusion

This study offers valuable insights into teachers' perspectives on assessment. Its primary contribution lies in presenting a model that illuminates assessment conceptions as a nuanced and multifaceted process. Similar educational, contextual, or cultural backgrounds are likely to yield comparable perceptions. Importantly, the conceptual framework encourages researchers to seek models tailored to specific contexts rather than adopting a universal international model.

The research findings reveal that various factors within the teaching environment intricately shape participants' assessment conceptions. This study underscores that even a globally validated survey, supplemented by replicated studies on assessment conceptions, may not uniformly apply across different educational settings. A key takeaway is the recognition that one standardized approach cannot universally accommodate all contexts. For instance, while the TCoA may suit settings like New Zealand or Australia with their low-stakes examination environments, its applicability in high-stakes contexts such as Indonesia requires adaptation. Future assessments of teachers' assessment conceptions should therefore consider socio-ecological factors to accurately capture these perspectives. Understanding teachers' assessment conceptions and the contextual influence on these conceptions is crucial groundwork for effective policy implementation in education.

B. Suggestion

This study aims to explore how contextual factors impact teachers' perceptions of assessment and whether these perceptions align with those observed among participants in this study. Future research employing the proposed components of the new conceptual framework would facilitate the adaptation and refinement of the TCoA model to better suit diverse educational contexts.

REFERENCES

Amorim, R D F C C D K., & Silva, M J P D. (2014, December 1). Effective/efficacious nonverbal communication in the classroom: the conception of the nursing lecturer. https://doi.org/10.1590/0104-07072014001710013

Barnes, N, Fives, H, Dacey, C. (2015). Teachers' beliefs about assessment. In H Fives, MG Gill (Eds.), International handbook of research on teacher beliefs, (pp. 284–300). New York: Routledge.

Brown, G. (2004). Teachers' conceptions of assessment: implications for policy and professional development. Assessment in Education: Principles, Policy & Practice, pp. 11, 301–318 https://doi.org/10.1080/0969594042000304609.

Brown, G. (2008). Conceptions of assessment: understanding what assessment means to teachers and students. New York: Nova Science Publishers.

Brown, GTL, Hirschfeld, GHF. (2008). Students' conceptions of assessment: links to outcomes. Assessment in Education: Principles, Policy & Practice, 15(1), 3–17.

Brown, GTL, Pishghadam, R, Sadafian, SS. (2014). Iranian university students' conceptions of assessment: using assessment to self-improve. Assessment Matters, 6, 5–33.

Chan, R C., Wang, Y., Li, H., Shi, Y., Wang, Y., Liu, W., & Huang, J. (2010, January 1). A 2-STAGE FACTOR ANALYSIS OF THE EMOTIONAL EXPRESSIVITY SCALE IN THE CHINESE CONTEXT. https://doi.org/10.2117/psysoc.2010.44

Coombe, C., Vafadar, H., & Mohebbi, H. (2020, May 26). Language assessment literacy: What do we need to learn, unlearn, and relearn? https://doi.org/10.1186/s40468-020-00101-6

Cortazzi, M, Jin, L, Zhiru, W (2009). Cultivators, cows, and computers: Chinese learners' metaphors of teachers. In Internationalising the University, (pp. 107–129). London: Palgrave Macmillan.

Cumming, A. (2009, March 1). LANGUAGE ASSESSMENT IN EDUCATION: TESTS, CURRICULA, AND TEACHING. Annual Review of Applied Linguistics, 29, 90–100. https://doi.org/10.1017/s0267190509090084

Douali, L., Selmaoui, S., & Bouab, W. (2022, August 26). The Pitfalls in Assessment Practices in Distance Learning. International journal of technology in education, 5(3), 486-498. https://doi.org/10.46328/ijte.238

Entwistle, NJ, & Entwistle, A. (1991). Contrasting forms of understanding for degree examinations: the student experience and its implications. Higher Education, 22(3), 205–227.

Fives, H, & Buehl, MM (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In KR Harris, S Graham, T Urdan (Eds.), APA educational psychology handbook: individual differences and cultural and contextual factors, (vol. 2, pp. 471–499). Washington, DC: APA.

Fulcher, G. (2010). Practical language assessment. London: Hodder Education.

Fulcher, N G. (2012, April 1). Assessment Literacy for the Language Classroom. https://doi.org/10.1080/15434303.2011.642041

Fulmer, GW, Lee, ICH, Tan, KHK. (2015). A multi-level model of contextual factors and teachers' assessment practices: an integrative review of research. Assessment in Education: Principles, Policy, and Practice, 22, 475–494. https://doi.org/10.1080/0969594X.2015.1017445.

Hakim, B. (2015). English language teachers' ideology of ELT assessment literacy. International Journal of Education & Literacy Studies, 3(4), 42–48.

Jin, L, & Cortazzi, M (2011). The changing landscapes of a journey: educational metaphors in China. In J Ryan (Ed.), Education reform in China: Changing concepts, contexts, and practices, (pp. 113–131). London: Routledge.

Lee, H., McAuley, J H., Hübscher, M., Kamper, S J., Traeger, A C., & Moseley, G L. (2016, April 1). Does changing pain-related knowledge reduce pain and improve function through changes in catastrophizing? https://doi.org/10.1097/j.pain.0000000000000472

López, AA (2008). Potential impact of language tests: examining the alignment between assessment and instruction. Saarbrucken: VDM Publishing.

López, MAA, & Bernal, RA. (2009). Language assessment in Colombia: a call for more teacher education and teacher training in language assessment. Profile: Issues in Teachers' Professional Development, 11(2), 55–70.

Malvini Redden, Shawna. (2017). Metaphor Analysis. 10.1002/9781118901731.iecrm0154.

Marlowe, BA, & Page, ML (2005). Creating and sustaining the constructivist classroom, (2nd ed.,). Thousand Oaks: Corwin Press.

Marton, F, & Säljö, R (1997). Approaches to learning. In F Marton, D Hounsell, N Entwistle (Eds.), The experience of learning: implications for teaching and studying in higher education, (2nd ed., pp. 39–59). Edinburgh: Scottish Academic Press.

Mirallas, C A. (2021, May 21). Students' conceptions of a scientific writing course: SFL Genre Pedagogy in an EFL context. https://doi.org/10.17058/signo.v46i86.15900

Mohamed, S, Kamis, A, Ali, N (2016). Gauging the assessment literacy of Malaysia's home economics teachers: an empirical study. Geografia - Malaysian Journal of Society and Space, 12(3), 130–138.

Muñoz, A. (2021, March 26). Peer Review as a Tool for Telecommunication Studies Teaching During COVID. https://scite.ai/reports/10.33422/etconf.2021.03.79

Nimehchisalem, V., & Hussin, N I S M. (2018, August 1). Postgraduate students' conception of language assessment. https://doi.org/10.1186/s40468-018-0066-3

Niveen, RE, Ain Nadzimah, A, Sabariah, MR. (2017). Malaysian instructors' assessment beliefs in tertiary ESL classrooms. International Journal of Education & Literacy Studies, 5(2), 29–46.

Nurdiana, N N. (2021, January 10). Language Teacher Assessment Literacy: A Current Review. https://doi.org/10.30813/jelc.v11i1.2291

Ooi, H X., Hamzah, R A., & Thien, L M. (2022, December 17). Postgraduate Students Academic Writing Skills: The Needs for Supervision and Self-esteem. https://doi.org/10.6007/ijarped/v11-i4/15168

Pavlenko, A. (2007). Autobiographic stories as data in applied linguistics. Applied linguistics, 28(2), 163–188.

Pot, A., & Weideman, A. (2015, January 2). Diagnosing academic language ability: An analysis of the Test of Academic Literacy for Postgraduate Students. https://doi.org/10.1080/10228195.2014.986665

Ramsden, P. (1997). The context of learning in academic departments. In F Marton, D Hounsell, N Entwistle (Eds.), The experience of learning: implications for teaching and studying in higher education, (2nd ed., pp. 198–217). Edinburgh: Scottish Academic Press.

Razavipour, K, & Rezagah, K. (2018). Language assessment in the new English curriculum in Iran: managerial, institutional, and professional barriers. Language Assessment in Asia, 8(1), 1–18.

Rezvani Kalajahi, SA, & Ain Nadzimah, A. (2016). Assessing assessment literacy and practices among lecturers. Pedagogika, 124(4), 232–248.

Richmond, HJ. (2002). Learners' lives: a story analysis. The Qualitative Report, 7(3), 1–14 Retrieved from https://nsuworks.nova.edu/tqr/vol7/iss3/4.

Roebber, P J. (2005, April 1). Bridging the Gap Between Theory and Applications: An Inquiry into Atmospheric Science Teaching. Bulletin of the American Meteorological Society, 86(4), 507-518. https://doi.org/10.1175/bams-86-4-507

Singh, M K M. (2014, July 30). Challenges in Academic Reading and Overcoming Strategies in Taught Master Programmes: A Case Study of International Graduate Students in Malaysia. https://doi.org/10.5539/hes.v4n4p76

Stiggins, R. (2002). Assessment crisis: the absence of assessment for learning. Phi Delta Kappan, 83(10), 758–765.

Segers, M, Dochy, F, Cascallar, E. (2006). Optimizing new modes of assessment: in search of qualities and standards. Dordrecht: Kluwer Academic Publishers.

Servant-Miklos, V. (2019, December 1). A Revolution in its Own Right: How Maastricht University Reinvented Problem-Based Learning. Health Professions Education, 5(4), 283-293. https://doi.org/10.1016/j.hpe.2018.12.005

Shepard, L A. (2000, October 1). The Role of Assessment in a Learning Culture. Educational Researcher, 29(7), 4-14. https://doi.org/10.3102/0013189x029007004

Suah, S., & Ong, SL. (2012). Investigating assessment practices of in-service teachers. International Online Journal of Educational Sciences, 4(1), 91–106.

Tong, SYA, & Adamson, B. (2015). Student voices in school-based assessment. Australian Journal of Teacher Education, 40(2), 2.

Vygotsky, LS. (1978). Mind in society: the development of higher psychological processes. Cambridge: Harvard University Press.

Xu, W J., & Zammit, K. (2020, January 1). Applying Thematic Analysis to Education: A Hybrid Approach to Interpreting Data in Practitioner Research. International Journal of Qualitative Methods, p. 19, 160940692091881–160940692091881. https://doi.org/10.1177/1609406920918810

Zainuriyah, A. (2012). A Study on English Teachers' Assessment Practices on the School-based Assessment for English Language in Malaysia. Retrieved from https://goo.gl/5RCm3X

Zolfaghari, F., & Ahmadi, A. (2016, November 7). Assessment literacy components across subject matters. Cogent Education, 3(1), 1252561-1252561. https://doi.org/10.1080/2331186x.2016.1252561

APPENDICES