Materials for Language Assessment

Materials for Language Assessment

Edited by

Asma Aftab

Cambridge
Scholars
Publishing

Materials for Language Assessment

Edited by Asma Aftab

This book first published 2025

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-0364-5693-1
ISBN (Ebook): 978-1-0364-5694-8

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

BRIAN TOMLINSON

This book is a development from the MATSDA/Centro Studi Mugello Conference held in Rufina, Italy on June 21st-22nd, 2023. The theme of the Conference was Materials for Testing and that is the focus for this book, with the chapters being written by contributors to the Conference. 'Testing' is used as a convenient term to include evaluation, assessment and examining and materials could be for the preparation, conducting or follow up to 'tests' of the ability of learners of second or foreign languages.

MATSDA is an international materials development association which I founded in 1993 to bring together researchers, teachers, materials developers and publishers in a joint effort to improve the effectiveness of language learning materials. MATSDA publishes a journal, Folio, runs materials development workshops and organises international conferences on specific topics related to significant themes and issues in the field of materials development.

The 2023 Conference focused on issues related to materials developed by teachers and by examiners. Some of the presenters were critical of typical testing, examination and research procedures for their tendency to only measure what is easily measurable, for their insistence on measuring discrete competences separately, for their conflict with many of the learning procedures in the classroom and especially for their typical lack of learning opportunities. Many of these presenters were also critical of the way that tests and examinations dictate the content and procedures of syllabuses, of coursebooks and of what typically happens in classrooms. Most of the presenters who offered critical appraisals also offered suggestions for making tests and examinations less dominant, less stressful for learners and teachers, more authentic in what they measure and how they measure it as well as much more valuable to learners in the preparation, taking and follow up to tests and examinations. Other presenters reported on classroom-based research on 'testing' and on local projects which were trialling innovative ways of testing or examining learners.

The general feeling of the Conference was that 'testing' needs to be less stressful and more valuable for the learners. We need to be rewarding learners for what they can do rather than penalising them for what they

cannot do. We need to be testing the learners' ability to communicate rather than their knowledge of discrete grammatical or lexical items. And above all we need to be providing learning opportunities in the preparation, the taking and the follow up to 'tests', whether they be informal classroom tests or high stakes examinations.

This book reports the content of some of the significant presentations at the Conference and in doing so aims to stimulate the reader to reflect upon their own practice and on how the theories, ideas and practical suggestions in the book could contribute towards it. The book is written for under-graduate students, post-graduate students, teachers, materials developers, assessors and researchers. The chapters are written to be academically rigorous but at the same time to be accessible to newcomers to the field and to experienced experts alike.

Professor Brian Tomlinson
President and Founder of MATSDA

ACKNOWLEDGEMENTS

The editor wishes to thank the authors whose chapters have been included in this volume. Their discussed principles, proposals and practices are extremely valuable, insightful and pragmatic contributions to the field of language assessment, the main theme of the volume. In addition, I will like to express my appreciation to all contributors for their extreme enthusiasm for this project and their prompt responses and revisions.

Appreciation is also due to all attendees and presenters of the MATSDA/Bilingual International School 2023 Conference in Rufina (Florence, Italy) for making the event successful. Special gratitude needs to be expressed to Brian Tomlinson, Kristin Lanese (Director Bilingual International School) and Hitomi Masuhara for their valuable endeavours without which the conference would not have been possible.

I would like to extend gratitude to Adam Rummens and the team at Cambridge Scholars Publishing for their support and patience during the creation of this volume.

Last but not the least, many thanks are due to Brian Tomlinson for giving me the opportunity to be involved in this very useful project and edit this important volume.

Asma Aftab
Lahore, Pakistan
December, 2024

INTRODUCTION

ASMA AFTAB

Current perspectives on materials, language and assessment

The key words in the title of this volume, namely "materials", "language" and "assessment", need to be expounded according to the recent applied linguistics literature so that the main theme can be comprehensively elucidated. **Materials**[1] in the context of language education are anything including texts, activities, books, videos, images, charts, and webpages which can be utilized to stimulate and support acquisition (Tomlinson, 2012). Language materials have been discussed and reviewed expansively for the past two decades. The majority of the publications have proposed principles for materials development and/or evaluation with recent focus on material application. For instance, language activities should provide opportunities to students to "to say what they think is worth saying" (Tomlinson, 2003, 440). The learners need to be motivated by extensively engaging with "relevant, motivating and engaging" texts and undertaking "communicative behavior" (Tomlinson, 2003, 439; McGrath, 2016, 115). The materials should preferably be inventive, constructive, and emotionally engaging. The tasks ideally ought to be motivating, appropriate, and focusing on higher cognitive skills (for example, "connecting, predicting, interpreting, evaluating") and creative skills (Richards, 2001; Tomlinson, 2008, 6). Importantly, the activities can preferably be meaningful and open-ended, making students interact with inspiring and thought-provoking topics and encouraging them to use their own judgements (Mukundan 2009; Tomlinson and Masuhara, 2018). Recently literature (e.gs., Garton and Graves, 2014; Fernández and Berwick, 2022) has focused on application of materials in the classrooms. Most of the disseminated research (Humphries 2014; Seferaj 2014; Bosompem 2014; Fernández and Berwick, 2022) indicates that teachers' use and the degree of adaptation of prescribed materials are based on their perspectives and/or the nature, constructs and content of formal student assessment.

[1] In this volume materials, activities and tasks are used interchangeably and as such these three terms have the same denotation.

Linguistic publications have highlighted varied theories of **language**. The latest models have accentuated that language is a tool for communication which can be perceived as a complex, purposeful process dependent on situations, communicators' sentiments and audience reactions (Ruler 2018; Hargie, 2019). According to Larsen-Freeman's (2020) appliance of complexity theory, language use is flexible, multi-facet, and cooperative.

Assessment has been one of the most profusely analyzed and researched areas in education and applied linguistics. Assessment in any educational context incorporates varied, systematic practices and techniques that are utilized to evaluate and review students' performance and progress (Durdikuliyevna, Anvarovna and Zulayho, 2019). Since the past few decades, assessment focus has "moved from what the student knows to what the student can do" (Vassiliou, Papadima-Sophocleous and Giannikas 2022, 8). Thus, assessment is no longer assumed to be limited to tests and examinations: together with the advent of formative assessment (in contrast to formal summative assessment, for instance, examinations) new types of student appraisals, such as Assessment for Learning (AFL), Dynamic Assessment (DA), and Learning oriented Assessment (LOA), have been being introduced. According to Poehner and Infante (2016), formative assessment and AFL incorporate classroom tasks which attempt to determine where the students need additional support leading to subsequent changes in teaching instructions so that students can improve in these identified areas. Thus, these forms of assessment move beyond simple measurement of learners' acquisition of relevant skills/knowledge. In most scenarios DA also commonly adopts these aims and processes, but Poehner and Infante (2016) have adopted a different outlook. They view DA as the procedure in which both learner development and assessment complement each other and are in fact correlated. Assessment through appropriate mediation and support (for instance, clarifications and prompts) helps the learners acquire the relevant features which are under focus during learning. Jones and Saville (2016, 11) assert that assessment can be termed learning-oriented if a student, instead of focusing on obtaining "a good mark", is stimulated to be dynamically engaged, and to assimilate and apply concepts. These procedures contribute towards learning as expounded by Harlen and James (1997).

A summative procedure (for example, a test) can be designed to replicate formative assessment by focusing on the key objectives and authentic techniques as utilized during classroom teaching (Dolin, Black, Harlen, and Tiberghien 2018). Assessment is vital in all educational levels but especially so at college or university levels because students at these high-

er stages need to attain highly developed linguistic competence and advanced mental abilities (Khvatova and Krutskikh 2020).

Language assessment as explored in this volume

The contributions to this book discuss language assessment from fresh perspectives while elaborating and enriching the recent concepts including AFL, DA, "assessment as learning" (Cheng and Fox, 2017, 223) and evaluation of interactional skills. At the same time, the presented analysis and practices involve characterization and use of new and innovative constructs (e.g., language potentiality) and language assessment procedures (e.g., assessing authentic acquisition and empowering learners to assess their skills). Moreover, the book enhances the understanding and perception of language and language learning beyond the conventional beliefs (such as the functional and communicative approaches, language proficiency, and communicative competence). For instance, four chapters of the book focus on the expressive power of language and language use as behaviour reflecting sensitivity and attitudes. Language learning is equated to the acquisition of skills involved in communicating appropriately and successfully.

Importantly, the handling of both language assessment and materials in this volume is unique and untypical since these two apparently independent spheres of applied linguistics are being related. In other words, while language assessment and language materials have separately been profusely explored and analysed, no other publication has attempted to comprehensively discuss perspectives on *language assessment materials*. This volume in contrast moves beyond pure theorizing, conceptualization and quantitative analysis (for instance, as under focus in Fulcher and Harding, 2021) to expounding specific language assessment practices with exemplified assessment activities. The rationales and principles behind the exclusively developed materials are comprehensively elucidated. In addition, most chapters incorporate specific classroom context-based evaluative research, applications and responses/feedback related to the presented materials. Bachman and Damböck (2018) have also proposed assessment tasks for different ages/levels, but their presentation does not incorporate learners' responses/views related to the specific materials.

The different chapters examine interesting and stimulating assessment activities which provide opportunities to the students to use their mental and creative abilities and involve them affectively. Most of these materials are flexible, contextualized, focus on authentic, natural language and/or require unpredictable responses. In short, the presented tasks reflect effec-

tive features as has been proposed in the previous publications on language materials (for example, Richards, 2001; Tomlinson, 2003, 2008; Mukundan 2009; McGrath, 2016; Tomlinson and Masuhara, 2018). In addition, chapters four to nine discuss the application of the assessment activities in classrooms and in this sense the book exemplifies and contributes to the principles related to the usage of materials.

Notably, the volume presents effective language assessment tools, procedures and ideas which can be applied and so it is valuable for both practitioners and researchers. The contributors of the book are practicing language teachers as well as experienced educators and applied linguists. Thus, their discussed principles and studies focus on specific contexts with real students. The presented procedures and assessment tasks have been tried on actual learners, and so the frameworks/tasks can be utilized and adapted in other teaching scenarios with useful consequences.

Jones and Saville (2016, 12) assert that "the strength of large-scale assessment lies in the degree of control and standardisation of the conditions under which performances are elicited and evaluated". "Large-scale assessment" implies high-stakes summative testing of a large number of candidates belonging to diverse contexts. However, as indicated in this book, testing being high-stakes, summative and/or extensive/global does not necessitate regulation and uniformity reflecting hegemony and predictability.

Successful educational programmes should aim towards attaining highest degree of cogency by aligning learning goals with teaching procedures including assessment. This level of compatibility between learning and assessment will ensure that student appraisal tasks are enhancing acquisition together with verifying the degree learning has been achieved. Similarly, Popkova (2020) states that assessment procedures should focus on educational benefits and facilitation of learning (for example, by providing learners opportunities to take risks). However, these aims are rarely given significance even when latest testing techniques are being employed. In chapters five to nine, the contributors discuss assessment tasks which reflect constructive learning goals and outcomes.

Both language and assessment are multidimensional concepts/ processes. Assessment is complex "because we cannot know with certainty what understanding exists inside a student's head" (Bennett 2011, 16); thus, at best we have to resort to deduction and interpretation while evaluating learners' abilities and performances. Moreover, when assessment involves appraisal of language usage/skills, it becomes additionally complex. As a result, language appraisals have to be based on the well-defined goals reflected in the usage of target language, and on the contexts/domains in which the students and candidates are required to com-

municate. The chapters of this book delve on language assessment in its complexity, from exploring the authentic nature of acquisition/learning and language usage (as discussed in chapters one and two) to the presentation and implementation of assessment materials meant for specific purposes, groups of learners, and contexts (as elucidated in chapters three to nine).

Detailed overview of the volume

The volume consists of theoretical expositions with exemplifications, and evaluative and application-based studies, expounding effective language assessment. The chapters present materials reflecting the discussed theoretical beliefs (e.g., the incompatibility between tests and learning); and, inventive procedures which can be used to evaluate useful constructs (like genuine acquisition, language potentiality and communicative power). The contributions highlight comprehensively key strengths and weaknesses of formal examinations, classroom-based appraisal tasks, and information technology and artificial intelligence based assessment techniques. In some cases, students' responses and feedback have additionally been disseminated. The presented principles, tasks and analysis highlight the importance of using *appropriate* procedures to assess language learners' acquisition. Since these approaches have been proposed theoretically, analysed, and implemented by theorists, researchers and practitioners belonging to varied contexts, the publication assembles diverse viewpoints and stances which are nevertheless related by their originality, authenticity, and pragmaticism.

Brian Tomlinson in chapter one introduces the main focus of this volume by highlighting and exemplifying that assessment should involve the use of effective evaluative materials which cater to language learners' needs. From this point of view, assessment activities should preferably aim towards appraising and contributing to the target students' language acquisition. In chapter two, **Alan Maley** utilizes personal anecdotes and experiences, and expertise professed by specialists, professionals, and successful educationists to provide an insightful evaluation of the general educational scenario and specific assessment practices. The concluding sections of the chapter detail some constructive and applicable measures which are being implemented to ameliorate the assessment of students.

In the third chapter **Sakae Onoda** evaluates the high-stakes Japanese university entrance examination questions from the point of view of reliability, validity and washback effects. The chapter identifies the weaknesses of examination questions which can lead to limited appraisal of the candi-

dates' academic English skills and indicates the probable negative influence on high school English teaching (for example, by placing misplaced focus on obsolete lexis and using ineffective reading tasks). Sakae concludes by presenting relatively more valid assessment tasks. **Maria Stec** in chapter four discusses some specifically developed on-line procedures for assessing the academic and professional English of Polish PhD students. The distance learning and assessment involved the use of the platform Moodle. Stec examines the virtual materials using Gibb's reflective cycle (1988) so that the evaluation is systematic and insightful. In chapter five **Stephanie Ashford** presents and evaluates tasks utilizing artificial intelligent (AI) applications (e.g., ChatGPT) for concomitantly developing and assessing professional English language skills of students enrolled in a German university. The chapter highlights the limitations and apprehensions related to generative artificial intelligence (GenAI) applications including the generation of inappropriate language and readily available opportunities to plagiarize. At the same time, Ashford comprehensively discusses and illustrates the productive, effective and perceptive use of these applications in the context of formative holistic assessment of professional writing skills. **Shree Deepa** and **Geetha Durairajan** examine procedures and tasks involving constructive and pragmatic collaborations between English language students and teachers in chapter six. Deepa and Durairajan illustrate the benefits of the focus on learners as multilingual adults in an Indian university context and the productive value of student empowerments during language skills assessment. The chapter demonstrates the application of the teaching and assessment approach for adults termed as 'anthrogogic'.

Sakae Onoda in chapter seven discusses the effectiveness of tasks based on Newton and Nation's (2021) four-strands framework for assessing Japanese university students' interactive oral English skills. The presentation of the three activities derived from the four-strand approach also incorporates the rubrics, responses and the students' viewpoints to exemplify and explicate the effective use of these assessment procedures. In chapter eight **Geetha Durairajan** and **Shree Deepa** expound the concept of language potentiality as reflecting different degrees of sensitivity to the audience. The chapter illustrates viable procedures for assessing the construct of language potentiality in a tertiary level anthrogogic Indian context by presenting specific materials, and the students' responses and feedback.

Asma Aftab in chapter nine elucidates the concept of communicative power which is believed to reflect the attribute of impact as the main aim of communication. The chapter presents the framework, principles, and

rubrics for assessing communicative power. Significantly, it demonstrates the application of the suggested procedure on groups of undergraduate Pakistani students by showcasing specific assessment activities and student responses and feedback. The **concluding chapter** discusses the implications of the presented procedures and materials and recommends proposals to make language assessment more needs-oriented, student-centred, and effectual.

The discussed assessment procedures and applications in most chapters (that is, three to nine) are delimited to tertiary educational levels. However, these expositions could inspire future large-scale research projects analysing the impacts of the identified paradigms and materials in varied educational contexts leading to an improvement in assessment policies/practices. This could in turn result in enhancing and enriching language learning by the adoption of more valid approaches in diverse teaching scenarios.

References

Bachman, Lyle, F., and Damböck, Barbara. 2018. Language Assessment for Classroom Teachers. Oxford: Oxford University Press.

Bennett, Randy Elliot. "Formative Assessment: A Critical Review." Assessment in Education: Principles, Policy and Practice, Vol. 18, No. 1 (2011): 5-25. DOI: 10.1080/0969594X.2010.513678.

Bosompem, Esther, G. "Materials Adaptation in Ghana: Teachers' Attitudes and Practices". In International Perspectives on Materials in ELT, edited by Sue Garton and Kathleen Graves, 104-20. (Basingstoke UK: Palgrave Macmillan, 2014).

Cheng, Liying and Fox, Janna. Assessment in the Language Classroom. 2017. London: Macmillan Palgrave.

Dolin, Jens, Black, Paul, Harlen, Wynne, and Tiberghien, Andrée. "Exploring Relations Between Formative and Summative Assessment". In Transforming Assessment: Through an Interplay Between Practice, Research and Policy, edited by Jens Dolin and Robert Evans, 53-80. (Cham, Switzerland: Springer International Publishing, 2018).

Durdikuliyevna, Abdiyeva, Adolat, Anvarovna, Choriyeva, Durdona, and Zulayho, Abdunazarova. "Formative and Summative Assessment in Teaching Process." European Journal of Research and Reflection in Educational Sciences, Vol. 7, No. 12 (2019): 235-237. https://www.idpublications.org/wp-content/uploads/2020/01/Full-Paper-FORMATIVE-AND-SUMMATIVE-ASSESSMENT-IN-TEACHING-PROCESS.pdf .

Fernández, Claudia, and Berwick, Adon, ed. 2022. Using Language Learning Materials: Theory and Practice. Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Fulcher, Glenn, and Harding, Luke, ed. 2021. The Routledge Handbook of Language Testing (second edition). Abingdon, Oxfordshire, UK: Routledge.

Garton, Sue, and Graves, Kathleen, ed. 2014. International Perspectives on Materials in ELT. Basingstoke UK: Palgrave Macmillan.

Gibbs, Graham. 1988. Learning by Doing: A Guide to Teaching and Learning Methods. Oxford: Oxford Brooks University.

Hargie, Owen. 2019. The Handbook of Communication Skills (fourth edition). Abingdon, Oxon, UK: Routledge.

Harlen, Wynne, and James, Mary. "Assessment and Learning: Differences and Relationships between Formative and Summative Assessment." Assessment in Education: Principles, Policy and Practice, Vol. 4, No. 3 (1997): 365-379. https://doi.org/10.1080/0969594970040304.

Humphries, Simon. "Factors Influencing Japanese Teachers' Adoption of Communication-Oriented Textbooks". In International Perspectives on Materials in ELT, edited by Sue Garton and Kathleen Graves, 253-269. (London: Palgrave Macmillan, 2014).

Jones, Neil, and Saville, Nick, ed. 2016. Learning Oriented Assessment: A Systemic Approach. Cambridge: Cambridge University Press.

Khvatova, Elena, and Krutskikh, Elena."Summary Writing as a Form of Integrated Skills Assessment in Tertiary Setting". In Changing Language Assessment, edited by Sahbi Hidri, 133-158. (London: Palgrave Macmillan, 2020).

Larsen-Freeman, Diane. "The Chaos/Complexity Theory for Second Language Acquisition/Development". In The Encyclopedia of Applied Linguistics, edited by Carol A. Chapelle. (Hoboken, New Jersey, USA: Wiley, 2020). DOI: 10.1002/9781405198431.

McGrath, Ian. 2016. Materials Evaluation and Design for Language Teaching (second edition). Edinburgh: Edinburgh University Press.

Mukundan, Jayakaran. "Are there really Good Reasons as to Why Textbooks should Exist?". In Readings on ELT Materials III, edited by J. Mukundan, 92-100. (Petaling, Jaya: Pearson, Malaysia, 2009).

Newton, Jonathan, M., and Nation, Ian, S., P. 2021. Teaching ESL/EFL Listening and Speaking. Abingdon: Routledge.

Poehner, Matthew, E., and Infante, Paolo. "Dynamic Assessment in the Language Classroom". In Handbook of Second Language Assessment, edited by Dina Tsagari and Jayanti Banerjee, 275-290. (Berlin, Boston: De Gruyter Mouton, 2016).

Popkova, Ekaterina, M. “The Pedagogic Role of Assessment in Improving Learner Outcomes”. In Changing Language Assessment, edited by Sahbi Hidri, 23-52. (London: Palgrave Macmillan, 2020).

Richards, Jack. C. 2001. Curriculum Development in Language Teaching. Cambridge: Cambridge University Press.

Ruler, v., B. “Communication Theory: An Underrated Pillar on which Strategic Communication Rests.” International Journal of Strategic Communication, Vol. 12, No. 4 (2018): 367-381. https://doi.org/10.1080/1553118X.2018.1452240.

Seferaj, Kristjan. “Coping with New Teaching Approaches and Materials: An Eastern European EFL Teacher’s Interpretation of Communicative Teaching Activities”. In International Perspectives on Materials in ELT, edited by Sue Garton and Kathleen Graves, 89-103. (Basingstoke UK: Palgrave Macmillan, 2014).

Tomlinson, Brian. 2003. Developing Materials for Language Teaching. London: Continuum.

Tomlinson, Brian. 2008. English Language Learning Materials: A Critical Review. London: Continuum.

Tomlinson, Brian. “Materials Development for Language Learning and Teaching.” Language Teaching, Vol. 45, No. 2 (2012): 143-179. DOI: https://doi.org/10.1017/S0261444811000528.

Tomlinson, Brian, and Masuhara, Hitomi. 2018. The Complete Guide to the Theory and Practice of Materials Development for Language Learning. Hoboken, New Jersey: Wiley Blackwell.

Vassiliou, Skevi, Papadima-Sophocleous, Salomi, and Giannikas, Christina, Nicole. 2022. Formative Assessment in Second Language Learning: A Systematic Review and An Annotated Bibliography. Voillans, France: Research-publishing.net.

List of Abbreviations

AE	American English
AFL	Assessment for Learning
AI	Artificial Intelligence
ALTE	Association of Language Testers in Europe
AQP	'Anthrogogic Question Paper'
BICS	Basic Interpersonal Communication Skills
BE	British English
CALL	Computer-Assisted Language Learning
CALP	Cognitive Academic Language Proficiency
CCSARP	Cross-Cultural Speech Act Research Project
CEFR	Common European Framework of Reference
CGPA	Cumulative Grade Point Average
CLIL	Content and Language Integrated Learning
CLSS	Comprehensive Learning System
DBR	Design-Based Approach
DCQ	Discourse Completion Questionnaire
DCT	Discourse Completion Task
DIY	Do It Yourself
DA	Dynamic Assessment
DHBW	Baden-Württemberg Cooperative State University (Germany)
EALTA	European Association for Language Testing and Assessment
EAP	English for Academic Purposes
EFL	English as a Foreign Language
EMI	English-Medium Instruction
ESP	English for Specific Purposes
GenAI	Generative Artificial Intelligence
IOR	Inter Organisational Relationship
IT	Information Technology
KEPT	Kanda University of International Studies
LOLA	Learning Oriented Language Assessment
LOA	Learning Oriented Assessment
MATSDA	Materials Development Association
MEXT (Japan)	Ministry of Education, Culture, Science and Technology
mHE	'mainstream Higher Education'

MT	Machine Translation
NMT	Neural Machine Translation
PKG	‘By the teacher for the teacher' (in Bahasa Indonesian)
PQP	‘Pedagogic Question Paper’
SAC	‘Students in an Anthrogogic Classroom’
TOEIC	Test of English for International Communication
WDCT	Written Discourse Completion Task

CHAPTER ONE

ASSESSMENT FOR ACQUISITION

BRIAN TOMLINSON

Introduction

In 2005 I published an article in *ELT Journal* entitled Testing to Learn (Tomlinson, 2005). In this article I advocated:

- "making the provision of opportunities for learning the main objective of language testing" (39)
- recognizing 'the need for tests to be fair, valid, and reliable' but asserting "the priority of 'learning validity' in order to prevent time being wasted … on tests, and the preparation for them" (39)
- insisting that learners should be provided with opportunities for learning in their preparation for tests, whilst taking tests, and whilst receiving constructive feedback during and after taking tests.

Since 2005 I have been giving presentations around the world arguing the case for the three main principles summarised above but focusing on assessment for acquisition rather than just assessment for language learning. In this chapter I will be elaborating and exemplifying these principles and suggesting how adherence to them can facilitate the development of communicative ability.

I will be using the term 'assessment' as an umbrella term to cover classroom tests, internal and external examinations and all other forms of assessment and evaluation. However, I will be focusing primarily on classroom tests and institutional assessments rather than on high stakes examinations, as I believe that what happens in educational institutions could influence improvement in high stakes examinations and in coursebooks and classrooms too. My starting point is that most of the assessments in educational institutions seem to be replicas of those used in high stakes tests (either currently or in the past) and therefore, that such assessment

types influence the content and format of coursebooks. For example, the predominant activity type in many coursebooks is ten closed questions (usually multiple choice or Yes/No questions) as this is still the predominant assessment type in high stakes examinations and such questions can easily be set for classroom tests and can be quickly and reliably marked. This is what I was told by a publisher when I was advocating more open-ended activities in coursebooks and fewer closed questions. It is understandable that publishers cater for what they perceive the market demands, but the market is driven by the needs of administrators and teachers and by those of the learners. Unfortunately, this can deny learners the opportunities for thinking, for communicating, for interacting and for engagement which open-ended questions and activities can afford.

I am still advocating the principles I promoted in my 2005 article because according to my experience most assessment still does not promote learning and can actually prevent it because of the time wasted in preparing for tests and examinations which do not provide learners with opportunities for learning. I am also now advocating that assessment should promote language acquisition and not just language learning. By language acquisition I mean the gradual development of the ability to use the language accurately, fluently, appropriately, confidently and, most important of all, effectively.

To help learners achieve language acquisition we need to make sure that classroom tasks, classroom tests, institutional assessments and, ideally, high stakes examinations meet relevant criteria such as:

- rich, relevant and recycled exposure to the language in communicative use
- affective and cognitive engagement
- opportunities to make discoveries about how the language is used
- opportunities to use the language for purposeful communication.

For discussion of these and other criteria for language acquisition see Tomlinson (2023) and Tomlinson and Masuhara (2018, 2021).

Very few classroom tests, very few institutional examinations, very few (if any) high-stakes examinations and therefore very few coursebook activities currently meet criteria for helping learners to develop communicative ability. Fortunately, though, most experts on testing, assessment and evaluation now stress in their publications the need to focus on the provision of opportunities for learning and a number of powerful movements have developed with a view to radically change how language learners are assessed (e.g., the Comprehensive Learning System [CLSS], Assessment

for Learning [AFL] and Learning Oriented Language Assessment [LOLA]). Unfortunately, though these publications and movements have not yet been very successful in actually bringing about change in assessment practices in schools, colleges and universities around the world. There is hope though, as action is being taken to influence change by, for example, the work of the British Council, of the Common European Framework of Reference (CEFR), of the Association of Language Testers in Europe (ALTE), of the European Association for Language Testing and Assessment (EALTA) and of the Centre for Research in English Language Learning, University of Bedfordshire. For information about current thinking on learner assessment see Gitsaki and Coombe (2016), Jones and Saville (2016), Sahbi (2020) and Britton (2021).

Unfortunately, though, despite the many publications clamouring for change, nothing much seems to have actually changed in most institutions around the world regarding their practises in the assessment of their students. Most publishers are still profiting from the demand for books offering less than useful practice of multiple-choice, Yes/No questions, filling in the blank activities and matching questions. Most learners are wasting precious learning time preparing for and doing tests and examinations which they gain very little from and many learners are suffering loss of self-esteem, confidence and motivation from excessive and inappropriate assessment.

The problems with assessment

Assessment tends to control coursebooks, teaching and learning, to intimidate many learners and to lead to untypical performance and demotivation. It also tends to measure what can be easily measured and to lead to the teaching of what can be easily taught. This restricts opportunities for rich exposure to language in use and opportunities to use the language for communication. Assessment also tends to feature closed questions in order to ensure reliability and ease of marking and this discourages the use of open-ended activities in coursebooks and the classroom (Freeman, 2014; Tomlinson, 2018). It could be that tests, by focusing mainly on constructed examples of correct use of language, typically fail to offer opportunities for authentic learner communication and instead penalise learners for the tiniest of transgressions.

Possible solutions

Ideally, I would like to dispense with institutional examinations and tests altogether. The information gained from them is rarely used to help learners to learn and the gains from then are minimal. Most institutions would still insist on retaining them though as they 'need to maintain standards', whatever that means. In that case I would advise institutions to reduce the amount of internal and external assessment, as the typical means of assessment take up valuable learning time without offering valuable learning opportunities. My advice would also be to assess what the learners have acquired and what learners can do rather than whether or not learners have learned what has been taught. I would also advise teachers to make more use of classroom and homework activities for indications of learner progress instead of one-off tests; to make assessment resemble typical classroom activities rather than the other way round; and, to make sure that test preparation, test-taking and test follow up is designed to prioritise opportunities for acquiring communicative ability.

Opportunities for acquisition

1. Preparing for assessment

My advice when preparing learners for tests and examinations is not to waste time on preparation for discrete item tests. If you focus on promoting language acquisition but provide some practice just before the assessment, the learners will have a better chance of success than if you had spent weeks practising answering typical assessment questions. On the PKG Programme in Indonesia we spent a year with one class in each junior secondary school preparing the students to be able to communicate effectively in English. Then for a short time at the end of the year we gave them a little bit of practice in answering the sort of discrete item questions they would face in the end of the year national examination. Our students dramatically out-performed the students who had spent most of the year answering examination type questions (Tomlinson, 1990). I would also advise making sure the assessment activities are indistinguishable from the learning activities, as we did with primary school students in Vanuatu who enjoyed participating in what are now called task-based activities both in the classroom and in their examinations (Tomlinson, 1981). In addition, I would suggest not mentioning to students that they are taking part in an assessment (as I did with university freshmen in Kobe University, Japan with their consent). My other suggestions would involve making any as-

sessment preparation experiential and purposeful, in the sense that the activities have learning goals and do not just involve mindless practice (for example, getting students to devise reading comprehension questions for other students to answer) and also getting students to do an activity, provide each other with feedback and then do a similar activity as part of the assessment procedure (for example, making a paper aeroplane from oral instructions, getting feedback and then making a paper boat from oral instructions as part of an assessment procedure, as I did in a secondary school in Zambia). Personally and controversially, I would not play the role of invigilator during assessment activities but instead I would be a facilitator available to give help provided it was asked for and given in the target language (as I did during assessment activities at Kobe University). As I see it, providing such help is not being unfair but is part of the assessment procedure of finding out what the learners are able to do with the target language. If learners can seek clarification fluently and successfully, can understand the advice given by the teacher and can apply it effectively then they are deserving of applause rather than punishment.

2 Assessment activities

My advice is not to assess knowledge of isolated skills or isolated language features. This does not give useful information about communicative ability and can lead to excessive and explicit teaching of discrete items and skills at the expense of developing communicative effectiveness. Instead, I would devise integrated activities such as communicative tasks, projects, problem solving etc. which both contribute to language acquisition and provide information about the learners' ability to communicate. Wherever possible I would assess learners in socially cohesive groups working on collaborative tasks (see Sato & Ballinger, 2015 for supportive evidence for such an approach) but sometimes I would assess individual contributions towards achieving the group objectives. If I was required to grade the activities, I would establish context related criteria and make them explicit in the task rubric (for example, 'Rewrite the passage on water conservation so that it is a short and compelling text in a coursebook for ten year old children. Aim to include the main points of the passage in a way which is informative, meaningful and engaging for ten year old children.') I would also use broad bands for marking rather than discrete marks (as we did when marking a national secondary school examination in Zambia when we marked in pairs and allocated essays to one of five piles).

All these suggestions are aimed at making assessment an enjoyable and positive experience for learners, and especially at providing oppor-

tunities to learn from assessments and to develop the ability to communicate effectively.

3 Feedback activities

My advice is to make sure that students always get feedback and get it as soon as possible after the assessment activity. The feedback should always be informative, constructive and non-judgemental and should never just consist of a grade and critical comments. Ideally the feedback should focus on a salient feature of the learner's performance and should aim to help the learner to gain from the experience of being assessed. Wherever possible the feedback should be followed by a similar activity enabling the learners to make use of the feedback they have received (for example, if the assessment activity involved reporting what was said at a meeting and the feedback focused on the use of direct and indirect speech, it would then be useful to ask the students to write a report of what was said at another meeting).

Feedback does not always have to come from a teacher. Learners could be helped to give feedback to themselves and to other learners. This could happen during the assessment activity before a task performance is finalised or after a task activity is completed. It could involve referring to a learner's own notes written during a research activity (for example, about when and how the passive is used), referring to a dictionary or textbook (when, for example, the task involves using genre specific vocabulary), or getting a reaction from another learner or group about the achievement of an objective (For example: 'Would you be likely to help me if you received this request from me?').

Examples of useful assessment activities

An integrated skills activity

- The learners are asked to read a section of a secondary school textbook about water conservation. This activity is not assessed.
- The learners individually write a page of a primary school textbook for ten year olds about water conservation, making use of the information in the text they read in 1 above.

The learners are given grades as follows:

Understanding of the source text (1-5); Suitability of approach and content for ten year old students (1-5); Accuracy of English (1-5); Intelligibility for ten year olds (1-5); Potential for engagement (1-5).

- Learners in groups invent a device for saving water which will be effective, efficient and cheap.

  This activity is not assessed.

- Learners in groups write an e-mail in English to an international company informing them of their invention and requesting a meeting to discuss it.

  The learners as a group are given grades as follows:
  Giving sufficient but not too much information (1-5); Appropriacy of tone (1-5); Accuracy of English (1-5); Potential for engagement (1-5); Likelihood of achieving objective (1-5).

- The learners as a group present their invention orally to a senior representative of an international company in order to persuade them to fund their invention. Each member of the group takes a turn in the presentation. Then each learner takes a turn answering a question about the invention

  The learners as a group are given grades as follows:
  Potential for impact (1-5); Giving sufficient information (1-5); Clarity of the presentation and answers (1-5); Accuracy of English (1-5); Likelihood of achieving objective (1-5).

- The learners as a group devise a television advertisement to sell their device and later perform it to the senior representative.

  The learners as a group are given grades as follows:
  Potential for impact (1-5); Giving sufficient information (1-5); Clarity of the presentation (1-5); Accuracy of English (1-5); Likelihood of achieving objective (1-5).

This activity is an integrated skills task involving reading, listening, writing and speaking in order to achieve a communicative purpose. I used it as an assessment activity with a group of freshmen Engineering students at Kobe University and they enjoyed the different tasks and had great fun when performing the television advertisement with props and musical accompaniments. It was designed so that it could be used either as an as-

sessment or an integrated classroom activity over a number of days and it aimed to stimulate affective and cognitive engagement, to provide exposure to the target language in purposeful use, to provide experience of using the target language for communication, and to contribute to the students' development of self-esteem, positivity and confidence through their contribution to collaborative tasks in socially cohesive groups. And, of course, to reveal information about the students' ability to communicate in order to achieve effect.

At Kobe University the task described above, as an assessment, could be used with subsequent classes of students as an extended classroom activity, just like the tasks used in primary schools in Vanuatu described above and included in Tomlinson (1981). There was no difference between what students did in class and what they did when being assessed.

A communication task for architects

- The teacher places the design for an innovative and impressive new civic building in a box at the front of the classroom. The teacher tells the students that they have a specified time to replicate the design in the box from information about it provided by the one member of the group who is allowed to look at the design.

  The students form groups and decide on a 'runner' who will run to the box, observe the design and then run back to the group to tell them what to draw in order to replicate the design in the box.
  The students can ask their runner questions about the design, can send the runner back up to the box three times to check the design again and the runner can give feedback to the student(s) drawing the design.
  The learners as a group are given grades as follows:
  Similarity to the original in size (1-5); Similarity to the original in colour;(1-5); Similarity to the original in shape (1-5); Similarity to the original in detail (1-5); Amount of the design completed (1-5).

- Each learner is given a copy of the original design and writes a critical review of the design for the potential clients, pointing out its positives and negatives and advising the clients what their response should be.

  Each learner is given grades according to the following criteria:
  Ability to describe the design accurately (1-5); Ability to highlight positive features (1-5); Ability to highlight negative features (1-5);