

# IDENTIFYING THE CRITERIA FOR SELECTING LEARNING OUTCOMES ASSESSMENT TOOLS: THE CASE OF TOEIC AND IELTS

**Dr. Vu Thi Phuong Anh**

*University of Economics and Finance (UEF)*

**Dr. Nguyen Quang Tien**

*University of Economics and Finance (UEF)*

**Ms. Nguyen Thi Tan**

*Education Quality Training & Support Center  
Association of Vietnam Universities and Colleges*

# INTRODUCTION

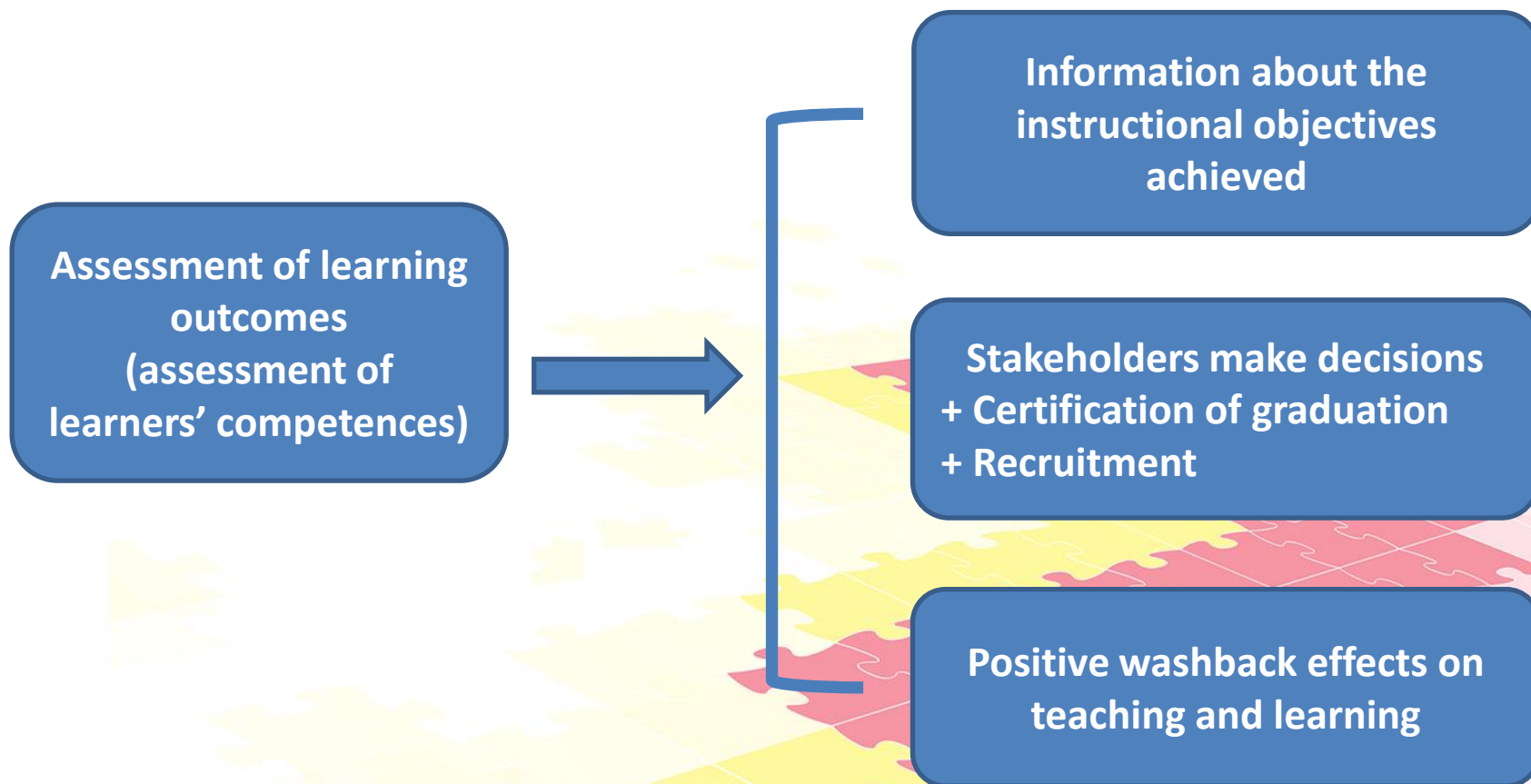
## IDENTIFYING THE CRITERIA FOR SELECTING OUTCOMES ASSESSMENT TOOLS

## APPLICATION OF THE CRITERIA TO ANALYZE TOEIC AND IELTS

## CONCLUSION



# Introduction



# English Language Learning Outcomes Assessment Tools

Language Testing Organizations	TESTS	
Pearson (USA)	PTE	Pearson Test of English
Educational Testing Service (USA)	TOEFL ITP	TOEFL Institutional Testing Program
	TOEFLE iBT	TOEFL Internet-based Testing
	TOEIC	Test of English for International Communication
Cambridge English Language Assessment (England)	IELTS	International English Language Testing System
	Cambridge Suite	Cambridge Suite of Examinations (KET, PET, FCE, CAE, CPE)
	BEC	Business English Certificates
	BULATS	Business Language Testing Service
British Council	APTIS	
National Foreign Languages 2020 Project (Vietnam)	VSTEP	Vietnamese Standardized Test of English Proficiency

Different standardized tests for different purposes, based on different point of view, and intended for different examinees



Different universities in Vietnam have selected different English proficiency tests mostly based on experience and intuition, regardless of whether it is suitable for their students and/ or instructional objectives.

**Teaching and learning  
efforts not paid off**



**Students' learning  
motivation destroyed**

**Achievement of  
instructional  
objectives hindered**

# Purposes of the Presentation

- To display the criteria for selecting learning outcomes assessment tools suitable for students and instructional objectives
- To analyze TOEIC and IELTS
- To make initial inconclusive remarks on the suitability of these two certificates for different types of students
- Not to aim to acknowledge or reject any certificate, but illustrate the importance and necessity of establishing criteria for the selection of certificates

# IDENTIFYING THE CRITERIA FOR SELECTING LEARNING OUTCOMES ASSESSMENT TOOLS

- 
- A decorative background of interlocking puzzle pieces in shades of yellow, pink, and red, with a small yellow star on one of the pink pieces.
1. The framework of “usefulness” of foreign language proficiency tests
  2. Suggesting a suitable analysis framework of language ability test of three components

# The framework of “usefulness” of foreign language proficiency tests

Bachman and Palmer’s (1996)  
framework of “usefulness”

## Concept of communicative language ability (CLA)

(focus on the role and impact of context on language use  
on the manifestation of users’ language ability)

“Usefulness” = a combination of 6 components  
(reliability + construct validity + authenticity + interactiveness + impact +  
practicality)



- Consistency/ stability of scores
- How to have reliability?
  - + Increasing the number of questions
  - + Increasing the degree of differentiation among questions



In order to avoid the situation of an increase in reliability and a decrease in validity, Bachman and Palmer (1996) define reliability in a language ability test as **consistency of language tasks** in different test versions. **In combination of authenticity** (the third component in the framework), consistency of language tasks both creates stability of scores and ensures test validity (Bachman & Palmer, 1996: 19-21).

- Appropriateness of interpretations based on scores
- **Accurate indication of the test taker's language communication ability**



The meaning and appropriateness of these interpretations **need to be proved by arguments and evidence**, not by assertion only (Bachman & Palmer, 1996:21-22).



- The correspondence between language tasks on a test and target language use (TLU)
- An indicator from which users' language ability is inferred in TLU



# Interactiveness

The extent to which test takers' various competences are employed to do language tasks on a test

(linguistic competence + other competences)

A well-designed language test is the one which allows test takers **to employ their existing knowledge to perform language tasks on a test** (Bachman & Palmer, 1996).



- At the individual level
  - \* Impact (positive or negative) on **the test takers** (motivation, learning habits, etc.)
- At the system level
  - \* Impact (positive or negative) on the teaching methods and outcomes of **the whole education system**



Ex: “Dump and deaf methods”

The word "practicality" is written in a white, lowercase, sans-serif font on a teal rectangular background with a black border.

## practicality

The correspondence between individual test takers' capability and necessary requirements to do the test successfully

Unlike the five previous components, practicality does not have any direct connection with the test quality, but with **practical conditions worthy of attention for taking a test** (Bachman & Palmer, 1996: 35-37).

# Suggesting a suitable analysis framework of language ability test of three components

Three suggested components as the basis for selecting English proficiency certificate tests

“Reliability” and “construct validity”  
in Bachman and Palmer’s (1996) framework

Communication features of language input and tasks  
(“authenticity” and “interactiveness” in the framework of “usefulness”)

The impact of a good test on teaching and learning  
(“impact” in the framework of “usefulness”)

# **APPLICATION OF THE CRITERIA JUST IDENTIFIED TO ANALYZE THE TWO INTERNATIONAL CERTIFICATES TOEIC AND IELTS**

1. The validity of the analysis framework
2. Communicative components of language input and tasks



## The validity of the analysis framework

	TOEIC	IELTS
Origin	Developed in 1979	Developed in 1989
Testing organization	ETS	Cambridge
Theory	<b>Theory of language</b>	<b>Communicative language competence</b> Language as a communication tool > The role of context of use > Interactiveness in language use
Assessment	Norm-referenced	Criterion-referenced
	No direct provision of learners' language ability ETS's guidelines not used to interpret TOEIC scores in reference to the CEFR	
Validity and reliability	MC Questions --> High validity	High validity and much reliability
Common practice	"Social moderation" Two language skills test (L & R) still used > This test should not be equated to the CEF Levels (in which 4 language skills are assessed) (Construct under representation)	4 language skills tested (objectively & subjectively)

## TOEIC (Listening – Reading)

No. of Questions/ Parts	TASKS	TOPICS
200 MC Questions		
* 100: Listening		Daily transactions
> 40: Individual statements	Selecting the statement best describing the picture (10)	Simple: booking, receiving guests, ...
	Selecting the best response to the question (30)	(Suitable for
> 30: Short conversations	Selecting the best response to each question	* officers
> 30: Short talks		* administrators
* 100: Reading		* sales representatives
> Sentence level	Filling the blank with a word or a phrase	* ....
> Paragraph level	Selecting the best answer to the questions	(Not for professional fields)
> Short passage level		

# IELTS

No. of Questions/ Parts	TASKS	TOPICS	Skills tested
4 parts for 4 language skills		Daily life	
Various types of Qs		Academic settings	
* MC Qs			
* Essay Writing			
* Interview			
<b>Listening Test (40 Qs)</b>			
Section 1: A conversation of 350-400 words	Filling the blanks	Social context	Listening for factual details
Section 2: A monologue of 350-400 words	MC Qs, Filling the blanks	Daily topics	Listening for gist
Section 3: A conversation of 400-600 words	Matching	Educational or training context	Recognizing attitudes and purposes
Section 4: A lecture of 650-800 words	Labeling,	An academic subject	Following the development of arguments
<b>Reading Test (40 Qs)</b>		> <b>Authentic:</b> textbooks, journals . . .	Reading for gist
3 texts of 700 - 1000 words each	MC Questions	> for non-specialist audience	Reading for specific points
	Matching	> Various styles:	Making inferences
	Labeling	* Scientific	Understanding text purposes
	Completion: Form/ note/ table/ flow-chart/ summary completion	* Descriptive	etc.
	Short answers	* Discursive	
	(> More or less simulate TLU)	* Argumentative	
<b>Writing Test</b>			
First Task	Describing a chart/ a table		
Second Task	Writing an essay		Presenting opinions and/ or arguments
<b>Speaking Test</b>			
Part 1: Introduction & Interview	General questions	Home, family, work, interests	
Part 2: A talk on a particular topic	Talking about a particular topic	A given topic	
	Answering some questions		
Part 3: In-dept discussion	follow-up questions related to P. 2		

TESTS	VOCABULARY	6-level framework
TOEIC (L & R)	5,000 general	Levels 3 and 4 of low end
	600 business	(NOT Levels 5 and 6
		Bc lang tasks without high authenticity
		(Unlike target language use)
IELTS	7,000 general	Levels 4 and 5
	> 600 technical	(can be inferred approximately
		one level lower or higher,
		i.e., Level 3 or Level 6)

- In terms of communication of tasks and language input, **IELTS is a better language ability test than TOEIC**, since the IELTS has a much higher degree of authenticity and interactivensess.
- the IELTS advises that **the band score below 4.0 should not be considered**, since its interpretation may not be accurate.



## Remarks on TOEIC

Topics/ Contexts	On a daily basis
	Daily transactions
6-Level equivalence	3 or 4, or lower
Interpretations of language ability	Higher levels need to be closely considered
Language input & tasks	Suitable for general positions
	(employees, office managers, assistants. . .)
Myth	Suitable for all disciplines & positions
Recommendations	Drawing on other sources of information
	when making a decision on using the test or not

- As the TOEIC test had been developed long before the CEFR was, **different levels of language ability of the TOEIC test are not linked to a certain corresponding series of English textbooks** to support English teaching and learning. This “**something learned, something else assessed**” may limit the effectiveness of language teaching.
- Due to the only use of multiple choice questions and limited language input in the TOEIC test, it is **better to be cautious about possible negative washback effects of the test** on teachers’ teaching methods and test takers’ English learning.

## Remarks on IELTS

Topics/ Contexts	General & Academic contexts
	(suitable for many schools)
6-Level equivalence	4 or 5, or higher
Interpretations of language ability	Assuring the lang ability of those who hold IELTS Certificate
Language input & tasks	In relation to assessment of communicative language competence
	Suitable in academic contexts
	* Reading professional documents
	* Writing international conference reports
	* Dealing with issues related to diplomacy
	medicine, technology, business, commerce
Myth	There is only one module - Academic.
	(Bc the scores of IELTS allow holders to enroll in degree programs in many English-speaking countries.)

- In reality, university graduates in Vietnam have demands for consulting foreign specialist documents in relation to their studies, and this is the orientation of the IELTS test.
- However, **programs of study taught in English in many universities in Vietnam do not usually impose this requirement.**
- Thus, in order to obtain an expected score of, for example, 4.5 or 5.0 (a score which corresponds to CEF Level B1 or Level 3 in Vietnam's English 6-level language ability framework), it is essential that **test takers need to have more academic reading and writing practice.**
- It is **unfair to require every test taker to take the IELTS test to complete their programs of study** when the curriculum does not ask them to practice academic Reading and Writing skills.



In general, **IELTS is rather difficult to most graduates in Vietnam**, since the language input and tasks on the test are set at Levels 4 and 5, or higher than the scores test takers need to obtain.

However, due to the fact that the format of the test closely follows the latest theory of assessment of communicative language competence, **IELTS is evaluated as one with positive backwash effects on previous learning and teaching**. This has been proved many times in published research.

# Conclusion

- A certain certificate can be appropriate for a certain type of test takers, but not for another.
- It is **unscientific to select and impose A CERTAIN CERCIFICATE ONLY** to assess students' learning outcomes without analyzing its suitability for test takers and instructional objectives. It is also unfair to test takers.
- Doing so can entail **possible negative washback effects** on the whole system, decreasing the effectiveness of prior teaching and learning.

# Conclusion

- This situation can explain the fact that though it has been required of graduates in many schools to submit international language proficiency certificates, **their language ability is still underestimated by recruiters.**
- There is **need for empirical research on effects of each type of certificate on language learners in Vietnam** – a gap needed to be bridged so as to navigate the National Foreign Language 2020 Project.



# Conclusion

- The remarks presented earlier aim not to affirm the appropriateness or inappropriateness of TOEIC and IELTS in any certain situation, but **to suggest the process of and criteria for selecting a language proficiency test as an outcomes assessment tool.**
- When there are more and more different language proficiency tests in Vietnam, it is **necessary to select the right one or modify/ develop one to completely suit language learners.**
- Doing so will contribute **positive effects on the system and enhancement of quality teaching and bring us close to an ambitious objective of the 2020 Project:** By 2020, foreign language ability will be one of Vietnamese' strengths.

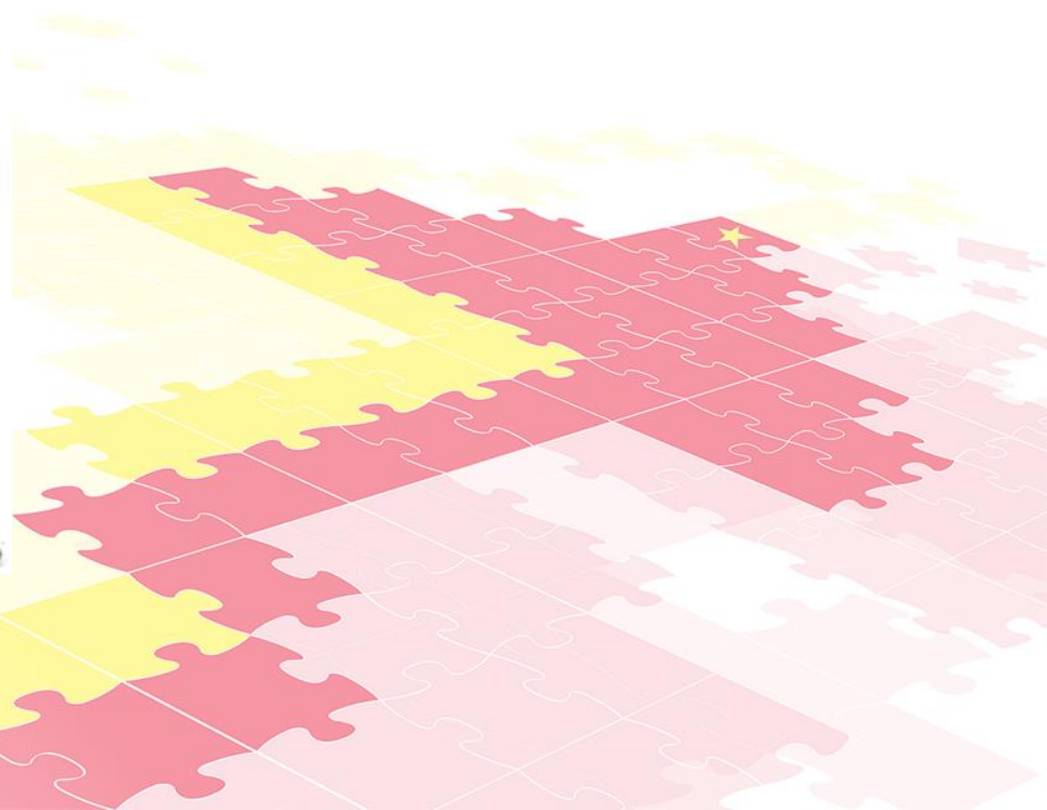


## REFERENCES

- Bachman, L.F. (1990) *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L.F. and A.S. Palmer (1996) *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Feuer, M.J et al (Eds) (1999) *Uncommon measures: Equivalence and linkage among education tests*. Washington, D.C: National Academy Press. Accessed online on Dec 28, 2014 at <http://files.eric.ed.gov/fulltext/ED440984.pdf>
- Chujo, K. and Genung, M (2005) “TOEIC tests: *The qualification of vocabulary-usage levels and the extraction of characteristically used words*”. TOEIC Research Report Number 3. Accessed online on November 19, 2015 at [http://5d.biglobe.ne.jp/~chujo/data/KiyomiChujo\\_E-TOEIC-Report.pdf](http://5d.biglobe.ne.jp/~chujo/data/KiyomiChujo_E-TOEIC-Report.pdf)
- Johnson, N (2007) “*A consideration of Assessment validity in relation to classroom practice*”. Cambridge Assessment. Accessed online on Nov 28, 2015 at [http://iaea.infor/documents/paper\\_1162b25504.pdf](http://iaea.infor/documents/paper_1162b25504.pdf)
- Lupp, L. et al (no year). *Exploring consequential validity of alternate assessments based on alternative achievement standards. Report to the States of the results of an online teacher survey in a three state consortium*. Assessed online on Nov 28, 2015 at <http://www.education.pa.gov/Documents/K-2/Special1%20Education/Assessment/Consequential1%20Validity%20-%20Report%20to%20the%20State.pdf>
- Taylor, L. and S. Chan (2015) “IELTS equivalence research project – Appendices” CRELLA. Accessed online on Oct 28, 2015 at [http://www.gmc-uk.org/IELTS\\_Equivalence\\_Research\\_Project\\_Appendices.pdf\\_63059074.pdf](http://www.gmc-uk.org/IELTS_Equivalence_Research_Project_Appendices.pdf_63059074.pdf)



#00537212



## Dr. Vu Thi Phuong Anh

*University of Economics and Finance, UEF*

Email: [vtpanh@gmail.com](mailto:vtpanh@gmail.com) or [anhvtp@uef.edu.vn](mailto:anhvtp@uef.edu.vn)



## Dr. Nguyen Quang Tien

*University of Economics and Finance, UEF*

Email: [nguyenquangtien@gmail.com](mailto:nguyenquangtien@gmail.com) or [tiennq@uef.edu.vn](mailto:tiennq@uef.edu.vn)

## Ms. Nguyen Thi Tan

*Education Quality Training & Support Center*

*Association of Vietnam Universities and Colleges*

Email: [tannguyenht@yahoo.com](mailto:tannguyenht@yahoo.com)