City University of Hong Kong
Course Syllabus

offered by Department of English
with effect from Semester A 2024/25

Part I Course Overview

Course Title:	Testing and Evaluation in Language Studies
Course Code:	EN5466
Course Duration:	1 semester
Credit Units:	3
Level:	P5
Medium of Instruction:	English
Medium of Assessment:	English
Prerequisites: *(Course Code and Title)*	Nil
Precursors: *(Course Code and Title)*	Nil
Equivalent Courses: *(Course Code and Title)*	Nil
Exclusive Courses: *(Course Code and Title)*	Nil

Part II Course Details

1. Abstract

This course aims at helping students to identify key issues in language testing and evaluation. Students will learn to describe basic test statistics and analyse the characteristics of good language assessments. They will apply this knowledge to the construction, revision, and administration of valid and reliable language assessments.

2. Course Intended Learning Outcomes (CILOs)

(CILOs state what the student is expected to be able to do at the end of the course according to a given standard of performance.)

No.	CILOs	Weighting (if applicable)	Discovery-enriched curriculum related learning outcomes (please tick where appropriate)		
			A1	*A2*	*A3*
1.	Designing standard procedures in assessment, and applying this knowledge of standard procedures to the construction, revision and administration of assessments	20%	✓	✓	✓
2.	Critiquing and selecting appropriate assessment tools to assess different language skills	30%	✓	✓	✓
3.	Analysing the validity, reliability and washback effect of a language test/method of assessment	30%	✓	✓	✓
4.	Analysing and interpreting test scores	20%	✓	✓	✓
		100%			

A1: *Attitude*
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: *Ability*
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to real-life problems.

A3: *Accomplishments*
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

3. Learning and Teaching Activities (LTAs)

(LTAs designed to facilitate students' achievement of the CILOs.)

LTA	Brief Description	CILO No.						Hours/week (if applicable)
		1	2	3	4			
1.	**Reading the course book and recommended readings** Students will complete the reading assignments related to language testing.	✓	✓	✓	✓			5 hours / week over 13 weeks
2.	**Interactive lecture and student in-class activities** Students will learn key concepts related to language testing and evaluation during the lectures. Students will also analyse different assessments by applying the key concepts in language assessment and relate the concepts to actual test/assessment constructions and administrations that they are familiar with.	✓	✓	✓	✓			2.5 hours / week over 11 weeks
3.	**Online discussion** Students will conduct online discussion related to language testing. The discussion allows students to share their own language testing and assessment practice and viewpoints on the issues with their classmates. They can practise applying, analysing, interpreting and evaluating concepts covered in the course.	✓	✓	✓	✓			0.5 hr/week over 11 weeks
4.	**Conferencing** Students will participate in individual/small group consultations which they will receive feedback on their application of the concepts to students' own test development.	✓	✓	✓	✓			45 mins in 2 meetings

4. Assessment Tasks/Activities (ATs)

(ATs are designed to assess how well the students achieve the CILOs.)

Assessment Tasks/Activities	CILO No.						Weighting	Remarks
	1	2	3	4				
Continuous Assessment: 100%								
Online discussion posting Questions/issues related to language testing and evaluation will be posted on the online discussion board to help students identify, describe and apply the concepts covered in the course. Students are encouraged to do continuous revision throughout the course and to read and comment on their classmates' questions and viewpoints.	✓	✓	✓	✓			10% of course total (Individual work; to be assessed individually)	
Individual Assignment One individual assignment will be given to students to help them apply concepts covered in the course to the development and revision of an assessment. To complete the assignment, students need to follow two key steps that a classroom test/assessment developer needs to go through before re-administering a test/method of assessment to another group of students. In following both steps, students need to apply concepts learnt in the course to generate their own assessment. (Cont. on next page)	✓	✓	✓	✓			90% of course total (Individual work; to be assessed individually)	

Individual Assignment (cont.) Step 1 – Critique an existing test / assessment & suggest how it can be improved. Step 2 – Revise/Construct the test / assessment & pilot it. At the end of the two steps, students need to submit a 15-page report describing, explaining and evaluating the processes taken to revise and develop the test/assessment. In Step 1, students need to apply concepts learnt in the first few lessons of the course to evaluate the test/assessment they have chosen. After completing Step 1, students need to submit their comments briefly in bullet points. In Step 2, students need to implement the revision plan they have suggested in Step 1, interpret and analyse the pilot findings and evaluate the final test/assessment using all concepts learnt in the course. After completing Step 2, students need to submit the final version of the test/assessment with annotations of piloting findings. After each step, students will receive comments on their work in a face-to-face consultation. Students will also receive suggestions on their individual test development practice related to the assignment.								
Examination: <u>0%</u>								
							100%	

Marking breakdown and CILOs		Weighting
Online discussion posting • Applying concepts covered in the course accurately throughout the course (CILOs 1-4)		10%
Individual assignment		90%
• Constructing and administering assessments following standard development procedures (CILO 1)	18%	
• Choosing appropriate assessment tools to assess different language skills (CILO 2)	15%	
• Evaluating the validity of a test (CILO 3)	10%	
• Evaluating the reliability of a test (CILO 3)	10%	
• Evaluating the test scores of a test (CILO 4)	18%	
• Evaluating the washback effect of a test (CILO 3)	7%	
• Revising a test effectively (CILO 2)	12%	

5. Assessment Rubrics

(Grading of student achievements is based on student performance in assessment tasks/activities with the following rubrics.)

Applicable to students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter

Assessment Task 1: Online Discussion Posting **Criteria: Content, adequacy, purpose, language**

Excellent (A+, A, A-)	**Good (B+, B, B-)**	**Fair (C+, C, C-)**	**Marginal (D)**	**Failure (F)**
♦ Shows full understanding of main concepts and their application**;** ♦ All relevant information is included in discussion and analysis; ♦ The topic is comprehensively analysed and explained; ♦ The purpose of the analysis of concepts is completely achieved; ♦ Style and tone are highly appropriate.	♦ The main concepts are competently discussed and applied**;** ♦ The information included in discussion and analysis of concepts is sufficient; ♦ The topic is sufficiently analysed and explained; ♦ The purpose of the conceptual analysis is achieved; ♦ Style and tone are appropriate.	♦ The concepts selected for analysis are sufficient and partially applied**;** ♦ Only partial information is included in discussion and analysis of concepts; ♦ Only partial analysis is provided; ♦ The purpose of the conceptual analysis is partially achieved; ♦ Style and tone are somewhat appropriate.	♦ The concepts selected for analysis are sketchy and inadequate**;** ♦ Incomplete information is included in discussion and analysis of concepts; ♦ The analysis is not informative or comprehensive**;** ♦ The purpose of the conceptual analysis is not adequately achieved; ♦ Style and tone are inappropriate.	♦ The concepts selected for analysis are highly inadequate**;** ♦ Very limited or inaccurate information is incorporated in conceptual analysis; ♦ The analysis is not at all comprehensible**;** ♦ The purpose of the conceptual analysis is not achieved in any way; ♦ Style and tone are completely inappropriate.

Assessment Task 2: Individual Assignment **Criteria: Content, presentation, purpose, language**

Excellent (A+, A, A-)	**Good (B+, B, B-)**	**Fair (C+, C, C-)**	**Marginal (D)**	**Failure (F)**
♦ The topic is extremely well-presented and analysed**;** ♦ All relevant information is excellently covered; ♦ The purpose of analysing and presenting the material is completely achieved; ♦ Style and tone are highly appropriate.	♦ The topic is competently presented and very well analysed**;** ♦ The information is sufficiently covered; ♦ The purpose of analysing and presenting the material is achieved; ♦ Style and tone are appropriate.	♦ The topic is adequately presented and is analysed reasonably well; ♦ Only part of the information is covered; ♦ The purpose of analysing and presenting the material is partially achieved; ♦ Style and tone are somewhat appropriate.	♦ The topic is sketchily presented and analysed inadequately**;** ♦ Only limited information is included; ♦ The purpose of analysing and presenting the material is not fully achieved; ♦ Style and tone are inappropriate.	♦ The topic is not adequately presented and is not analysed**;** ♦ No relevant or accurate information is included; ♦ The purpose of analysing and presenting the material is not achieved in any way; ♦ Style and tone are completely inappropriate.

Applicable to students admitted from Semester A 2022/23 to Summer Term 2024

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B)	Marginal (B-, C+, C)	Failure (F)
1. Online Discussion Posting	Content	• Shows full understanding of the main concepts and their application. • All relevant information is included in discussion and analysis.	• Shows understanding of the main concepts and their application. • The information included in discussion and analysis of concepts is sufficient.	• Shows some understanding of the main concepts and their application. • Only partial information is included in discussion and analysis of concepts.	• Shows little or no understanding of the main concepts and their application. • Incomplete information is included in discussion and analysis of concepts.
	Purpose	• The purpose of the analysis of concepts is completely achieved.	• The purpose of the conceptual analysis is achieved.	• The purpose of the conceptual analysis is partially achieved.	• The purpose of the conceptual analysis is not adequately achieved.
…	Language	• Style and tone are highly appropriate.	• Style and tone are appropriate.	• Style and tone are somewhat appropriate.	• Style and tone are inappropriate.
2. Individual assignment	Content & presentation	♦ The content is extremely well-presented and analysed. • All relevant information is excellently covered.	• The content is well presented and analysed. • The relevant information is sufficiently covered.	• The content is adequately presented and is analysed reasonably well. • Only part of the relevant information is covered.	• The content is not adequately presented and is not analysed. • No relevant or accurate information is included.
	Purpose	• The purpose of the assignment is completely achieved.	• The purpose of the assignment is achieved.	• The purpose of the assignment is partially achieved.	• The purpose of the assignment is not achieved in any way.
	Language	• Language use is accurate. • Style and tone are highly appropriate.	• Language use is mostly accurate. • Style and tone are appropriate.	• Language use is sometimes inaccurate causing minor confusion. • Style and tone are somewhat appropriate.	• Language use is frequently inaccurate causing major confusion. • Style and tone are inappropriate.

Part III Other Information (more details can be provided separately in the teaching plan)

1. Keyword Syllabus

(An indication of the key topics of the course.)

Basic Concepts

- Constructive alignment
- Proficiency tests and achievement tests
- Subjective testing and objective testing
- Norm-referenced testing and criterion-referenced testing
- Assessment of learning and assessment for learning
- Washback
- Validity
- Reliability
- Fair testing practices
- Code of ethics for language testers

Testing tools and testing procedures

- Assessing different learning skills
- Evaluating with tests and without tests
- Performance assessments
- Diagnostic assessments
- Standardised tests
- Writing multiple choice items
- Steps in test development
- Test specifications
- Trialling and piloting
- Moderating items
- Setting assessment criteria
- Training markers

Test evaluation and interpretation of test scores

- Item analysis and content analysis
- Item facility, analysis of distractors, item discrimination
- Dichotomous scale
- Partial credit scoring
- Item-test correlation
- Classical testing theory
- Item response theory
- Inter-rater reliability and intra-rater reliability

2. Reading List

2.1 Compulsory Readings

(Compulsory readings can include books, book chapters, or journal/magazine articles. There are also collections of e-books, e-journals available from the CityU Library.)

1.	Hughes, A. (2003). *Testing for language teachers.* Cambridge, UK: Cambridge University Press.

2.2 Additional Readings

(Additional references for students to learn to expand their knowledge about the subject.)

1.	Alderson, J.C. (2000). *Assessing reading.* Cambridge, UK: Cambridge University Press.
2.	Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment.* Continuum.
3.	Alderson, J. C., Brunfaut, T., & Harding, L. (2014). Towards a theory of diagnosis in second and foreign language assessment: insights from professional practice across diverse fields. *Applied Linguistics 2014*: 1–26.
4.	Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14*(2), 115-129.
5.	Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation.* Cambridge, UK: Cambridge University Press.
6.	Bachman, L.F. (2004). *Statistical analyses for language assessment.* Cambridge, UK: Cambridge University Press.
7.	Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly, 2*(1), 1-34.
8.	Bachman, L. & Damböck, B. (2017). *Language assessment for classroom teachers*. Oxford, UK: Oxford University Press.
9.	Bachman, L.F. & Palmer, A.S. (1996). *Language testing in practice: Designing and developing useful language use.* Oxford, UK: Oxford University Press.
10.	Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford: Oxford University Press.
11.	Berry, R. (2011). Assessment trends in Hong Kong: Seeking to establish formative assessment in an examination culture. *Assessment in Education: Principles, Policy & Practice, 18,* 199–211.
12.	Buck, G. (2001). *Assessing listening.* Cambridge, UK: Cambridge University Press.
13.	Carr. N.T. (2011). *Designing and analyzing language tests*. Oxford, UK: Oxford University Press.
14.	Carless, D., Joughin, G., & Liu, N.F. (2007). *How assessment supports learning: Learning-oriented assessment in action*. Hong Kong: Hong Kong University Press.
15.	Cheng, L. (2008). The key to success: English language testing in China. *Language Testing, 25*(1), 15-37.
16.	Condon, W. (2013). Large-scale assessment, locally developed measures, and automated scoring of essays: Fishing for red herrings? *Assessing Writing 18,* 101-108.
17.	Davidson, F., & Lynch, B. K. (2002). *Testcraft: A teacher's guide to writing and using language test specifications.* Yale: Yale University Press.
18.	Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T. and McNamara, T. (1999). *Dictionary of language testing.* Cambridge, UK: Cambridge University Press.
19.	Douglas, D. (2010). *Understanding language testing.* Hodder Education.
20.	Fulcher, G. & Davidson, F. (2007). *Language testing and assessment- An advanced resource book.* New York: Routledge.
21.	Fulcher, G. & Harding L. (2022). *The Routledge Handbook of Language Testing (2nd Ed.).* London & New York: Routledge.
22.	Green, A. (2021). *Exploring language assessment and testing (2nd Ed.).* Routledge.
23.	Huhta, A. (2008). Diagnostic and formative assessment. In B. Spolsky & F. Hult (Eds), *Handbook of Educational Linguistics* (pp. 469–82). Blackwell.
24.	Kane, M. (2012). Validating score interpretations and uses. Messick Lecture, Language Testing Research Colloqium, Cambridge April 2010. *Language Testing, 29*(1), 3-17.
25.	Lo, Y. Y., & Fung, D. (2018). Assessments in CLIL: the interplay between cognitive and linguistic demands and their progression in secondary education. *International Journal of Bilingual Education and Bilingualism*, 1-19. Retrieved from https://doi.org/10.1080/13670050.2018.1436519.
26.	Luoma, S. (2004). *Assessing speaking.* Cambridge, UK: Cambridge University Press.

27.	McNamara, T.F. (1996). *Measuring second language performance.* New York: Longman.
28.	McNamara, T.F. (2000). *Language testing.* Oxford, UK: OUP.
29.	McNamara, T.F., Knoch, U. & Fan, J. (2019). *Fairness, justice, and language assessment: The role of measurement.* Oxford: Oxford University Press.
30.	Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*(3), 241-256.
31.	Moskal, B.M. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. *Practical Assessment, Research & Evaluation, 8*(14). Retrieved from http://PAREonline.net/getvn.asp?v=8&n=14.
32.	Purpura, J.E. (2004). *Assessing grammar.* Cambridge, UK: Cambridge University Press.
33.	Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing, 18*(4), 429-462.
34.	Rea-Dickins, P., & Gardner, S. (2000) Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing, 17*(2), 215–43.
35.	Read, J. (2000). *Assessing vocabulary.* Cambridge, UK: Cambridge University Press.
36.	Read, J. and Chapelle, C.A. (2001). A framework for second language vocabulary assessment. *Language Testing, 18*(1), 1-32.
37.	Wiegle, S.A. (2002). *Assessing writing.* Cambridge, UK: Cambridge University Press.
38.	Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Palgrave Macmillan.
39.	Zhang, Y., & Elder, C. (2009). Test review: Measuring the speaking proficiency of advanced EFL learners in China: The CET–SET solution. *Language Assessment Quarterly, 6*(4), 298-314.
40.	Zou, S. & Xu, Q. (2017) A washback study of the Test for English Majors for Grade Eight (TEM8) in China. *Language Assessment Quarterly, 14*(2), 140-159.