



THE ASSOCIATION OF TEACHING ENGLISH
AS A FOREIGN LANGUAGE
IN INDONESIA

The nd 62 TEFLIN

International
Conference
2015

Denpasar, 14th - 16th September 2015

PROCEEDINGS

*Teaching and Assessing L2 Learners
in the 21st Century*



ENGLISH DEPARTMENT
FACULTY OF LETTERS AND CULTURE
IN COLLABORATION WITH
POST GRADUATE STUDY PROGRAM, UDAYANA UNIVERSITY

BOOK 3

PROCEEDINGS

The 62nd TEFLIN

International Conference 2015 Teaching and Assessing L2 Learners in the 21st Century

Internal Reviewers

- | | |
|--|--|
| 1. Ni Luh Ketut Mas Indrawati
(Udayana University) | 16. I Komang Sumaryana Putra
(Udayana University) |
| 2. I Gusti Ayu Gde Sosiowati
(Udayana University) | 17. I Wayan Mulyawan
(Udayana University) |
| 3. Ni Luh Nyoman Seri Malini
(Udayana University) | 18. Ida Ayu Made Puspani
(Udayana University) |
| 4. Ni Wayan Sukarini
(Udayana University) | 19. Putu Weddha Savitri
(Udayana University) |
| 5. Ni Made Ayu Widiastuti
(Udayana University) | 20. Made Sena Darmasetiyawan
(Udayana University) |
| 6. I Made Rajeg
(Udayana University) | 21. I Gusti Ngurah Parthama
(Udayana University) |
| 7. A A Sagung Shanti Sari Dewi
(Udayana University) | 22. I Nyoman Udayana
(Udayana University) |
| 8. Putu Ayu Asty Senja Pratiwi
(Udayana University) | 23. I Ketut Wandia
(Udayana University) |
| 9. I Gusti Agung Istri Aryani
(Udayana University) | 24. I Gede Budiasa
(Udayana University) |
| 10. Ni Ketut Sri Rahayuni
(Udayana University) | 25. I Made Netra
(Udayana University) |
| 11. Yana Qomariana
(Udayana University) | 26. I Gede Putu Sudana
(Udayana University) |
| 12. Ni Ketut Alit Ida Setianingsih
(Udayana University) | 27. I Ketut Tika
(Udayana University) |
| 13. Sang Ayu Isnu Maharani
(Udayana University) | 28. I Nyoman Aryawibawa
(Udayana University) |
| 14. Ni Putu Lirishati Soethama
(Udayana University) | 29. I Nengah Sudipa
(Udayana University) |
| 15. I Nyoman Tri Ediwan
(Udayana University) | 30. Ni Luh Putu Laksminy
(Udayana University) |

External Reviewers

- | |
|---|
| 1. Benedictus B. Dwijatmoko
(Universitas Sanata Dharma) |
| 2. Chuzaimah Dahlan Diem
(Universitas Sriwijaya) |
| 3. Diemroh Ihsan
(Universitas Sriwijaya) |
| 4. Gusti Astika
(Universitas Kristen Satya Wacana) |
| 5. Emi Emilia
(Universitas Pendidikan Indonesia) |
| 6. Cayandrawati Setiono
(Universitas Lambung Mangkurat) |
| 7. Oikurema Purwati
(Universitas Negeri Surabaya) |
| 8. Setyadi Setyapranata
(Universitas Negeri Malang) |
| 9. Yazid Basthomi
(Universitas Negeri Malang) |
| 10. Lies Amien Lestari
(Universitas Negeri Surabaya) |
| 11. Fuad Abdul Hamied
(Universitas Pendidikan Indonesia) |
| 12. Handoyo Puji Widodo
(Politeknik Negeri Jember) |
| 13. Nur Arifah Drajadi
(SMA Labschool Jakarta) |
| 14. I Made Hery Santoso
(Universitas Pendidikan Ganesha) |

Setting and Typeset

Gede Primahadi Wijaya
Made Artadi Gunawan
Artika Putri
Gusti Agung Ngurah Dwi Suryawan
Moh. Noval Ashari
I Wayan Gede Agus Wirawan
Ni Wayan Manik Septianiri Putri
I Made Yoga Dwi Angga
Ni Luh Putu Sisiana Dewi

Cover

I Gede Juniasta Datah

Publisher:

Udayana University Press
in collaboration with
English Department Faculty of Letters and Culture
and Post Graduate Study Program, Udayana University
Email : sasingunud@gmail.com

ISBN: 978-602-2940-66-1

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means: electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without prior written permission from the writers.

THE ROLE OF CAMPUS CLIMATE, FACULTY AND PEERON ENGLISH MAJOR STUDENTS' SUCCESS

Bayu H. Wicaksono ~ 193

THE IMPROVEMENT OF ENGLISH SPEAKING SKILL IN ELEMENTARY SCHOOL STUDENTS THROUGH THE MONTESSORI METHOD

Herlina ~ 201

FIRST LANGUAGE APPROACH IN EFL LEARNING: HOW DO STUDENTS AND TEACHERS SEE IT?

Magdalena Kartikasari Tandy Rerung ~ 207

APPLYING THE ACTIVITIES OF CASE BUILDING TECHNIQUE FOR THE STUDENTS' WRITING IMPROVEMENT

Nur Fitri, S. & Fatimah Hidayahni Amin ~ 215

IMPLEMENTING TEAM TEACHING IN TEACHING ENGLISH FOR YOUNG LEARNER

Fida Anisah, Fathin Anjani Hilman, & Lestiyani Sunarto ~ 223

IELTS SPEAKING TEST: REVIEW, LIMITATION, STANDARDISATION AND REVISION TO ENSURE VALIDITY AND RELIABILITY

Andy ~ 231

"DELIVERING ENGLISH E-LEARNING: A STRATEGY FOR IMPLEMENTATION" (A REVIEW OF BEST PRACTICES IN JAPAN)

Ben Porter & Irene Irawaty ~ 239

USING BLENDED ON LEARNING, WEBLOG AND E-LEARNING TO IMPROVE STUDENTS' WRITING ABILITY OF STUDENT FOURTH SEMESTER ENGLISH EDUCATION DEPARTMENT FACULTY OF LANGUAGE AND SCIENCE WIJAYA KUSUMA UNIVERSITY

Supeno & Lusy Tunik ~ 247

POWER AND LANGUAGE: DISCOURSE ANALYSIS IN THE CLASSROOM

Majid Wajdi ~ 257

DEVELOPING A MODEL OF TEACHING WRITING BASED ON READING APPROACH THROUGH REFLECTION READING STRATEGY FOR EFL STUDENTS OF STKIP YPUP

Eny Syatriana & Rita Roswita Duyo ~ 269

INVESTIGATING THE TEACHER'S TEACHING APPROACHES IN BILINGUAL CLASSROOM

Syarifah Kurniaty K. ~ 277

WASHBACK EFFECT OF NATIONAL EXAMINATION ON EFL TEACHING

Sholeh Setiaji ~ 285

BLENDED CULTURE AS A MODEL OF ENGLISH LANGUAGE TEACHING AT SECONDARY SCHOOL LEVELS

Margana ~ 297

IELTS SPEAKING TEST: REVIEW, LIMITATION, STANDARDISATION AND REVISION TO ENSURE VALIDITY AND RELIABILITY

Andy

andyuni23@yahoo.com

University of Kanjuruhan

Malang, East Java

Abstract

IELTS is high stakes and gate keeping test for international students to participate at institutions and universities in English speaking countries. Considering its big impact to the future of test-takers (the students) as well as fulfilling the demand of language proficiency to succeed in the study at universities, and to understand better the need of different stakeholders. Therefore, ongoing research for development, revision, and also evolution have been done by the British Council, IDP: IELTS Australia and Cambridge English Language Assessment (jointly are the owners of IELTS) in order to address the problem, limitation and disadvantages, moreover mainly to assure its validity, reliability and consistency in doing the assessment, particularly of IELTS speaking test. This paper illustrates the review, limitation, standardization and revision. In terms of review, Brown and Hudson (2002, p.72) suggest that the disadvantages of performance test are "difficult to create; take considerable time to administer; may result in increased costs; causes logistical problems; creates reliability and validity problems; increase the risk of security breaches". For the limitation, verbal interview has subjective interpretation and also factor of gender of raters and test-takers that influence the scoring procedure. To maintain scoring consistency and steadiness, standardisation of test management is done through training, retraining, certification, monitoring, multiply rated, and video-rating. In 2001, interview format and criteria underwent revision to ascertain IELTS speaking test remains fair and unbiased. This paper tries to enlighten necessary consideration to speaking assessment developers to successfully provide evidence of representativeness of the skills and knowledge required.

Keywords: IELTS speaking test, review, limitation, standardisation, revision, validity, reliability, speaking assessment developers.

1 INTRODUCTION

The need to pursue higher education opens access to national even international tertiary institutions; it is therefore standardized language preparatory system to enter those institutions is needed to ensure the readiness and the success of the teaching and learning process. Institutions in English-speaking countries use test to measure the preparedness of international students (non-native speakers of English), one of English language test that is growing internationally trusted to be able to provide evidence of representativeness of linguistics and language skills required to succeed is the IELTS (International English Language Testing System) test. Considering its high-stakes and gate keeping test for international students both to study and work in English-speaking countries, therefore, the British Council, IDP: IELTS Australia and Cambridge English Language Assessment (jointly are the owners of IELTS) have done ongoing research for development, revision, and also evolution to address the problem, limitation and disadvantages of the IELTS test. Mainly to assure the validity, reliability and consistency of it in doing the assessment. There are four English skills tested in IELTS test: listening, reading, writing and speaking, among those skills, testing speaking skill has many drawbacks considering that it is performance test. As claimed by Brown and Hudson (2002, p.72) that

performance test are “difficult to create; take considerable time to administer; may result in increased costs; causes logistical problems; creates reliability and validity problems; increase the risk of security breaches”. Moreover, the limitation of IELTS speaking test which employ verbal interview has subjective interpretation and also factor of gender of raters and test-takers that influence the scoring procedure. Considering many variables and factors influencing the IELTS speaking test, therefore to maintain scoring consistency and steadiness, standardisation of test management is done through training, retraining, certification, monitoring, multiply rated, and video-rating.

2. IELTS BACKGROUND

IELTS, the International English Language Testing System, is designed to assess the language proficiency of candidates who need to participate in academic courses at institutions in English-speaking countries. It is produced by the University of Cambridge Local Examinations Syndicate (UCLES), and jointly managed by British Council and IDP: IELTS Australia. IELTS was launched worldwide in 1989 in order to replace its original English Language Testing Service (ELTS), which was developed in the late seventies and had been used around the early eighties.

IELTS functions as high stakes test as well as gate keeping test. Apart from the UK, Australia, and New Zealand, where IELTS is the preferred entry qualification among universities, over 170 universities in the USA now welcome students with the required IELTS score. Moreover, Low, E., Chong, S., & Ellis, M. (2014) affirm that “IELTS has been used by many universities in more than 120 countries as a means for ensuring baseline levels of proficiency in English required for entry into their respective programmes”.

3. TEST PURPOSE

McNamara (2000) discusses that IELTS is a performance-based test of English for academic purposes used to measure the English of international students wishing to study in universities in the English-speaking world. Even though the test modules are categorised into academic and general training versions, the only objective is education; either higher education or vocational and secondary education. According to IELTS Annual review 2001/2002, 81.06 percent of all academic candidates took IELTS in order to get higher education.

Low, E., Chong, S., & Ellis, M. (2014) state that “IELTS is not meant to certify whether candidates have passed or failed the test. Instead, institutions must determine the minimum selection band score for entry into their programmes and courses”. The International. Moreover, English Language Testing System (IELTS) has been selected because it has a high reputation as an internationally accredited test of language proficiency with a strong track record for language profiling purposes (Taylor & Jones, 2006).

4. TEST FORMAT

IELTS covers the four language skills; listening, reading, writing and speaking. The first three modules; Listening, Reading and Writing, must be completed in one day. The Speaking Module may be taken, at the discretion of the test centre, in the period of seven days before or after the other modules. All candidates take the same Listening and Speaking Modules but need to elect to take either the Academic or General Training versions in the Reading and Writing modules. While the Academic version assesses readiness to study in higher education, the General Training version focuses on the survival skills necessary in social and secondary education contexts. The tests are task-based, and are designed based on the criterion of the reality

in which candidates are likely to encounter. Candidates must complete all four components in order to receive a score. The total test time is 2 hours and 45 minutes.

A band score is awarded for each module, ranging from 1, where the candidate has no ability to use the language beyond a few isolated words, to a maximum of 9, where the candidate is the expert user of language. Each of the component scores is equally weighted and the overall band score is obtained by taking the mean of the sum of scores obtained in all four components. The overall band is calculated from the average of the band scores of each module, and this is reported on a scale at 0.5 intervals. There is no standard required band score for entry to academic courses as they vary in terms of linguistic demands. However, very generally speaking, a band score of around 6.0 is required for most undergraduate studies and 6.5 to 7.0 for post-graduate studies. Some institutions also specify a minimum score in each module. (IELTS Handbook, 2003)

5. THE REVIEW OF IELTS SPEAKING TEST

IELTS Speaking sub-test is concerned here in this essay since it is claimed to be the direct test, which is the most valid test of speaking. According to O'Loughlin (2001):

"speaking component' of the IELTS is an example of direct speaking tests and direct tests are the most valid procedures as measures of global speaking proficiency because of the close relationship between the test context and 'real life'. In other words, direct tests more authentically reflect the conditions of the most common form of 'real world' communication, face-to-face interaction" - O'Loughlin, 2001, p.6

Fulcher (1997, p.77) asserts the benefit of direct test that "...the oral proficiency interview (OPI) was popular...as a direct test of speaking ability, the OPI was seen to have high face validity". These claims are supported by Weir (1990, p. 12) who said "Tests of communicative language ability should be as direct as possible (attempt to reflect the 'real life' situation) and the tasks candidates have to perform should involve realistic discourse processing". However the topic of reliability in the direct test and speaking test has been argued. While Cronbach's alphas has been used to report the reliability and consistency of the Reading and Listening tests, Writing and Speaking tests' reliability is assured through training, certification and monitoring of examiners. (IELTS Annual Review 1998/1999, p.18) Speaking test, moreover, is performance test, not paper-and-pencil test like the other modules, so it's quite hard to make the score reliable and constant. This review, therefore, specifically focuses on the Speaking test.

6. IELTS SPEAKING TEST FORMAT

The Speaking module consists of an oral interview between the candidate and an examiner, which is recorded on audio cassette. Regarding task types, the revised version used since July 2001 has three main parts. In part one, candidates answer general questions about themselves and their life, then in part two the candidate is given a card about particular topic on which the candidate need to talk. The candidate has one minute to prepare before speaking at length one or two minutes. The examiner then asks one or two rounding-off questions. Finally in part three, the examiner and candidate discuss more abstract issues, which are thematically linked to the topic in part two. The overall interview takes about fifteen minutes in maximum.

7. LIMITATIONS

Brown and Hudson (2002, p.72) suggest that the disadvantages of performance test are "difficult to create; take considerable time to administer; may result in increased costs; causes logistical

problems; creates reliability and validity problems; increase the risk of security breaches”.

IELTS Speaking test is the direct, controlled-interview, performance test based on eliciting tasks. The more specific limitations found in this sort of speaking test are general practical constraints, e.g., the high administrative costs and payment for a large number of examiners as well as time consuming problem. Also claimed as another limitations are the test authenticity, and reliability and standardisation of examiners. In detailed and descriptive interviews, it is difficult to replicate all the feature of real life communication such as motivation, purposes and role appropriacy, so the test authenticity and criterion need to deal with the problem of reality (Weir, 1990)

The reliability of the test material, which is item-based, cannot be reported by Cronbach' alpha as done with Reading and Listening modules. So it depends on examiners who need to be highly qualified and experienced. It is thus very necessary for UCLES to make sure that the face-to-face training and re-certification process of examiners have to be done consistently.

Regarding examiners' steadiness, there is still no guarantee that candidates will be asked the same questions in the same manner, even by the same examiner. (Weir, 1990, p.66). A recent study by Brown (2003, p.1) asserts that “The interviewers differed...An analysis of verbal reports produced by some of the raters confirmed that these differences resulted in different impressions of candidate's ability: in one interview the candidate was considered to be more 'effective' and 'willing' as a communicator than in the other”. She supports this idea that the unpredictable or impromptu nature of the test interaction is also likely to lead to a lack of standardisation of examiners across interviews. Moreover, she mentions that the revised version of IELTS Speaking uses 'interlocutor frames' which constrain interviewer behaviour.

As the issue of examiner has been considered, the impact of gender is also concerned as a problem. O'Loughlin (2002) discusses the gender effect on IELTS interview. As a number of research in the field of language and gender (cited in O'Loughlin, 2002 as, Maltz and Borker, 1982; Tannen, 1990; Coates, 1993; Thwaite, 1993) suggests that male and female conversational styles are quite different, female conversational style is assumed to be collaborative, co-operative, symmetrical and supportive whereas its male equivalent is portrayed as controlling, unco-operative, asymmetrical and unsupportive. Reed and Cohe (2001) also claim by citing the work of Sunderland, 1995, and Porter, 1991, that gender of both raters and test-takers plays a role. Such claims imply that gender neutrality does not exist in the construct of communicative ability. However, O'Loughlin (2002) argues that most of this research reveals some kind of gender effect on test scores but the effect is not always the same; some studies point that candidates received higher score by male interviewers while others argue that female interviewers scored more. A case study, thus, was raised in O'Loughlin (2002) by having sixteen different students (eight male and eight female) and eight accredited IELTS interviewers (four male and four female) participated. Each of candidates were interviewed on two different times by a male and a female interviewer, totally 32 interviews, and each of the interviews were audio-taped as done in the real IELTS Speaking test. At the end, the results from the discourse and test score analyses were compared and it was claimed that gender did not have a significant impact on IELTS Speaking test in this case study. Both male and female participants showed their ability to make supportive and collaborative contributions to the interviews. Test scoring also does not depend on gender either. O'Loughlin concludes this case study, which seems to contradict former researches, that there might be other factors enhancing the impact of gender difference, such as characteristics of the test context and participants, the purpose of the test, the language being tested as well as the social identities of the interviewer and candidate (including their gender, age, ethnicity, and perceived status). It can be assumed from the case study that gender

bias is not the big issue in IELTS Speaking test compared to the other limitations.

The issue of gender, therefore, is concerned by IELTS stakeholders and owners which is portrayed by The IELTS Annual Review. (2010). In this review, the scores of both male and female candidates were displayed as part of the report for further research and future consideration regarding to IELTS validity and reliability. These figures below show the mean overall and individual band scores achieved in 2010 by Academic and General Training candidates according to their gender.

Table 1. The IELTS Annual Review (2010)

Mean band scores for female candidates					
	Listening	Reading	Writing	Speaking	Overall
Academic	6.1	6.1	5.6	5.9	6
General Training	6.2	5.8	5.9	6.3	6.1

Mean band scores for male candidates					
	Listening	Reading	Writing	Speaking	Overall
Academic	5.9	5.9	5.4	5.8	5.8
General Training	6.2	5.8	5.8	6.2	6.1

8. TEST DEVELOPMENT AND REVISION

UCLES itself has acknowledged the limitations in IELTS Speaking sub-test, so that the revision has been planned and done in order to develop a clearer specification of tasks, in terms of input and expected candidate output; to increase standardisation of test management by the introduction of an examiner frame; and to revise the rating scale descriptors. (Boddy, 2001) The plan for the IELTS Speaking test Revision Project was first draw up in 1998 and introduced worldwide in July 2001.

The revision project set out to revise the assessment criteria and rating scale as described below.

9. ASSESSMENT CRITERIA

From July 2001 IELTS employs new Speaking test format, with similar length to the former format. The revised Speaking test format comprises three phases as mentioned above, while the prior one comprises five phases which is claimed to push the candidate progressively to his or her 'linguistic ceiling' in phases 3 and 4. It has been argued that these phases of the test did not always elicit a richer performance; moreover, it led to variations in amount and type of examiner-talk. As a result, these 2 phases has been deleted so that candidates need no longer to move towards a 'linguistic ceiling'. (Taylor, 2001)

10. RATING SCALE

The former holistic or global rating scale for IELTS has been replaced with four analytical subscales: fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation. This decision is in line with the claim of Canale and Swain in McNamara (1996) that language knowledge includes sociolinguistic competence, strategic competence and

grammatical competence. The last competence includes “knowledge of lexical items and of rules of morphology; syntax; sentence-grammar semantics; and phonology” (McNamara, 1996, P. 61). As a result, the four subscales need to be discretely concerned.

In addition, Taylor and Jones (2001) raise the question from Lazaraton (1998)’s work that how well the holistic scale and its descriptors could articulate the key features of candidates’ performances at different levels or bands. They point out that a clearer specification of performance features at different proficiency levels might enhance standardisation of assessment. For this reason, they support the IELTS’ deconstruction of rating scale.

UCLES believes that the analytical scales are able to allow key features of candidates’ spoken language production at different proficiency levels to be recognized more easily. Overall speaker performance then can be reflected more sensitively in the assessment. The analytical approach also helps to focus raters’ attention on salient features and so contributes to standardization of assessment. (IELTS Annual Review, 2000/2001).

11. TEST VALIDATION

A high quality language proficiency assessment therefore begins with a test that has been validated by research and that shows close alignment between what the tests set out to measure and the assessment modes. In addition to selecting a reliable, internationally benchmarked test of language proficiency, it is equally important for teacher education institutes to be knowledgeable about how to interpret the test scores (Banerjee, 2003)

After the revision of IELTS had been done, UCLES’ validation work was carried out by focusing on “setting up an experimental study to investigate the assessment criteria and scale functioning. The research design involved gathering a sample of video performances using the revised IELTS test format and then arranging for these to be multiply rated by experienced IELTS examiners. The video-rating option was preferred on the grounds that examiners rating audio-performances are inclined to under-rate...” The subjects are 10 male and 10 female candidates with 15 different first languages scoring from IELTS Band 3 to Band 8. (Taylor and Jones, 2001)

12. EXAMINER TRAINING

With the new IELTS Speaking format, the retraining procedure is necessary due to test revision. Since 1999, UCLES has encouraged IELTS examiner to have face-to-face training. (Boddy, 2001). During the revision of IELTS speaking test, the new examine training material was developed to be collaboratively used with the prior method of face-to-face training. The new set of materials included an IELTS Examiner Introduction pack with accompanying video and work sheet and an IELTS Examiner Training pack with 2 accompanying videos and detailed Notes for Trainers. These training materials are given to examiners to train themselves before being ‘(re)trained’ during the face-to-face training session in IELTS centre. (IELTS Annual Review, 2001/2002)

Fulcher (1997, p.83) asserts “Rater training is designed to change an individual’s perception of the world so that he or she conforms to an institutional standard of rating; this is the definition of rater reliability...”

Bachman (1990, p. 76) supports that “Tests such as the oral interview ... involve the use of rating scales are necessarily subjective scored, since there is no feasible way to ‘objectify’ the scoring procedure”. IELTS Speaking test employs oral interview, therefore it is a ‘subjective’ test. He points out that (1990, p.76) “In an objective test the correctness of the test taker’s response is determined entirely by predetermined criteria so that no judgment is required on the

part of scorers. In a subjective test, on the other hand, the scorer must make a judgment about the correctness of the response based on her subjective interpretation of the scoring criteria". In relation with this, IELTS examiner training is really worthwhile to be conducted to maintain consistencies.

13. FEEDBACK AFTER REVISING

UCLES claims that "... feedback from both trainers and examiners was very positive and this is one measure of the success of the world-wide (re)training programme" (IELTS Annual Review, 2001/2002). Taylor (2001) also supports the benefit of the examiner frame developing during the revision that in the feedback, examiners appreciate the examiner frame since it allows them to focus their attention on assessing rather than thinking what to say next.

14. CONCLUSION

It can be said that the development and revision of IELTS Speaking test UCLES has progressed for many years is effective in terms of the ability to solve the problem of limitations, consistency and validity. The outcome of the IELTS revision consists of evolution of assessment criteria and rating scale; standardisation of test management is done through examiner training and (re)training as well as examiner frame, in addition also examiner certification, monitoring, multiply rated, and video-rating. Therefore, it can be assumed that IELTS Speaking test has been standardised worldwide to maintain scoring consistency and steadiness.

This paper tries to enlighten necessary consideration to speaking assessment developers to successfully provide evidence of representativeness of the skills and knowledge required. Assessment criteria and rating scale is worthwhile to be considered since it is the starting point of deciding the scoring rubric and doing the rating. Maintaining the rating by standardisation of test management is the next step to be paid attention to remain fair and unbiased. Consistency and clear cut format is compulsory for speaking assessment developers, especially those at local or national level to be at approximate level to the international calibre test like IELTS in order to replicate and ensure its quality and representativeness in providing necessary evidence of the English speaking proficiency. Doing interrater (or in IELTS terminology multiply rated) is another attempt to ensure test validity. Factors like subjective interpretation and also gender of raters and test-takers can possibly influence the scoring procedure in a test involving verbal interview like the IELTS speaking test.

Upon completion of addressing these aforementioned factors, local and national speaking assessment developers are able to create and develop a testing system with a sound approach to reliability and validity in providing evidence of the test takers representative language mastery to adequately succeed in studying and working in English-speaking countries both at higher educational institutions and workplaces.

At smaller scope namely English teachers who are willing to run a speaking test at their local institution, by starting to put those factors mentioned earlier into consideration in developing their own version of speaking test; the issue of validity and reliability can help them in facilitating their effort to be fair and unbiased to their students in doing the rating. In fact, if their speaking test version has adopted and adapted the IELTS speaking test format; this experience of sitting the test can be as preliminary practice in joining the real IELTS test in the future. As we know that joining the real IELTS test is quite a rare chance for most of the students in Indonesia. Limited cities which have IELTS authorised centers and also the costly test fees are the obstacles for the test takers. Therefore, the opportunity to join test like version of IELTS is necessary, since it is growing internationally accepted English test for

studying and working in English-speaking countries. As gate keeping test, IELTS opens the access to participate in international connection with possible better earnings and future both for the students as individual and as human resources of Indonesia (support the development of national human resources).

REFERENCES

- Alderson, J.C., Krahnke, K. & Stansfield, C. (1987). *Reviews of English Language Proficiency Tests*, Washington, DC: TESOL.
- Bachman, L. (1990). *Fundamental Considerations In Language Testing*. Oxford: Oxford University Press.
- Banerjee, J. (2003). Using Language Proficiency Test Scores: The Admissions Process in British university. Paper presented at the Language Testing research Colloquium, University of Reading, 22 -25 July.
- Boddy, N. (2001). The revision of the IELTS Speaking test. *JALT Testing & Evaluation SIG Newsletter*, 5 (2), 2-4.
- Brown, J. D. & Hudson, T. (2002) *Criterion-referenced Language Testing*. Cambridge: Cambridge University Press.
- Brown, A. (2003). Interviewer Variation And The Co-Construction Of Speaking Proficiency. *Language Testing Journal*, 20 (1),1-25.
- Davies, A., Brown,A., Elder, C., Hill, K., Lumley,T., & McNamara,T. (1999) *Dictionary of Language testing (Studies in language testing 7)*. Cambridge: Cambridge University Press.
- Fulcher, G. in Clapham, C. & Corson, D. (eds), (1997), *Encyclopedia of Language and Education. Language Testing and Assessment*, 7, 75-85.
- Low, E., Chong, S., & Ellis, M. (2014). Teachers' English Communication Skills: Using IELTS to Measure Competence of Graduates Froma Singaporean Teacher Education Program. *Australian Journal of Teacher Education*, 39(10). Retrieved at <http://dx.doi.org/10.14221/ajte.2014v39n10.5>
- McNamara, T. F. (1996) *Measuring Second Language Performance*. London and New York : Addison Wesley Longman
- McNamara, T. (2000). *Language Testing*. Oxford: Oxford University Press.
- O'Loughlin, K. (2001). The Equivalence of Direct and Semi-Direct Speaking Tests, Cambridge: Cambridge University Press.
- O'Loughlin, K. (2002). The Impact Of Gender In Oral Proficiency Testing. *Language Testing Journal*, 19 (2),169-191.
- The IELTS Annual Review. (1998/1999). Cambridge: University of Cambridge Local Examinations Syndicate.
- The IELTS Annual Review. (2000/2001). Cambridge: University of Cambridge Local Examinations Syndicate.
- The IELTS Annual Review. (2001/2002). Cambridge: University of Cambridge Local Examinations Syndicate.
- The IELTS Annual Review. (2010). British Council, IDP: IELTS Australia and University of Cambridge ESOL Examinations(Cambridge ESOL).
- The IELTS handbook. (2003). Cambridge: University of Cambridge Local Examinations Syndicate.
- Taylor, L. (2001). Revising the IELTS Speaking Test; Development In Test Format And Task Desing. *Research Note 14*. [online] : cambridge-efl.org Directory: news/index.cfm
- Taylor, L. & Jones, N. (2001). Revising the IELTS Speaking Test. *Research Note 11*. [online] : cambridge-efl.org Directory: news/index.cfm
- Taylor, L., & Jones, N. (2006). *Cambridge ESOL Exams and the Common European Framework of Reference (CEFR). Research notes 24/1* (pp 2-5), University of Cambridge ESOL Examinations.
- Weir, C.J. (1990). *Communicative Language Testing*, Hertfordshire: Prentice Hall International.



www.teflin2015.org

Udayana University Press

in collaboration with

English Department Faculty of Letters and Culture
and Post Graduate Study Program, Udayana University

Jalan Pulau Nias 13 Sanglah Denpasar

Bali Indonesia 80114

Email : sasingunud@gmail.com

ISBN 978-602-294-066-1



9 786022 940661 >