

Tamansiswa International Journal in Education and Science (TIJES)

<http://jurnal.ustjogja.ac.id/index.php/TIJES>

## **The New Speaking Progress Test Design for Economics and Business Students in Indonesia**

**Adhi Kusuma<sup>1</sup>, Warissara Audjam<sup>2</sup>**  
**Universitas Sarjanawiyata Tamansiswa<sup>1</sup>, Indonesia<sup>1</sup>**  
**Chiang Rai Rajabath<sup>2</sup>, Thailand <sup>2</sup>**

### **To cite this article:**

Kusuma, Adhi. & Audjam, Warissara (2019). The New Speaking Progress Test Design for Economics and Business Students in Indonesia. *Tamansiswa International Journal in Education and Science (TIJES)*, 1 (1), 21-27.



## The New Speaking Progress Test Design for Economics and Business Students in Indonesia

Adhi Kusuma<sup>1</sup>, Warissara Audjam<sup>2</sup>  
Universitas Sarjanawiyata Tamansiswa<sup>1</sup>, Indonesia<sup>1</sup>  
Chiang Rai Rajabath<sup>2</sup>, Thailand<sup>2</sup>  
adhikusuma@ustjogja.ac.id<sup>1</sup>

---

### Article Info

#### Article History

Received:

14 September 2019

Accepted:

30 October 2019

---

#### Keywords

Speaking Test  
IELTS model

---

### Abstract

This study is to see the application of a new designed speaking progress test for the English 4 subject which has not done yet at one of the university in Indonesia. This aims to assess students during their learning and needs, and to evaluate the course. The new speaking test design was a progress test run formally on semester 4. It was adopted the model of IELTS test which consisted of three parts, used the TSE scoring scale and ETS rater system. All aspects in this test follow the compulsory requirements of a test such as reliability, validity, practicality, and impact. The test was a face-to-face interview test (Weir, 2005, p. 143). Time allocated for the entire test was not more than 15 minutes. The data taken from the five chosen trial students of semester 5. The actual test design was intended for the student of semester 4, but the examiner and the rater were the real ones. However, the age, learning background, and the general proficiency level of these five trial students were nearly similar to the intended candidates for the test design. Since the TSE was not familiar to the students, the students got two kinds of score. The first final result of the TSE score were between 40–50. The second final result was scored using the range of 100. It showed between 66.7–83.3. Both of the final results showed that this new design really tested the progress of the students. The design met the students' needs and evaluates the course. This innovative model of test was proposed to be trialled in the progress test in semester 4 next year.

**Keywords: Speaking Test, IELTS model**

---

## Introduction

Speaking is one of the most demanding skills in the daily life. Every person needs to communicate with others through speaking. Speaking plays an important role in making a social interaction with another people in order to gain information. Thus, it is necessary for every student to have a good speaking skill and ability and test designed is made based on the IELTS Speaking Test section as the model, with an expectation that the designed test really measures the progress of students' speaking ability like the real IELTS speaking test does.

Hence, this research is intended to propose a new assessment test, putting weight on the process and the product of learning by adapting and modifying the IELTS test for the Economics Business Students. The new designed test is especially intended to measure the speaking ability of the fourth semester students of Economics Business Students.

The act of giving an assessment (test) always has a purpose. In one of the founding documents of modern language testing, (Carroll, 1961) states: 'The purpose of language testing is always to render information to aid in making intelligent decisions about possible courses of action'.

The purpose of such testing is primarily related to the needs of the teachers and learners working within a particular context. Tests that are under local control are mostly used to place learners into classes, to discover how much they have achieved, or to diagnose difficulties that individual learners may have. Although it is very rarely discussed, teachers also use tests to motivate learners to study.

In line to the above statement, the study is limited to the finding of a new speaking progress test for the fourth semester students of Economics Business Students by using the IELTS Speaking Test section as the model.

## Method

### The Target Test Participants

The new speaking test design is a progress test runs formally on the fourth semester of Economics Business Students. The test aims to assess students “*during learning*” (Cowie, B. & Bell, B., 1999, pp. 101-116) and needs, and to evaluate the course. Each student has a compilation module taken from IELTS speaking section. The intended candidates for the test are the fourth semester students of Economics Business Students and in upper-intermediate level. They are between 18 and 21 years old. However, the trial students for the test were five students of semester 5 chosen randomly.

### Grand Design of the New Speaking Progress Test

The test is a face-to-face interview test (Weir, 2005, p. 143) resembles the high standard of IELTS speaking test format which consists of three parts (Simon Haines, 2006, p. 7) with slightly difference on Part II. Time allocated for the entire test is 11-15 minutes. Part I, which weighs 10 %, deals with introduction and familiar topics such as the candidates’ family and hobbies (Cusack, B. & McCarter, S., 2007) This takes two minutes including the registration. Part II, which weighs 40%, is a seven-minute-extended session consisting of three minutes preparing a talk and maximum four minutes describing particular topics given by the examiner. The following-up section which exists on the IELTS model is moved on to the Part III due to the degree of difficulty. So, Part II has less critical thinking than in Part III. In Part II, the examiner shows the topics list, gives instructions to the candidate, and takes it back after the end of the Part II.

Part III is a five-minutes test, which weighs 50%, is a “constructed response format” (McNamara, 2000) consisting of a follow-up session and a discussion on more abstract ideas linked to the Part II. In this part, the candidate is allowed to answer based on his/her experience and knowledge if needed. The one-minute closing is to allow the candidate to relax and to ask question and the examiner gives feedback orally which is also recorded. All instructions in this test are clear and straightforward as the examiner reads from the interview outline. However, on the Part II, the instruction is also provided on the task sheet.

The Test of Spoken English (TSE) scale which has scale between 20 and 60 is chosen to score each part (ETS, 2001 cited in (Luoma, 2004, p. 69). As at the department uses a 100 scale, the final scores are in two versions: the TSE version and the 100-scale score with certain formulas.

### The Trial

According to (McNamara, 2000), the trialing, which involves careful design of data collection to see how well the test design is working on the group of people resemble in relevant respects such as age, learning background, general proficiency level to the target test candidates, is needed to the new test materials and procedures before its use under operational condition. The trial of the new speaking progress test of the English 4 speaking test at the one of isolated room at department with the help of the lecturer. She is the English lecturer for semester 3 and 5 of the department. In this trial, she was the examiner. She chose randomly her five students to do this trial test. The five chosen trial students, who are now in semester 5, were not the intended test participants but the examiner and the rater are the real ones. However, the age, learning background, and the general proficiency level of these five trial students were nearly similar to the intended candidates for the test design.

The test environment was set as the real test. The five students chosen were directed to the isolated room. Since this is a new test, before the trial, the lecturer explained the test model to make them familiar with the test. Moreover, she also gave fifteen minutes for them to relax. They were tested one by one with a strict time allocation for each part. This trial test was recorded digitally using a Nikon Coolpix camera.

(McNamara, 2000) states that feedback after the trial should be done to gather data about perceptions of the level of difficulty of particular questions, the clarity of the instructions, and general attitude to the materials and tasks. Therefore, after the trial, the examiner, the teaching assistant and the five trial students were asked to complete a short simple questionnaire capturing aspects that McNamara suggests. Finally, the interview records, the feedback sheets, and the filled questionnaires were gathered and analyzed.

## Results and Discussion

Because the TSE score is not familiar at the department, the students get two kinds of score. First is the TSE score. The results show that the final scores are between 40 and 50. Each part is score using the TSE scoring scale and weighting of each part. Consequently, each part has score with certain calculation. The final score shows that there are two students with final score of 50 or 83.3% of the TSE Scale and the rest gains final score of 40 or 66.7% of the TSE.

Table 1: The Trial Result and Discussion

No	Participant	Score				Percentage
		Part I	Part II	Part III	Final	
1.	Std #1	5	20	25	50	83.3%
2.	Std #2	4	16	20	40	66.7%
3.	Std #3	4	16	20	40	66.7%
4.	Std #4	5	20	25	50	83.3%
5.	Std #5	4	16	20	40	66.7%

As at Economics Business Students uses numeral score of 100, the score results from TSE scoring scale are counted into 100 score. Using the specific formulas, the final scores can be seen below.

No	Participant	Score			Final Score
		Part I	Part II	Part III	
1.	Std #1	8.3	33.3	41.7	83.3
2.	Std #2	6.7	26.7	33.3	66.7
3.	Std #3	6.7	26.7	33.3	66.7
4.	Std #4	8.3	33.3	41.7	83.3
5.	Std #5	6.7	26.7	33.3	66.7

In the score sheet given to the students, all these scores are provided. Even though the students get feedback from the examiner and the rater, by knowing each score for each part and listening from their recordings, students can analyze and do self-feedback. In the real test, this final score gained, then, is added to the final-test score to see whether one student passes the English 4 unit or not.

From the data above, it can be seen that the new test designed seems easy so that the test takers could get good score. However, there are several considerations regarding the results. First, the number of students who took the test trial was only five students. This is a very small data. (McNamara, 2000) states that on a trial of a new test design, at least 100 students are required. The number of students who will take the progress test in semester 4 at the department is around 200 students. Only picking up five students, of course, does not represent the entire students.

Second, the five students chosen were not the intended candidates. Those students passed English 4 last semester so that they experienced in describing and analyzing data from the figures. If the chosen students were those who were in semester 3, it would be difficult because they have not introduced to terms used in describing figures such as graphics, charts, and tables. In addition, the semester 3 students have not been familiar with the IELTS test and the interview test yet. By viewing the trial test result, it can be seen that the test is reliable viewed from the test item discrimination point. (Killen, R., 2005, p. 106) mentioned that the reliability increases if more test takers answer correctly as they know and understand. This is understandable as the test includes items which are familiar to the test takers' fields.

Third, there was too much time for preparing the talk in Part II given to the candidate. The three-minute preparation before describing the table was too long. The feedback from the examiner may be taken into account. She mentioned that because it was too long, the test taker could even write sentences in this three-minute preparation. The reason is may be the candidates learnt how to describe data from figures in their previous semester. As a result, it was easy for them to deal with Part II task. Fourth, the questions in Part III requires more time for students to answer. Almost all the test takers could not finish in five minutes. As the time allowed is only five minutes, the examiner had to stop the discussion even though sometimes the candidates still wanted to talk.

Regarding to the test usefulness, designing a test considers reliability, validity, authenticity, practicality and impact (Bachman, L. & Palmer, A., 1996, pp. 17-19). The new speaking test design completely respects and follows these considerations. To increase the reliability of the test, IELTS speaking test model, TSE scoring scale and ETS scoring system are used. (Killen, R., 2005, p. 105) states that the large number of question allows a better sample of the test participants' knowledge and understanding. The idea of making questions in this test is from English course books for Economics Business Students, which are close to the candidates' knowledge, and from IELTS course book which provides grammatical items used to describing figures. This is in line what (Douglas, 2000, p. 61) states that the combination between the language features and tasks may be set to meet the students' specific situation. The higher-order thinking (Killen, 2005, p. 14) is applied as the test takers are on the level where they are engaged in complex, and contextualized thinking to manipulate given information and abstract ideas to give meanings on information.

The TSE scoring scale is chosen as model as it combines holistic and analytic rating scales (Luoma, 2004, p. 69). Its holistic scale is to express the test takers' ability in one score while its analytic scale is to provide a number of criteria as a detailed guidance for the examiner (Luoma, 2004). Thus, it does not require much time to score. The ETS system; Online Scoring Network (OSN), enables raters to score anonymously (Hines, 2010). In other words, the raters only score what is in front of them. (Pickford, R. & Brown, S., 2006, p. 124) stated that the practical skills should be assessed by the most appropriate raters in the suitable place. The examiners and the raters of the new test are the lecturers who do not examine and score the same students. For example, the examiner of class A scores the candidates in the class B.

In term of test validity, the test uses criterion-referenced test as all tasks are targeted to the certain learning outcomes and address the criteria in questions (Stephen H. Moore, 2009, p. 7) by considering content relevance and representativeness to what have been taught and used for exercises (Killen, R., 2005, pp. 107-113). Based on the syllabus in the module, all materials taught in semester 4 are related to spoken English used in teaching elementary students and in completing the unit, the test participants are expected to be able to speak English in the teaching environments (Munandar, 2007). Thus, the contents of the test are in line with the syllabus and module.

There are two aspects of LSP testing: the authenticity of task and the interaction between language knowledge and specific purpose content knowledge (Douglas, 2000) ; (Hutchinson, 1987). (Nunan, 1993, p. 59) stated that a communicative task involved learners in comprehending, manipulating and producing in target language with meaning rather than the form as its focus. The table chosen on Part II is the authentic material got from the real teaching practice situation. By describing and interpreting data from the provided table, the candidates need to have language knowledge on this specific knowledge such as numbers, trends, comparison-contrast and cause-effect (Kerridge, 1990)

Moreover, students may face the same condition in the future and they need to use their language and knowledge in the same time in their real situations. Thus, the test focuses on more exposing the candidates' language productivity (Kitao, S. K., & Kitao, K., 1996). Moreover, in their final test, students present a product presentation which includes describing features, and in the next semester, they enroll at the TOEFL and IELTS preparation classes. In other words, this test is like a preparation for students' academic and future job needs so that students have high motivation on this test.

A test can be labeled as a practical test if it is easy and cheap to construct, administer, score and interpret (Hughes, 2003). These all points have been fulfilled in this new test. The test is planned to be done on week 4 to 8 of semester 4 and is recorded digitally. As there are 50 students in one class and the duration is for one meeting of 180 minutes, there are ten students tested in one meeting in an isolated room. Meanwhile, the forty other students have classrooms activities handled by the teaching assistant. As this test is not run in the exam period, supervisors are not needed so that it cuts the budget. The TSE scoring scale also enables the raters to score quickly because it does not look the detail features on speaking skills such as pronunciations and accents but mainly on the message delivered. Thus, it can be inferred that the new test is designed to challenge the students and teachers with new test procedures and systems, to measure students' progress and needs, and to evaluate the course.

(Hughes, 2003, p. 1) states that testing affects on teaching and learning known as backwash. On the teaching and learning process, every activity in the classroom is closely related to preparing for the test. The curriculum framework asks the lecturers to ensure that the students are fully aware of the assessment process and the assessment criteria. The syllabus should be redesign to meet the need of the test. The syllabus now focuses only on the preparing students for individual presentation where terms used to and how to describe figures are taught on the week 4 (Munandar, 2007). Because of this new test, the week 4 materials should be moved on week 1. On

week 2 and 3, students prepare for the progress test. In these weeks, the lecturers introduce IELTS speaking model and how to deal with it. Regarding the score, there are two scores provided because at the department applies different scoring system from TSE system. Moreover, this is due to the calculation for passing the unit which is a sum up of the progress and summative test. As this test is run on week 4 to 8, the teaching assistants have to be ready handling the class without the lecturers. It may be good as they can practice their capabilities in tutoring. Commonly, the candidates are easily getting stressed (Hughes, 2003, p. 124). To eliminate it, they have to be prepared and during the test, the introduction and closing parts are included to make them relax.

### **Feedback and Suggestion**

Regarding the test format, the examiner, the teaching assistant, and all the trial students agree that this new test design is applied at Economics Business Students. However, some considerations take into accounts which lead to the improvements for this new test design. Firstly, both the examiner and the rater give feedback to the student. The examiner, suggested that the rater should also give a feedback on the scoring sheet. It may be useful for students as the test is a progress test. This means that the students are still in learning progress. Thus, an additional sheet for feedback is attached to the score sheets delivered to the rater.

Secondly, the examiner of the trial also suggested that the recordings should be given to the students so that they could learn from their performance and mistakes. The suggestion is really good and will be more useful as the students will do self-feedback. (Lynch, 2009) states that one student may learn by comparing his/her performance with other students. Thus, giving the recordings to the students may promote self-feedback and peer feedback in the learning system at economics business school.

Thirdly, the rater system may invite problem. The examiner of the trial argued that this system will not run well due to the large number of candidates and timing. The number of candidates is around 500 while the raters are only ten lecturers. It means that one lecturer has to examine 50 students in his/her class then he/she has to listen to the 50 recordings from other class and scores them.

Regarding to the time, in giving the test, the lecturer does not need extra time as the test is run during the lecture week. However, in scoring, the lecturer has to bring the recordings and listen to them outside the lecture week. The problem will appear when the lecturer who is in charge of the scoring is very busy and does not have enough time to do scoring. To avoid this possible problem, the rater system may be changed in case this matter really happens. The lecturer, who plays as the examiner, should also the rater. It means that the examiner scores the candidate directly after his/her performance. Consequently, time consumed for testing one student should be lengthened to give time for scoring. However, if it does so, the reliability of the rater is decreasing. Thus, there is no final suggestion for this case as this is still debatable.

Fourthly, the test seems easy for students on semester 5, but this may be moderate and appropriate for semester 4 students. Only time allocated to the test needs to be altered. Thus, the revised test should redesign the time allocated for the test.

Fifthly, the new test design requires some training for examiners, raters and teaching assistants. As discusses above that the examiners who are also the raters of this test may face problem in practicality and timing in scoring, they need to be trained. Especially in scoring, the raters need to practice. (McNamara, 2000) suggests that to train the raters, they need to have a certain meeting where they practice doing moderation. In this case, they listen to the recordings of some candidates' performance, then, score them based on the scoring scale. In this meeting, having arguments is acceptable. Especially in scoring, the raters need to practice. (McNamara, 2000) suggests that to train the raters, they need to have a certain meeting where they practice doing moderation. In this case, they listen to the recordings of some candidates' performance, then, score them based on the scoring scale. In this meeting, having arguments is acceptable. The teaching assistant, also gave comments that the test system was good as he could have more time to practice in handling the classroom without the lecturer. However, he was quite nervous because of his lack of experience in teaching.

Sixthly, the absence of the shifting instruction for each part made the candidates comfortable. Generally, the trial students said that they did not realize that the test consisted of three parts because in every part, they said that the examiner did not mention each part. However, this made them comfortable as the interview flowed as it was like a discussion. Moreover, the examiner greeted them friendly and asked easy questions at the beginning. This made them relax and prepared before the next parts. Another possibility why the test takers were comfortable was that both the test takers and the examiner knew each other. This test also motivated them as they realized that in the

future, they have to take the similar test for their jobs' requirement. According to (Killen, 2005, p. 37), students with high motivation to succeed will demonstrate greater persistence and effort. From the trial test takers' feedback and Killen's idea, it can be predicted that the new design speaking test at business economics should be run in the real test and condition.

Seventhly, the idea of giving different type of figures on Part II for each test meeting may increase the reliability of the test. There were two students giving a suggestion that the types of the authentic figures used in Part II should be more than one. They meant that the figure used in every week during the test is different. For example, on week 4, the figure used is a table taken from a restaurant but on week 5, the figure used is a chart taken from a hotel. In other words, as the test is run in five weeks, there should be five test models as well. The reason why they suggested this system because they experienced that if one student had done his/her test, other students asked items in the test or even the questions. As a result, students who are tested later get higher score as they have known the questions and prepared the answers before the test. This is very interesting as this phenomenon is not related to the validity of the test content but this relates to the culture among the students. So, even though the test tasks have already kept track by the examiners and the raters, students who get the later test will have higher score. This, of course, will decrease the reliability of the test. If the test models are made in five models, it requires more time for the test designer to find other authentic materials. The hardest one is that the contents of all models must be equal. However, designing five models with the same contents and weighs should be done to keep the reliability of the test.

Eighthly, the need to have another trial is urgent if this new test will be operated at business economic school next year. (McNamara, 2000) mentions that in introducing a new version of a test, it needs a testing cycle which the new test design should be trialed, analyzed, revised, investigated the validity and the usefulness of the test under the operational condition, and periodically revised. It means that new designed speaking test still needs time for this new test design to be used at the department due to the need of this test to be trialed on the intended students. Due to this aim, this new test is proposed to be used in the mid-term test in semester 4 at the department. The empirical data gathered from this mid-term test will be analyzed and monitored. If the result shows rational scores for all candidates, the test can be run for the following semester 4.

### **The Final Version of the New Speaking Progress Test**

After analyzing the finding from the trial test result and feedback, the maximum time provided in Part II and III are altered. One candidate has to finish the test not more than 16 minutes. The allocated time for Part I is still same, 2 minute. Before the trial, the Part II required seven minutes for three-minute preparation of the talk and four minutes for the talk itself. After the trial, time allocated to the Part II is only six minutes: two-minute talk preparation and four-minute talk. There is two minutes added to the Part II, from five minutes to seven minutes follow-up and discussion section. Therefore, there is no significant change on the test tasks.

However, the procedure is altered slightly. The raters give feedback so that there is additional sheet for it. Moreover, the students receive their own recording completed with score sheets which include the examiner's and the rater's feedback. (Luoma, 2004, p. 173) stated the feedback may describe the strengths and weaknesses of the test participants in detail. This is due to self-feedback, peer feedback, and teacher feedback. According to (Lynch, 2009), "speaking logs" encourage learners to analyze and make written notes on their own recordings. When the students receive the recordings, they may listen to their own speaking record and make notes on their mistakes. It means that they do self-feedback. If all students do the same, they may compare each other in paired discussions. The written feedback given by the examiner and the rater may be discussed in the class by the teacher reformulation on several common mistakes.

The variety models for five speaking test meetings cannot be provided soon on this revised test. The reason lays on the availability of the authentic materials. The authentic materials which are known and close to the candidates' real life cannot get quickly. Lastly, the debatable issue related to the rater system, it is also still in progress. Thus, it has not been provided on the revised test yet.

### **Conclusion**

The new speaking test design for semester 4 at the department uses the model of IELTS test which consists of three parts, uses the TSE scoring scale and ETS rater system. All aspects in this test follow the compulsory requirements of a test such as reliability, validity, practicality, and impact. Because of its innovative model, this test is proposed to be trialed in the progress test in semester 4 next year. Even though the trial test did not find any

major barriers, in the real test there may be some problems which have not covered yet. The predictable problems from the unready lecturers, teaching assistants, and the students themselves may affect on the other aspects of this test. The lecturers who are also the raters need to be trained on scoring. The teaching assistants have to practice handling the classroom without the existence of the lecturer. The students have to be introduced to the new format of the test earlier. In other words, these three parties require to be prepared before the real test. Overall, this new speaking test design really tests the progress of the students, meets the students' needs and evaluates the course.

## References

- Bachman, L. & Palmer, A. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Carroll, J. B. (1961). *Fundamental considerations in testing for English language proficiency of foreign students'* (reprinted in H.B. Allen and R.N. Campbell (eds.) (1965) *Teaching English as a Second Language: A Book of Readings*. ed.). New York: McGraw Hill.
- Cowie, B. & Bell, B. (1999). A model of formative assessment in Science Education, *Assessment in Education*, 6(1). *Assessment in Education: Principles, Policy & Practice*, 101-116.
- Cusack, B. & McCarter, S. (2007). *Listening and Speaking Skills*. Oxford: MacMillan Publishers Limited.
- Douglas, D. (2000). *Assessing Languages for Specific Purposes*. . Cambridge: Cambridge University Press.
- Hines, P. E. (2010, January Tuesday). *ets.org*. Retrieved Oktober Monday, 2018, from How ETS Scores the TOEIC Speaking and Writing Test Responses: [https://www.ets.org/research/policy\\_research\\_reports/publications/report/2010/itjy](https://www.ets.org/research/policy_research_reports/publications/report/2010/itjy)
- Hughes, A. (2003). *Testing for Language Teachers. 2nd Ed.* . Cambridge: Cambridge University Press.
- Hutchinson, T. &. (1987). *English for Specific Purposes: A learning-centred Approach*. Cambridge: Cambridge University Press.
- Kerridge, D. (1990). *Presenting facts and figures revised by Ellis, M., O'Driscoll, N., & Pilbeam, A. (Eds)*. Helvetica: Longman Singapore Publishers Pte Ltd.
- Killen, R. (2005). *Programming and assessment for quality teaching and learning*. Southbank Victoria: Thomson.
- Killen, R. (2005). *Programming and assessment for quality teaching and learning*. . Southbank Victoria: Thomson.
- Kitao, S. K., & Kitao, K. (1996). *Testing communicative competence*. Retrieved November Monday, 2018, from iteslj.org: <http://iteslj.org/Articles/Kitao-Testing.html>
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- Lynch, T. (2009). Responding to learners' perceptions of feedback: the use of comparators in second language speaking courses. *Innovation in Language Learning and Teaching*, 3(2).
- McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.
- Munandar, A. (2007). *Modul Pratikum Bahasa Inggris 4*. Yogyakarta: Diploma Ekonomika dan Bisnis.
- Nunan, D. (1993). Task-based syllabus design: selecting, grading and sequencing tasks. In G. C. Gass, *Tasks in a pedagogical Context: Integrating theory and practice* (pp. 55-68.). Clevedon: Multilingual Matters.
- Pickford, R. & Brown, S. (2006). *Assessing skills and practice*. New York: Routledge.
- Simon Haines, P. M. (2006). *IELTS Masterclass*. Oxford: Oxford University Press.
- Stephen H. Moore, a. M. (2009). A Case study of an assessment task bank. *Prospect : an Australian journal of TESOL*, 24(1), 3-16.
- Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. New York: Palgrave Macmillan.

---

## Authors Information

---

### Adhi Kusuma

Sarjanawiyata Tamansiswa University  
 Batikan UH-III/1043 Yogyakarta, Indonesia 55167  
 Contact :  
 E-mail Address: [adhikusuma@ustjogja.ac](mailto:adhikusuma@ustjogja.ac).

### Warissara Audjam

Chiang Rai Rajabath University  
 80 Moo 9 Phaholyothin Road, Ban Du, Mueang, Chiang  
 Rai, Chiang Rai, Thailand.