

# Topic Familiarity and Extraversion in ESL Oral Assessment: A Conceptual Review

Nurul Iman Ahmad Bukhari

Faculty of Language Studies and Human Development, Universiti Malaysia Kelantan, Bachok, Kelantan, Malaysia, Faculty of Educational Studies, Universiti Putra Malaysia, Serdang, Selangor

Lilliati Ismail

Faculty of Educational Studies, Universiti Putra Malaysia, Serdang, Selangor  
Corresponding Author Email: [lilliati@upm.edu.my](mailto:lilliati@upm.edu.my)

Noor Lide Abu Kassim

Kulliyah of Education, International Islamic University Malaysia  
Email: [noorlide@iium.edu.my](mailto:noorlide@iium.edu.my)

Nooreen Noordin

Faculty of Educational Studies, Universiti Putra Malaysia, Serdang, Selangor

Abu Bakar Razali

Faculty of Educational Studies, Universiti Putra Malaysia, Serdang, Selangor

To Link this Article: <http://dx.doi.org/10.6007/IJARPED/v13-i3/21695>

DOI:10.6007/IJARPED/v13-i3/21695

---

**Published Online:** 29 August 2024

## Abstract

Oral performance among test-takers in proficiency English tests has shown to be prone to variability due to its rater-mediated nature. Test-takers could better lead their assessment preparation if they understood more about how to optimise their strengths, which can be identified as test-taker characteristics. The purpose of this study is to determine whether topic familiarity and extraversion can be linked to oral assessment facets using the Conceptual-psychometric framework of rater-mediated assessment (Eckes, 2015). To understand how these factors affect test performance and task engagement, it focuses on the relationship between test takers' topic familiarity and extraversion traits. By reviewing existing literature, this review proposed a conceptual framework that explains the concepts inherent in ESL oral assessment to investigate test-taker individual differences in topic familiarity and extraversion. This review seeks to contribute to a deeper understanding of how these factors interact and may influence the results of ESL oral assessments by carefully examining the relationships between individual differences such as topic familiarity and extraversion and oral assessment. Furthermore, it displays the significance of involving affective and cognitive dimensions of test-taker behaviour when developing assessment

instruments that are dependable and valid for English as a Second Language (ESL) learners. Future research directions and the implications of these findings for ESL assessment practices are also addressed.

**Keywords:** Oral Assessment, Test-Taker Characteristics, Topic Familiarity, Extraversion, Individual Differences.

## **Introduction**

In the field of English as a Second Language (ESL) education at the university level, there is a substantial focus placed on the task-based approach, which is in line with the Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001; Fischer, 2020). This emphasis is evident in the transformation of the Malaysian University English Test (MUET), which now integrates CEFR-aligned descriptors into its test specifications as part of the English Roadmap 2015-2025 initiative (Bijani et al., 2017). Originally conceived in 1999 as a high-stakes language proficiency assessment for post-secondary students, evaluating listening, reading, writing, and speaking skills, the MUET underwent a significant overhaul in 2021 to conform to CEFR standards, showcasing Malaysia's dedication to enhancing English communication (Chonghui, 2019). As a crucial placement exam for tertiary institutions and a prerequisite for certain job positions (Bidin et al., 2019; Mahmud, 2018; Mahmud et al., 2021; Rethinasamy & Chuah, 2011), the MUET carries substantial implications for test-takers, particularly those seeking admission to universities. The English Language Education Reform in Malaysia, guided by the Roadmap 2015-2025, endeavours to attain CEFR B2 proficiency among post-secondary students by 2025, bolstered by the Post-Secondary English Language Curriculum Framework (PSELCF) tailored for Form 6, Matriculation, and Foundation programmes, thereby equipping students for higher education undertakings. The endorsement of MUET by the Malaysian Examination Council and Cambridge Assessment English validates its alignment with CEFR standards (Geranpayeh & Ahmad Zufrie, 2018; Malaysian Examinations Council, 2019), with the expectation that university graduates possess at least B2 proficiency, enabling them to effectively understand more complex texts, handle more abstract topics and technical discussions, partake in technical discussions, and communicate comfortably with native speakers (Badrasawi et al., 2021).

While the MUET serves as a crucial assessment tool for measuring English language proficiency, it is essential to examine deeper into the factors that influence oral performance, particularly in the context of task-based language testing. This conceptual review paper aims to propose a conceptual framework for studying the effects of topic familiarity and extraversion on oral performance in the MUET speaking test. By exploring the context between these variables and their impact on test-takers linguistic abilities, this paper seeks to contribute to the understanding of individual differences in second language acquisition and their implications for language assessment.

## **The Research Gap**

For the proposed study, there are two distinctive research gaps derived from previous proponents related to test-taker characteristics in investigating speaking test validity: topic familiarity and extraversion in ESL oral assessment. According to Weir's (2005) socio-cognitive framework for validating speaking tests, test-taker characteristics can be divided into 3 categories: physical/physiological, psychological, and experiential. These research gaps form

the basis of inquiries into the problem to be explored. Details of the discussions on these research gaps are presented in the following sections.

The research on topic familiarity in oral performance, particularly in the context of high-stakes English language tests like the Malaysian University English Test (MUET), highlights several gaps that need to be addressed. Firstly, the impact of topic familiarity on test-taker performance has been largely overlooked in many tests (Banerjee, 2019; Brown, 2003; Huang et al., 2016; Qiu, 2020). This is concerning, especially in the MUET format, where the inclusion of different degrees of topic difficulty may affect test fairness towards test-takers who lack familiarity with a given topic (Huang et al., 2018). Secondly, the study of topic familiarity in oral performance has primarily focused on cognitive complexity, neglecting the potential influence of test-taker personality traits like extraversion (Khabbzbashi, 2017). Nakatsuhara's (2011) study on Japanese test-takers suggests that a test-taker's extraversion levels can significantly impact their discourse in group oral tests. This raises concerns about the potential effects of extraversion on speaking performance among Malaysian test-takers with varying degrees of extraversion. Lastly, the literature on extraversion and shyness among Asian test-takers (Berry, 1993; Bonk & Ockey, 2003; Nakatsuhara, 2011) highlights the need to consider cultural differences in the context of high-stakes language tests. This study aims to address these gaps by investigating the influence of extraversion on speaking performance among Malaysian test-takers in oral presentations. These research gaps underscore the importance of considering both topic familiarity and test-taker personality traits like extraversion in oral performance assessments to ensure test validity and fairness.

### **Topic Familiarity in Oral Performance**

In recent years, researchers have delved into the impact of topic familiarity on oral performance, revealing intriguing insights into how test-takers are affected during oral tests. Studies consistently demonstrate the effects of topic familiarity on oral performance (Khabbzbashi, 2017). Additionally, research has explored the connection between topic familiarity and test-taker attributes like confidence (Abu Kassim & Zubairi, 2006) and anxiety (Huang et al., 2016). Assiri (2019) investigated topic familiarity and facets in EFL monologue tasks, categorising topics into experiences/preferences and opinions/attitudes. While experiences/preferences led to more fluent language, opinions/attitudes required greater cognitive effort due to their unfamiliar nature, resulting in more complex and diverse language use. Opinions and attitudes showed a positive correlation with complexity, while preferences and experiences correlated with fluency and complexity.

### **Further Insights on Topic Familiarity in Task-Based Performance Research**

Topic familiarity remains a relatively unexplored area in Skehan's (1996), model of task-based performance. Bei (2010); and Skehan et al (2012), revisited past research by Skehan & Foster (1999), on narrative-retelling assessments, emphasising the impact of task design and learners' processing load on oral performance. Qiu (2020), introduced a novel construct on topic familiarity, illustrating content familiarity horizontally and task repetition vertically. Her study aligned with Skehan's cognitive process frameworks (Skehan, 2014, 2018) derived from Levelt (1989), suggesting an interaction between content familiarity and task repetition influencing learners' attention to form. By examining monologic tasks of 60 EFL learners in terms of complexity, accuracy, and fluency (CAF), Qiu's research shed light on the influence of content familiarity and task repetition, proposing a comprehensive construct of topic

familiarity. This interaction explains why task repetition enhances the structural complexity of learners' speech, aiding in clarity and the removal of redundant details. Therefore, it is vital to measure the variables that regulate topic familiarity of test-takers when selecting topics for oral assessment purposes.

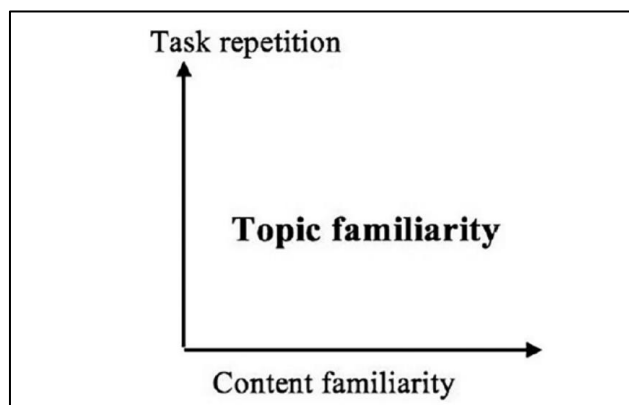


Figure 1: The Construct of Topic Familiarity (Qiu, 2020, p. 16)

### Extraversion in Oral Performance

There has been a significant amount of interest displayed by researchers in investigating the relationship between extraversion and oral performance in a second language (L2). Despite the fact that there is a widespread notion among teachers, academics, and students that extraverts are superior when it comes to language learning, particularly when it comes to the acquisition of oral skills, empirical investigations have produced contradictory results. According to the findings of a few studies, there is a positive connection between extraversion and oral performance. For example, Zafar et al (2017), found that around 69% of students who performed exceptionally well in speaking tasks were in the extrovert personality category. The findings of Bonk and Van Moere (2004), and Ockey (2006), demonstrated that candidates with greater degrees of extraversion tended to have better oral performance.

It is possible that extraversion interacts with other factors, such as the gender of the person taking the test, the qualities of the speaking test partner, and the emotional dynamics (O'Sullivan, 2000). However, the impact of extraversion appears to be varied. It has been found that the influence of extraversion on the outcomes of oral performance differs from study to study. Berry (1993), made the observation that extraverts did better in oral examinations when they were paired with other extraverts, whereas the scores of introverts were not affected by their partners. In contrast, Berry (2004), and Ockey (2011), discovered that extraverts who interacted with more introverted groups in group settings achieved greater ratings than those who did not have contact with such groups. On the other hand, Ahmad Bukhari et al (2023), Aljuaid (2022), Csépes (2002), Davis (2009), and Bonk & Van Moer (2004) showed that there were no significant differences in scores based on the levels of proficiency of peers.

According to the findings of qualitative studies (Iwashita, 1998; Norton, 2005), test-takers who are paired with partners who have a higher level of English proficiency may likely benefit from improved language input quality. In his study, O'Sullivan (2000) discovered a complicated relationship that involved three different factors: the proficiency levels of test-takers, their gender, and the gender of their test partners. In oral assessments, the number of participants

in the group can also affect the degree to which extraversion is emphasised. Extraversion was shown to be more visible in groups of four as opposed to groups of three, according to the findings of (Nakatsuhara, 2011). Ockey's (2011), research found that assertiveness, which is a component of the extraversion scale, was associated with a minor advantage in terms of communication skills and fluency with participants. Hence, although there is a widespread notion that extraverts are superior when it comes to oral performance in second language acquisition, the empirical evidence shows a contradictory picture. The influence of extraversion appears to be complex, as it interacts with a variety of elements, including the structure of the tasks, the dynamics of the group, and individual characteristics. To get a complete understanding of the connection between extraversion and oral performance in a second language, additional research is required.

### Conceptual Framework Development

The conceptual paper examines the theoretical foundation of the relationship between topic familiarity and extraversion towards oral performance. A conceptual-psychometric framework was produced by reviewing a well-known framework established by Eckes (2015) (see Figure 2) which outlines the whole process of a rater-mediated assessment, whether it is gauging speaking or writing performance. The framework begins the process presenting the distal facet involving external task factors such as features of examinees, features of raters, and features of situations. Subsequently, the involvement of proximal facets take part where the actual internal factors of the assessment is presented with the task difficulty and examinee proficiency (ability) directly affects performance. At this stage, rater severity, criterion difficulty, and scale structure is involved if chosen by the task designer. From the rater scores given, the assessment rating and observed score can be analysed using the Many Facet Rasch Measurement (MFRM) model and generate valuable data for tangible empirical measurements which shows examinee proficiency measures, rater severity measures, task and criterion difficulty measure, consistency analysis, rating scale functioning, interaction analysis, and differential facet functioning.

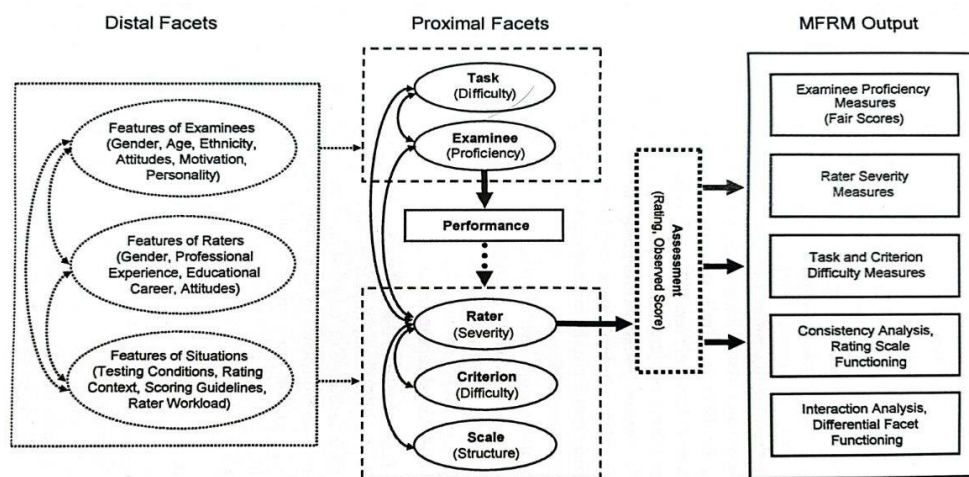


Figure 2: The Conceptual-psychometric framework of rater-mediated assessment (Eckes, 2015, p. 49)

### The Proposed Conceptual-Psychometric Framework for Topic Familiarity and Extraversion of Oral Assessment

Among oral assessment research, Aryadoust et al. (2021) cited that Eckes's (2015) Conceptual-psychometric framework for rater-mediated assessment is one of many prominent Rasch frameworks for language assessments. This framework has been implemented in various studies related to oral and writing performance assessment studies. For that reason, the proposed conceptual framework is adapted from Eckes's idea for a study which also implements the MFRM model, as shown in Figure 3.

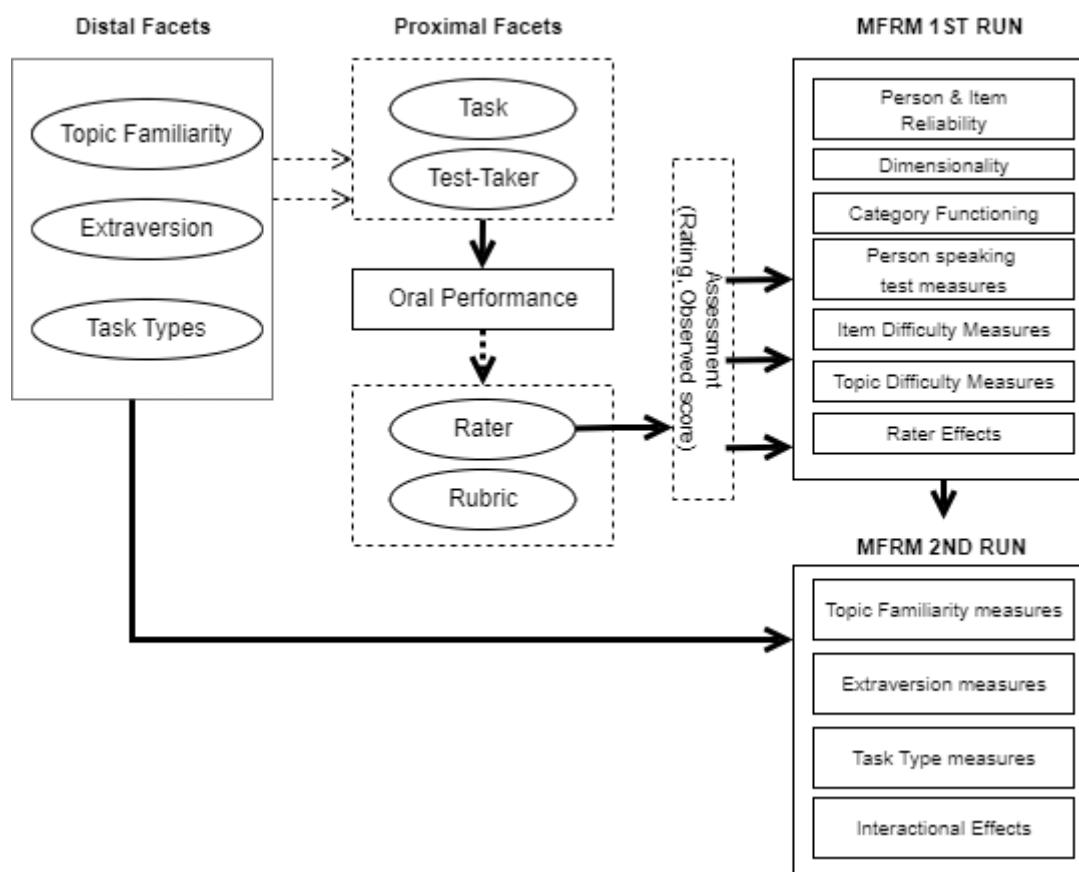


Figure 1: The Proposed Conceptual-Psychometric Framework of Oral Assessment for Topic

#### Familiarity and Extraversion Facets Using MFRM

In the proposed framework, attention is given specifically to the independent variables of topic familiarity and extraversion as the distal facets which are featured as test-taker characteristics; subsequently, in the framework, we include oral performance as the dependent variable. What makes this framework detailed is the MFRM output guided by Eckes (2015), where the expected outcome of future studies is closely guided.

#### Implications

The findings of the current study hold pedagogical, empirical, and practical significance for L2 oral assessment theories and practices. Practically, this study will enlighten language testing researchers about the influence of test-taker characteristics towards oral performance, specifically the interactions of topic familiarity and extraversion in oral presentation and group discussion of L2 speaking performance. Additionally, the present study could

potentially provide empirical evidence to support or contradict the scholarly claim that topic familiarity and extraversion would have influence on L2 oral performance.

### **Pedagogical Implications**

The findings of this study could well contribute to the body of knowledge in task-based language assessment, specifically in supporting Skehan's task difficulty perspectives in the Limited Attention Capacity Hypothesis (LACH). Various studies have been conducted on the different elements of task difficulty in language testing contexts (e.g. de Jong & Vercellotti, 2016; Figueras, 2019; Gan et al., 2017; Huang et al., 2016; Huang et al., 2018; Khabbzbashi, 2013, 2017; O'Grady, 2019; Préfontaine & Kormos, 2015; Taguchi, 2007; Tavakoli, 2009; Weir et al., 2006; Zalbidea & Sanz, 2020). However, there is still a need for findings on the influence of topic familiarity on speaking test task assessment. Studying the influence of topics in the speaking performance of oral presentation could help enrich understanding in the field of language testing towards theories and practices to best design speaking tests of enhanced fairness in terms of topical specificity. Studying test-taker characteristics will contribute very much to theories and models in the second language testing and assessment. One imperative contribution is that this study will be able to identify the extent of bias which test-taker characteristics could influence, specifically from a group of Form 6 MUET test-takers. Identifying sources of bias is often possible using knowledge of psychometric models such as Many Facet Rasch Measurement (MFRM). Hence, this study is hoped to contribute to models to identify possible sources of bias among MUET speaking test-takers which researchers and test designers could benefit from.

Investigating test-taker characteristics would lead to findings which would lend substantial evidence for the principle socio-cognitive framework for validating speaking tests proposed by Weir (2005), which was more recently adapted by Taylor (2011), by confirming whether the targeted test-taker characteristics can be applied as an influence towards context validity and cognitive validity.

### **Empirical Implications**

In spite of the fact that there is not many empirical research that have been conducted to investigate the effects of the two variables of interest, the findings of these studies are inconclusive. As previously demonstrated, the MUET serves as a prominent instance of a high-stakes examination in which the possible influence of the topic poses a threat to the test's validity (Messick, 1989). It is evident that the primary purpose of language proficiency assessments is to gauge competence, not pre-existing knowledge on specific topics. Thus, it is critical to assess the potential impact that could result from the MUET exam continuing to consist of "unfamiliar topics" in its speaking tests, as stated in its latest version of test specifications (Malaysian Examinations Council, 2019).

Examining test-taker characteristics is intended to contribute to the literature on MUET Speaking component testing by first guiding future researchers on how to analyse test-taker characteristics that may affect them in their institutional context. Second, this study hopes to be one of the first in Malaysia to inform researchers about CEFR-aligned MUET speaking test designs that may reduce the impact of construct-irrelevant variance caused by test-taker factors like topic familiarity and extraversion.

### Practical Implications

Practical implications of the proposed study include providing a scientific lens through which higher education administrators can more objectively evaluate the appropriateness of keeping the MUET speaking tests as part of proficiency tests for university entrance. The findings will also guide L2 practitioners in choosing oral assessments for their classes or preparing students for them. First, this project will expand on prior teaching contributions for MUET Speaking assessment to deepen the link between research into test-taker characteristics and instruction. This study will help teachers understand student traits that affect oral performance. Thus, teachers may better develop classroom interventions and tactics to ensure that linked traits do not hamper test-takers oral performance during the test. This study also proposes to examine extraversion, a personality trait that affects oral performance assessments (Almziad et al., 2020; Nakatsuhara, 2011; Zafar et al., 2017). Understanding how extraversion affects oral performance will help teachers help students maximise their strengths during tests. Teachers can also develop emotional skills to help them control extraversion when interacting with group members of varied extraversion degrees.

Likewise, an examination of experiential test-taker attributes would provide researchers with a more comprehensive understanding of how test design influences the cognitive dimensions of test-takers. Consequently, examining test-taker characteristics will contribute to the evaluation of the MUET speaking context in addition to aiding in the instruction of speaking. It is crucial to acknowledge that test-taker characteristics, while not quantified as language ability, have the potential to influence oral performance on the speaking test (Huang et al., 2016). Conducting research on experiential test-taker characteristics could provide researchers with valuable insights into the cognitive processes that Malaysian test-takers experience to a greater extent during the examination. Therefore, it is crucial that test designers and administrators take the greatest care regarding the potential impact of test-taker characteristics on MUET oral performance in order to i) guarantee that test-takers' scores accurately reflect their language proficiency rather than a product of influence from test-taker characteristics, and ii) prevent the covering of the intended speaking language proficiency construct (Purpura, 2004); and iii) make sure that test bias induced by test-taker characteristics does not affect the results (Weir, 2005).

Affective factors of psychological characteristics are considered vital in the MUET context due to the fact that MUET results serve as a criterion for admission requirements and programme allocation across numerous academic disciplines (Geranpayeh & Ahmad Zufrie, 2018; Harun et al., 2021; Sohaimi et al., 2021). Consequently, the outcomes of the MUET examination would significantly impact the test-taker's triggers beyond cognitive processes; the test-taker's performance on the speaking test could be significantly influenced by their emotions, personality, and attitude.

This study may also help the Malaysian Examination Council and test designers choose university-level speaking exam topics. The MUET may also utilise the information to build criteria to improve speaking exam fairness. Weir's socio-cognitive validation framework for the speaking test may be used to evaluate subjects' suitability for the target audience. By applying Weir's principles of context validity to the development of speaking assessments, the importance of considering the test-taker context could be brought to light. This research will



use CEFR-aligned criteria similar to those used by the Malaysian Examination Council. Hopefully, this research will shed light on the CEFR-aligned MUET speaking test.

### Conclusion

In this proposed study, a new conceptual framework of oral assessment is proposed by adapting a previous conceptual-psychometric framework of rater-mediated assessment by Eckes (2015) with a focus solely on test-taker facets of topic familiarity and extraversion as well as oral rater-mediated assessment. The conceptual framework comprises three main constructs, which are independent variables that include individual differences, specifically topic familiarity and extraversion. Oral performance is the dependent variable. This study offers a more thorough understanding of the impact of topic familiarity and extraversion on oral performance in rater-mediated assessments. This study addresses a gap in the current body of research by incorporating these specific qualities of test-takers into the assessment framework. This aids in guaranteeing that oral assessments are comprehensive and equitable. This theoretical contribution is crucial because it highlights the need to implement a tailored approach in language testing, which might potentially lead to more equitable assessment administration. The findings of this study are particularly relevant to educational settings that prioritise the assessment of English proficiency oral performance. They provide useful information for educators and policymakers to improve evaluation methods. Hence, this proposed conceptual framework is hoped to contribute to future studies where the outcome of these studies would serve as being many steps ahead in improving the inclusivity of test-takers in the design of oral rater-mediated assessment.

### References

- Kassim, N. L., & Zubairi, A. M. (2006). Interaction Between Test-taker Characteristics, Task Facets and L2 Oral Proficiency Test Performance. *Educational Awakening: Journal of the Educational Sciences*, 3(2), 139–159.
- Bukhari, N. I., Aziz, T. M. F., Hakim, M. L., Ismail, L., & Abdul Razak, A. (2023). Exploring the Link Between Extraversion Tendencies and ESL Proficiency. *BITARA International Journal of Civilizational Studies and Human Sciences*, 6, 19–27. <http://www.bitarajournal.com>
- Aljuaid, N. S. (2022). The Relationship between Personality Traits, Motivation Levels, and Speaking Performance among EFL Learners. *3L: Language, Linguistics, Literature*, 28(3), 181–200. <https://doi.org/10.17576/3L-2022-2803-12>
- Almziad, A., Samad, A. A., Razali, A. B., & Ismail, L. (2020). Using information gap activities to improve the speaking skills of Saudi Arabian introverted and extroverted learners. *PalArch's Journal Of Archaeology of Egypt/ Egyptology*, 17(6), 357–368.
- Aryadoust, V., Ng, L. Y., & Sayama, H. (2021). A comprehensive review of Rasch measurement in language assessment: Recommendations and guidelines for research. *Language Testing*, 38(1), 6–40. <https://doi.org/10.1177/0265532220927487>
- Assiri, M. S. (2019). Experiential versus attitudinal topic types and task performance in EFL monologues. *Indonesian Journal of Applied Linguistics*, 9(2). <https://doi.org/10.17509/ijal.v9i2.20238>
- Badrasawi, K. J. I., Kassim, N. L. A., Zubairi, A. M., Johar, E. M., & Sidik, S. S. (2021). English Language Speaking Anxiety, Self-Confidence and Perceived Ability among Science and Technology Undergraduate Students: A Rasch Analysis. *Pertanika Journal of Social Sciences and Humanities*, 29(3), 309–334. <https://doi.org/10.47836/pjssh.29.s3.16>
- Banerjee, H. L. (2019). Investigating the Construct of Topical Knowledge in Second Language

- Assessment: A Scenario-Based Assessment Approach. *Language Assessment Quarterly*, 16(2), 133–160. <https://doi.org/10.1080/15434303.2019.1628237>
- Bei., X. (2010). *The effects of topic familiarity and strategic planning in topic-based task performance at different proficiency levels*. Chinese University of Hong Kong, China.
- Berry, V. (1993). The Assessment of Spoken Language Under Varying Interactional Conditions. In N. Bird (Ed.), *Annual International Language in Education Conference* (pp. 491–507).
- Berry, V. E. (2004). A study of the interaction between individual personality differences and oral performance test facets. *Doctoral Thesis, PhD*(June).
- Bidin, S. J., Don, Z., Abdul Raof, A. H., Zubairi, A. M., & Mahat, N. I. (2019). *A Correlational Study Between MUET and IELTS*. Majlis Peperiksaan Malaysia.
- Bijani, H., Khabiri, M., & Branch, C. T. (2017). *Direct and Semi-Direct Validation: Test Takers' Perceptions, Evaluations and Anxiety towards Speaking Module of an English Proficiency Test*. 7(1), 25–41. [http://ttlt.stb.iau.ir/article\\_529069.html](http://ttlt.stb.iau.ir/article_529069.html)[http://ttlt.azad.ac.ir/article\\_529572.html](http://ttlt.azad.ac.ir/article_529572.html)
- Bonk, W. J., & Ockey, G. J. (2003). A many-facet Rasch analysis of the second language group oral discussion task. *Language Testing*, 20(1), 89–110. <https://doi.org/10.1191/0265532203lt245oa>
- Bonk, W. J., & Van Moere, A. (2004). L2 group oral testing: The influence of shyness/outgoingness, match of interlocutors' proficiency level, and gender on individual scores. In *Annual meeting of the Language Testing Research Colloquium, Temecula, CA*.
- Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language Testing*, 20(1), 1–25. <https://doi.org/10.1191/0265532203lt242oa>
- Chonghui, L. (2019). Boosting English Standards. *The Star Online*. <https://www.thestar.com.my/news/education/2019/06/09/boosting-english-standards>
- Council of Europe. (2001). the Common European Framework of Reference for Languages : Learning, Teaching, Assessment. *Council of Europe*. <https://doi.org/10.1017/S0267190514000221>
- Davis, L. (2009). The influence of interlocutor proficiency in a paired oral assessment. *Language Testing*, 26(3), 367–396. <https://doi.org/10.1177/0265532209104667>
- Jong, N., & Vercellotti, M. Lou. (2016). Similar prompts may not be similar in the performance they elicit: Examining fluency, complexity, accuracy, and lexis in narratives from five picture prompts. *Language Teaching Research*. <https://doi.org/10.1177/1362168815606161>
- Eckes, T. (2015). *Introduction to Many-Facet Rasch Measurement: Analyzing and evaluating rater-mediated assessments* (R. Grotjahn & G. Sigott (eds.); Second Edi). Peter Lang GmbH.
- Figueras, N. (2019). Developing and using tasks for the assessment of speaking. *Apples – Journal of Applied Language Studies*, 13(1), 133–149. <https://doi.org/10.17011/apples/urn.201903011693>
- Fischer, J. (2020). The underlying action-oriented and task-based approach of the CEFR and its implementation in language testing and assessment at university. *Language Learning in Higher Education*, 10(2), 301–316. <https://doi.org/10.1515/cercles-2020-2021>
- Gan, Z., Oon, E. P. T., & Davison, C. (2017). ESL students' oral performance in English language school-based assessment: results of an empirical study. *Language Testing in Asia*, 7(1). <https://doi.org/10.1186/s40468-017-0051-2>
- Geranpayeh, A., & Ahmad Zufrie, A. R. (2018). The Alignment Of Malaysian University English Test To The CEFR. *6th British Council New Directions in English Language Assessment*

*Conference: Standards in Learning Systems, October, 21.*

- Harun, H., Bosro, M. Z. M., Palpanadan, S. T., Ibrahim, M. Y., Sohaimi, N. A. M., & Jannaton, N. A. A. L. (2021). Malaysian public university students' challenging skills to pass Malaysian University English Test (MUET). *AIP Conference Proceedings*, 2347, 020031. <https://doi.org/10.1063/5.0052716>
- Huang, H. T. D., Hung, S. T. A., & Hong, H. T. V. (2016). Test-Taker Characteristics and Integrated Speaking Test Performance: A Path-Analytic Study. *Language Assessment Quarterly*, 13(4), 283–301. <https://doi.org/10.1080/15434303.2016.1236111>
- Huang, H. T. D., Hung, S. T. A., & Plakans, L. (2018). Topical knowledge in L2 speaking assessment: Comparing independent and integrated speaking test tasks. *Language Testing*, 35(1), 27–49. <https://doi.org/10.1177/0265532216677106>
- Huang, L. fen, Kubelec, S., Keng, N., & Hsu, L. (2018). Evaluating CEFR rater performance through the analysis of spoken learner corpora. *Language Testing in Asia*, 8(1), 0–17. <https://doi.org/https://doi.org/10.1186/s40468-018-0069-0>
- Khabbazbashi, N. (2013). *An investigation into the effects of topic and background knowledge of topic on second language speaking performance assessment in language proficiency interviews*. 2013(January), 395.
- Khabbazbashi, N. (2017). Topic and background knowledge effects on performance in speaking assessment. *Language Testing*, 34(1), 23–48. <https://doi.org/10.1177/0265532215595666>
- Mahmud, N. (2018). *Investigating the washback effect of the MUET as a university entry test on students in Malaysia* (Issue December). University of York.
- Mahmud, N., Lateh, N. H. M., Mahmud, N., Hassan, A. A., Mohamed, A. F., & Tarmizi, S. A. A. (2021). Washback impact of the MUET: The before and after effect of a high-stake university english test in Malaysia. *International Journal of Learning, Teaching and Educational Research*, 20(8), 1–17. <https://doi.org/10.26803/IJLTER.20.8.1>
- Malaysian Examinations Council. (2019). Malaysian University English Test (MUET). In *Batu Caves: Percetakan Warni*. [https://www.mpm.edu.my/images/dokumen/calon-peperiksaan/muet/regulation/Test\\_Specification\\_Regulation.pdf](https://www.mpm.edu.my/images/dokumen/calon-peperiksaan/muet/regulation/Test_Specification_Regulation.pdf)
- Messick, S. (1989). Validity in educational measurement. In *Educational Measurement*.
- Sohaimi, N. A., Harun, H., Palpanadan, S. T., Bosro, M. Z. M., Ibrahim, M. Y., & Abdul Latif Jannaton, N. A. (2021). Students' feedback on learning skills to enhance MUET results. *AIP Conference Proceedings*, 2347(1), 020046. <https://doi.org/10.1063/5.0052712>
- Nakatsuhara, F. (2011). Effects of test-taker characteristics and the number of participants in group oral tests. *Language Testing*, 28(4), 483–508. <https://doi.org/10.1177/0265532211398110>
- O'Grady, S. (2019). The impact of pre-task planning on speaking test performance for English-medium university admission. *Language Testing*, 36(4), 505–526. <https://doi.org/10.1177/0265532219826604>
- O'Sullivan, B. (2000). Exploring gender and oral proficiency interview performance. *System*, 28(3). [https://doi.org/10.1016/S0346-251X\(00\)00018-X](https://doi.org/10.1016/S0346-251X(00)00018-X)
- Ockey, G. (2011). Self-consciousness and assertiveness as explanatory variables of L2 oral ability: A latent variable approach. *Language Learning*, 61(3), 968–989. <https://doi.org/10.1111/j.1467-9922.2010.00625.x>
- Préfontaine, Y., & Kormos, J. (2015). The relationship between task difficulty and second language fluency in French: A mixed methods approach. *Modern Language Journal*, 99(1), 96–112. <https://doi.org/10.1111/modl.12186>

- Purpura, J. E. (2004). Validating questionnaires to examine personal factors in L2 test performance. In M. Milanovich & C. Weir (Eds.), *European language testing in a global context* (pp. 93–114). Cambridge University Press.
- Qiu, X. (2020). Functions of oral monologic tasks: Effects of topic familiarity on L2 speaking performance. *Language Teaching Research*, 24(6), 745–764. <https://doi.org/10.1177/1362168819829021>
- Rethinasamy, S., & Chuah, K. M. (2011). The Malaysian university English test (MUET) and its use for placement purposes: A predictive validity study. *Electronic Journal of Foreign Language Teaching*, 8(2), 234–245. <https://doi.org/10.2139/ssrn.2146007>
- Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions in narrative retellings. *Language Learning*, 49, 93–120.
- Skehan, P., Xiaoyue, B., Qian, L., & Wang, Z. (2012). The task is not enough: Processing approaches to task-based performance. *Language Teaching Research*, 16(2), 170–187. <https://doi.org/10.1177/1362168811428414>
- Taguchi, N. (2007). Task difficulty in oral speech act production. *Applied Linguistics*, 28(1), 113–135. <https://doi.org/10.1093/applin/aml051>
- Tavakoli, P. (2009). Investigating task difficulty: Learners' and teachers' perceptions. *International Journal of Applied Linguistics*, 19(1), 1–25. <https://doi.org/10.1111/j.1473-4192.2009.00216.x>
- Taylor, L. (2011). *Examining speaking: Research and practice in assessing second language speaking*. Cambridge University Press.
- Weir, C. J. (2005). Language Testing and Validation. In *Language Testing and Validation*. <https://doi.org/10.1057/9780230514577>
- Weir, C., O'Sullivan, B., & Horai, T. (2006). Exploring difficulty in speaking tasks: An intra-task perspective. In *International English Language Testing System (IELTS) Research Reports 2006: Volume 6*. IELTS Australia and British Council. <https://search.informit.org/doi/pdf/10.3316/informit.078778646705266>
- Zafar, S., Khan, Z. A., & Meenakshi, K. (2017). Extraversion-introversion tendencies and their relationship with ESL proficiency: A study of Chinese students in Vellore, India. *Pertanika Journal of Social Sciences and Humanities*, 25(2), 687–703.
- Zalbidea, J., & Sanz, C. (2020). Does learner cognition count on modality? Working memory and L2 morphosyntactic achievement across oral and written tasks. *Applied Psycholinguistics*, 1171–1196. <https://doi.org/10.1017/S0142716420000442>