Japanese Teachers' Attitudes Towards Classroom-based Assessment

Adam Murray
University of the Ryukyus
Taiko Tsuchihira
Seitoku University

Reference Data:

Murray, A. & Tsuchihira, T. (2023). Japanese teachers' attitudes towards classroom-based assessment. In P. Ferguson, B. Lacy, & R. Derrah (Eds.), *Learning from Students, Educating Teachers—Research and Practice*. JALT. https://doi.org/10.37546/JALTPCP2022-23

In this paper, we report on a pilot study about Japanese Teachers of English (JTEs) beliefs about classroom-based assessment (CBA). For this study, we used a revised version of the Classroom-based Assessment Self-Efficacy Scale (CBA-SES) which consists of four sections: teaching context, teacher beliefs, self-efficacy, and teaching practices. We administered this instrument to a convenience sample of 29 JTEs to get a better understanding of their work in various teaching contexts (primary, secondary, post-secondary) with the intention of revising the instrument and conducting a large-scale study next year. This indicated that respondents had strong teaching beliefs, with teachers at the university level having the strongest beliefs. Also, there was a connection between experience and beliefs. As teachers gained more classroom experience, their beliefs became stronger. We found a correlation between the tier of institution as measured by *hensachi*. Finally, teachers believe there was a need for assessment tasks that resemble real-life language use.

本稿は、日本人英語教師（JTEs）の教室内評価に関する信条について行ったパイロット研究について報告する。自己評価測定尺度（CBA-SES）の改訂版に準拠し、4つの範疇、すなわち学校種などの背景(5項目)、信条(11項目)、自己効力感(10項目)、教育実践(12項目)について質問紙を作成した。様々な学校種（小学校、中学校・高校、大学など）で教えているJTEsの現況についてより理解を深めるため、さらなる改訂を意図し、より多くの参加者に同尺度を実施するにあたり、今回は２９人のJTEsからの回答を分析した。結果として、JTEsは強い教育信条を持っており、特に大学教員にその傾向が最も強いことがわかった。教師の信条は経験と有意な関係があり、教室での経験を積むにつれて信条が強くなることがわかった。さらに、回答からは偏差値による学校レベルに応じた違いが認められた。また、教師が実際の言語使用に似せた測定タスクが必要だと考えていることも明らかになった。

In recent years, the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) has been using can-do-type criteria along with setting CEFR (Common European Framework of Reference) Levels of A1 and A2 as targets for junior and senior high school graduates to place a greater emphasis on actual language output. As a result, boards of education nationwide are implementing performance-based testing in their courses. For example, as outlined in the English Education Improvement Plan for 2022 (MEXT, 2022), there should be five speaking tests and five writing tests in junior and senior high school courses such as English Communication I. To better equip pre-service and in-service teachers with the skills and knowledge necessary for performance-based assessment, the "Core Curriculum for Teacher Education" (MEXT, 2016) clearly states the importance of developing teachers' ability to implement valid and proper student evaluation. However, it is unclear how well Japanese Teachers of English (JTEs) are dealing with these changes to the curriculum. In this paper, we report on an ongoing research project to develop an instrument to measure JTEs attitudes and beliefs about Classroom Based Assessment (CBA), and how CBA is implemented in their classrooms.

Background

In order to effectively implement classroom-based assessment, teachers need to be capable of managing all aspects of learning in their classrooms. Classroom teachers need to assess the needs of their students, set learning objectives, plan teaching and learning strategies, implement the instructional plan, and evaluate the instructional outcomes (IGI Global, n.d.). Because many aspects such as learning objectives (CEFR

Levels) and learning materials have been determined by MEXT's Course of Study, the main responsibilities of JTEs are implementing the instructional plan and evaluating the instructional outcomes. To do this, it is beneficial that JTEs have elevated levels of assessment literacy.

There are many descriptions and definitions of assessment literacy. Simply described, assessment literacy is practical knowledge (skills) along with "relevant background in measurement and language description" (Davies, 2008, p. 328). Fulcher (2012) defines assessment literacy in relation to three dimensions (practice, concepts, context) as:

> The knowledge, skills and abilities required to design, develop, or evaluate, large-scale standardized and/or classroom-based tests, familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice. (p. 125).

Taylor (2013), however, proposed the following elements: knowledge of theory, technical skills, principles and concepts, language pedagogy, sociocultural values, local practices, personal beliefs, and scores and decision-making. Regardless of which definition of assessment literacy is used, there are many theoretical and practical aspects to be considered.

Because JTEs are responsible for assessment, they need to have a working understanding of CBA. In general terms, CBA is any ongoing assessment that is conducted by those directly responsible for the teaching and learning (Davison & Leung, 2019). Hill and McNamara (2012) described the various forms of assessment and uses along with the roles of the teachers and learners: "any reflection by teachers (and/or learners) on the qualities of a learner's (or group of learners') work and the use of that information by teachers (and/or learners) for teaching, learning (feedback), reporting, management or socialization purposes" (p. 396). CBA consists of four phases: planning (the tasks along with the relationship to instruction), framing (if and how students are informed), conducting (types such as formal and incidental), and using (how it is used) (Hill & McNamara, 2012).

Because we are interested in how JTEs are coping with this stressed importance of being able to apply their prior knowledge of assessment by integrating classroom-based assessment in regular classroom practices, teacher self-efficacy (TSE) is a key factor that should not be overlooked. TSE beliefs can be defined as "teachers' beliefs in their abilities to support learning in task-, domain- and context- specific cognitive, metacognitive, affective and social ways" (Wyatt, 2018, p. 93). In recent years, there has been a growing interest in the self-efficacy of language teachers in Japan (Nishino, 2012; Thompson & Woodman, 2018; Wicking, 2017; Yada et al, 2019). In these studies, the researchers examined several areas such as teaching practices, L2 self-confidence, and communicative language teaching self-efficacy.

Our Objectives

During the 2019 academic year, the authors were part of a research group who created an instrument named the Classroom-Based Assessment Self-Efficacy Scale (CBA-SES) (Murray et al., 2020) which was based on the four phases of assessment (Hill & McNamara, 2012), the three dimensions of teacher assessment literacy (Fulcher, 2012), and the Michigan Assessment Literacy Standards (Michigan Assessment Consortium, 2017). The instrument consisted of three sections: teachers' beliefs, teachers' self-efficacy, and teaching practices. This instrument was piloted with 30 teachers. The respondents generally believed that tests for students should reflect real-life language use and that they could make and utilize such tests in their classrooms. They also felt that effective feedback and the use of learning targets are also important. However, some of the respondents reported having difficulty answering these questionnaire items because self-analysis and reflection were needed. Additionally, they rarely had time to think about these general beliefs and principles because their focus was on their day-to-day responsibilities.

About the instrument itself, the respondents commented that some of the items needed to be refined to better reflect the varied working contexts of English teachers in Japan. For example, teachers' attitudes may vary depending on the kinds of evaluation (formative or summative), and how much the school curriculum (at the primary and secondary levels) is dominated by entrance exam preparation. The institution (private vs. public, and/or competitive vs. regular schools) itself had an influence as well.

One objective of this pilot study was to gain insights into the beliefs, classroom practices, and the self-efficacy of JTEs in various teaching contexts. Additionally, we wanted to test the revised version of the CBA-SES to see how appropriate it is for JTEs, particularly those teaching at junior and senior high schools. Results will be used to inform a larger study to be conducted in the 2024 academic year.

Methods

Based on the feedback about the 2019 study, the CBA-SES was revised. Like the previous version, the instrument had sections about beliefs, self-efficacy, and classroom practice. An additional section was added about educational setting (junior or senior high

school, university, other), type of institution (public or private), and tier of institution (standardized test scores/*hensachi*). Furthermore, there were items about biographical information such as age and years of teaching experience.

Participants

An online questionnaire (via Google Forms) was distributed in Autumn 2022. Because the instrument was still in the piloting stage, it was decided that a convenience sample would be appropriate at that time. The researchers sent emails and asked colleagues to complete the questionnaire. Unfortunately, we were unable to get as many responses from junior and senior high school teachers as anticipated. Some teachers commented that it was a busy time of the academic year and that a different time of the year might be more convenient for data collection. In total, 29 participants completed the questionnaire. All the participants gave informed consent and could opt out at any time. The questionnaire was completely anonymous, and no personal identifiers were collected. Figure 1 shows the educational settings where the respondents were teaching. Most of the participants (79% / 23 people) were teaching at the university level. There were only four participants at the junior and senior high school levels and unfortunately, there were no elementary school teachers. We hoped to have more elementary school and junior high school teachers because we wanted to learn about how they are dealing with performance-based assessments within current MEXT educational reforms. In addition, some of the new items (e.g., public versus private institutions) were intended for teachers at the primary and secondary levels. Based on the feedback from the previous version of the questionnaire, we added more items about the teaching context.

Figure 1

Participants' Educational Settings

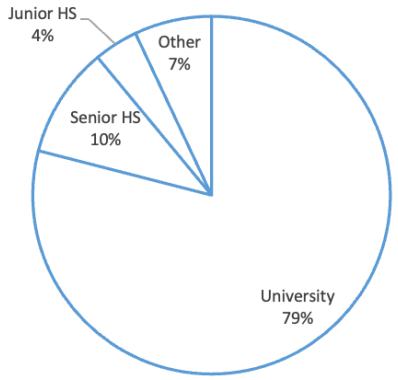
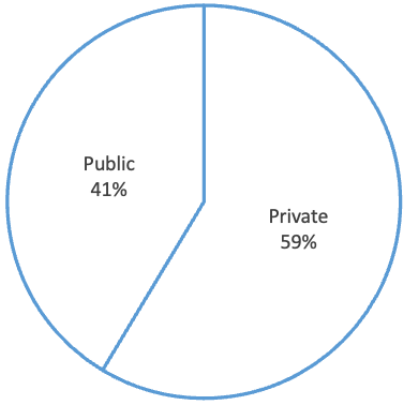


Figure 2 provides data about the type of institution. The distribution of institution type was comparable with 59% (17) of the participants teaching at private schools and 41% (12) at public ones.

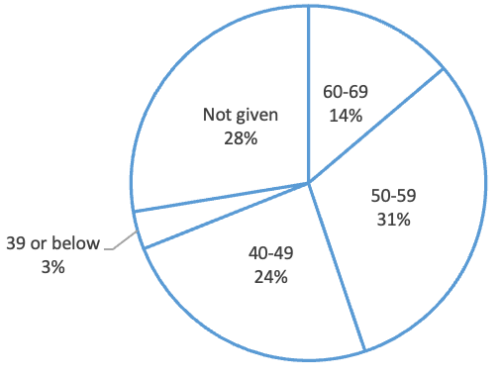
Figure 2

Participants' Institutional Type



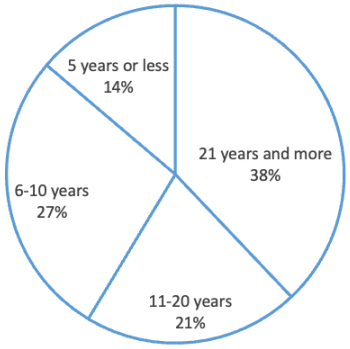
Another related new item was about the competitiveness of the institution. One widely published measurement of competitiveness is *hensachi*, which indicates the performance of students admitted into an institution or department. It is a norm-referenced test score with a mean of 50 and a standard deviation of 10 (Goodman & Oka, 2018), where the higher the score, the more competitive the institution is. Figure 3 shows which tier of institution the participants are members of. More than half of the participants (54% / 13 people) belonged to institutions that were one or two deviations higher than the mean. This means that the participants teach at competitive institutions. Many participants (28% / 8 people) did not report their *hensachi* scores. The remaining participants (27% / 8 people) belonged to less competitive institutions.

Figure 3
Participants' Institutional Tier



The final biographical question inquired about experience and Figure 4 provides information about years of teaching experience. The largest group were those with 10 or less years of experience (42% / 12 people). The next group of participants were veteran teachers with more than 21 years of experience (38% / 11 people), and the final group (21% / 6 people) had between 11 and 20 years of experience.

Figure 4
Participants' Teaching Experience



It should be noted that the teachers who participated in this study might not be representative of all JTEs. It is possible that these teachers are more interested in assessment than the typical teacher because of their willingness to complete our online questionnaire. Also, this is a convenience sample because most, if not all, the participants are known by one or both researchers. Keeping these facts in mind, the interpretation of the results should be regarded with some caution.

Questionnaire

The revised version of the questionnaire (Murray et. al, 2020) which was based on items about teacher disposition (Michigan Assessment Consortium, 2017), CLT self-efficacy (Nishino, 2012), and *The Japanese Teacher of English Teacher Efficacy Scale* (Thompson & Woodman, 2018) consisted of four sections: teaching context (five items), teacher beliefs (11 items), self-efficacy (10 items), and teaching practices (12 items). Belief refers to what teachers think effective (or ideal) teachers should do, while self-efficacy indicates how capable they think they are. The final section, teaching practices, is what they are doing in their classrooms. Except for the teaching context section, all items utilized 5-point Likert-type scales. The items on teachers' beliefs were on what ideal teachers do. For example:

- Teachers should understand and be able to use tests.
- Multiple assessments provide a more balanced picture of a student.

- Quality assessments are a critical attribute of effective teaching and learning.

(See Appendix A for all of these items along with their Japanese equivalents)

For each of these items, the participants indicated the strength of their agreement or disagreement.

In contrast to beliefs, the items concerning self-efficacy were about what teachers think they can do. For example:

- I can select and use various assessment methods appropriate to assessment purposes and learning targets.
- I can use grading practices that result in grades that are accurate, consistent, meaningful, and supportive of learning.
- I can use assessment results appropriately to modify instruction to improve student achievement.

For each of these items, the participants indicated how well they could do something.

Finally, there were items on their teaching practices. In other words, these items were about how assessment was integrated and how often it was conducted. For example:

- I use various assessment methods.
- I use formative assessment.
- I use summative assessment.

For each of these items, the participants indicated how often something was done. On the Google Form, each question was displayed in Japanese. The data were analyzed using Microsoft Excel by the authors.

Results

The results will be first given for all sections, followed by some comparisons of some items. Tables 1 and 2 show the overall descriptive statistics. Figures 5 to 10 show the results of the teachers' ratings on belief and practice items related to setting, tier, age, and experience.

For each of the statements about beliefs, efficacy, and classroom practices, a 5-point Likert scale was used. If a participant had maximum levels of agreement, self-efficacy, and classroom practices, a score of 165 would be given. Table 1 shows the descriptive statistics. The average score for the participants was 125.83 (*SD* = 15.94). Because the values for skewness and kurtosis are within generally agreed-upon ranges, the data can be considered a normal distribution.

Table 1

Descriptive Statistics

N	Range	Min.	Max.	Average (SEM)	SD	Variance	Skewness (SEM)	Kurtosis (SEM)
29	58	100	158	125.83 (2.96)	15.94	254.36	0.436 (0.434)	-0.556 (0.845)

Beliefs

Figure 5 shows the average scores for the belief items for each group of teachers. All the participants showed high levels of agreement with all of the statements. However, a significant difference was found with B1 (Teachers should understand tests and be able to use them). University teachers agreed most strongly with this statement (4.87) while those in the junior high school and other contexts had the lowest level of agreement (4.0).

Figure 5

Average Responses on Belief Items

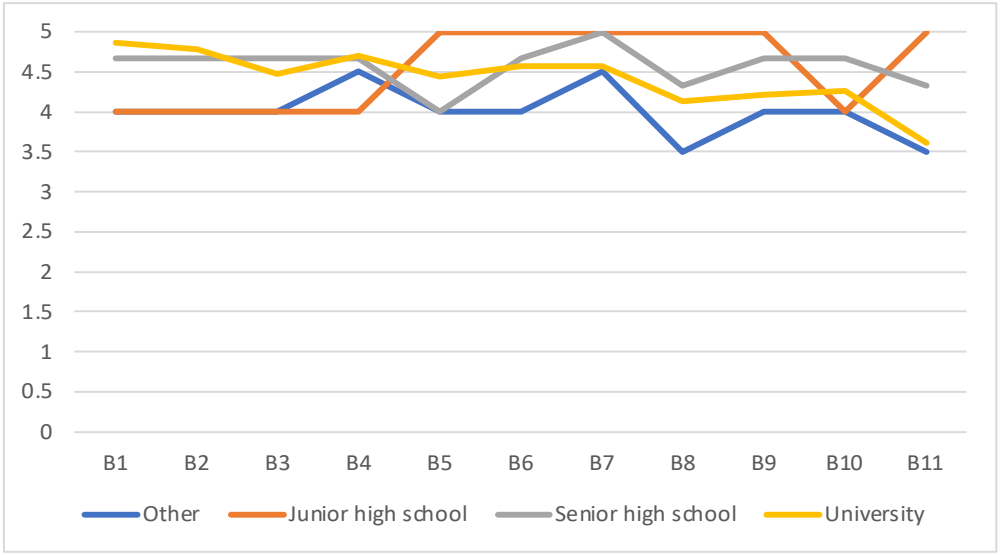


Figure 6 shows the average scores for the belief items and tier of institution. For most of the items, there were high levels of agreement. A significant difference was found with B11 (Language tests should resemble real-life language use). The teachers at the highest tier of institution (*hensachi* 60 or higher) had the lowest average score of 2.8. The one teacher at the lowest tier of institution (39 or below) had a score of 3. Teachers at above average institutions (50-59), and those teachers who had not given *hensachi* scores, had average scores of 3.75. Finally, the group (40-49) had the highest score of 4.43.

Figure 6
School Tier and Belief Items

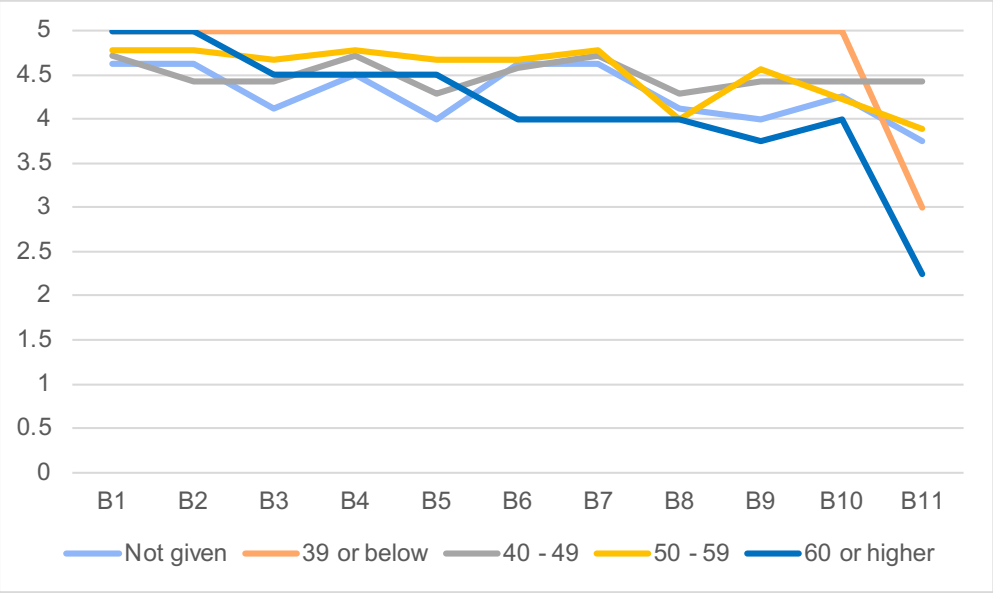


Figure 7 shows the average scores for the belief items and age. A significant difference was identified for B5 (Assessment results should be used to make instructional decisions to improve student learning). Older teachers (60 or over) had higher scores of 4.40 while the youngest teachers (20-29) had an average score of 3.33.

Figure 7
Age and Belief Items

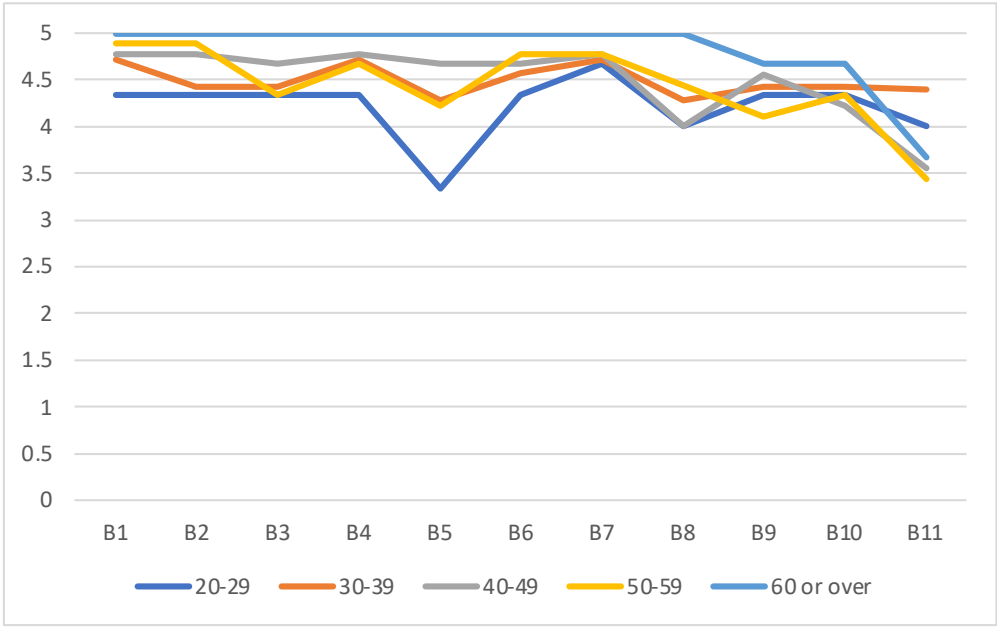
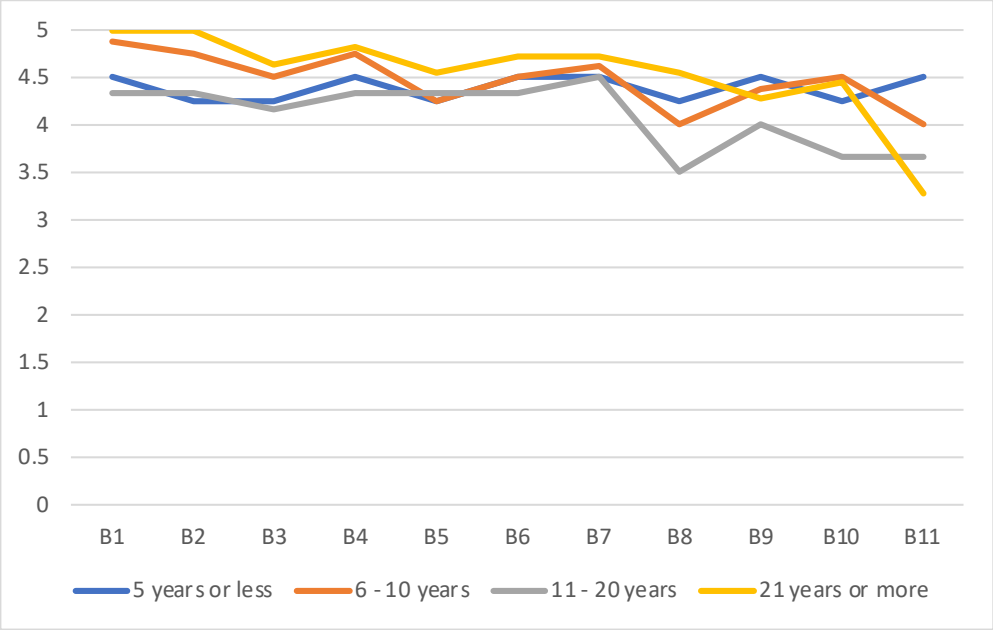


Figure 8 shows the average scores for the belief items and teaching experience. Two significant differences were identified between the groups. For B1 (Teachers should understand and be able to use tests), the veteran teachers (21 years or more) showed strong agreement with an average score of 5. The next highest group (6-10 years) had 4.88 and was followed by new teachers (5 years or less) with 4.5, and the lowest score was 4.33 for the experienced group (11-20 years). The next significant difference was found for B2 (Multiple assessments provide a more balanced picture of a student). A similar pattern to B1 emerged, where the veteran teachers (21 years or more) showed strong agreement with an average score of 5. The next group (6-10 years) had 4.88, followed by new teachers (5 years or less) with 4.5 and experienced teachers (11-20 years) with 4.33.

Figure 8

Experience and Belief Items



Practice

Figure 9 shows the average scores for the practice items and age. A significant difference was found for P5 (I provide timely, descriptive, and actionable feedback to students). The highest group (60 or over) had an average score of 3.83 followed by the 50-59 group with 3.66. The middle-aged group (40-49 years) had a score of 3.55. The youngest groups had scores of 2.6 (30-39 years) and 3.33 (20-29 years).

Figure 9

Age and Practice Items

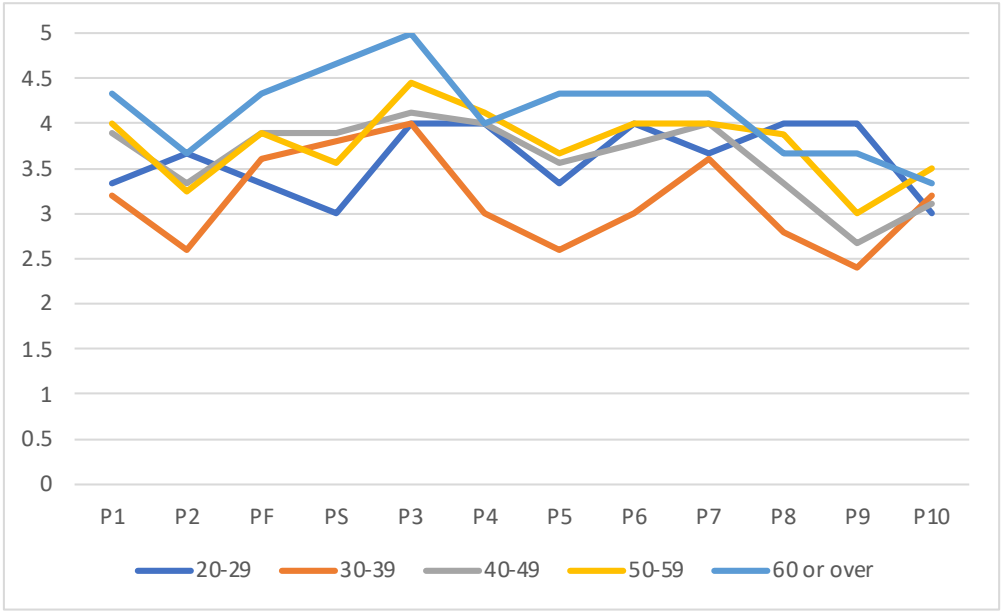
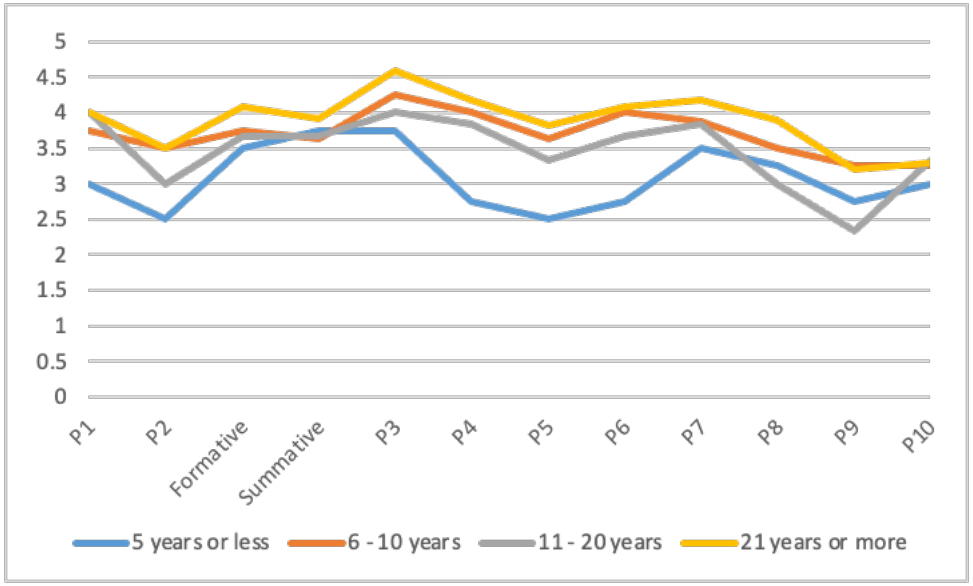


Figure 10 shows the average scores for the practice items and teaching experience. Two significant differences were found. The first difference was P4 (I use assessment results to make appropriate instructional decisions), where teachers with more experience had higher scores than those with less experience. The veteran teachers (21 years or more) had an average score of 4.18 while the new teachers (5 years or less) had an average score of 2.75. A similar pattern was found for P5 (I provide timely, descriptive, and actionable feedback to students), where veteran teachers had a score of 3.82 while the new teachers had a score of 2.5.

Figure 10

Experience and practice items



Feedback from Participants

In this revision, we had an open-ended item for suggestions or comments from the participants. There was one suggestion about providing concrete examples of formative and summative assessment:

> It would be easier for us to answer the questionnaire if the examples of formative assessment and summative assessment were provided. Also, it was a good chance for me to reflect my own practice since I could see the difference between the research-based testing and assessment by researchers and those in real classrooms.

This indicates that it was not very easy for JTEs to have a clear understanding of those assessments at their own schools. In fact, testing and assessment are not covered very well in many teacher training courses.

Testing and assessment topics receive little attention in the textbooks written for language teachers (Ito, et. al 2010), which reiterates the need for higher levels of language assessment literacy in Japan. This comment also indicates that participating in this study was a good chance for them to reflect on their own teaching practices in their classrooms. This further confirms the need for hands-on assessment training in teacher training courses along with improving the language assessment literacy of JTEs. Additional open-ended questions should be added in the next revision so that we can obtain more insights from our participants about language assessment literacy in Japan. Also, structured interviews would also help us better understand the participating JTEs.

Discussion

Beliefs

All the participants strongly believed that teachers understood tests and were able to use them (B1 average 4.76 / 5). However, there was a significant difference between teaching contexts. University teachers agreed most strongly with this statement (4.87) while those in the junior high school and other contexts had the lowest amount of agreement (4.0). One possible explanation is that tests, particularly summative tests such as midterm and final examinations, become more prevalent and of greater importance at higher levels of learning. Also, many university educators are solely responsible for all aspects of the four phases of assessment in their courses, so they know the difficulty and importance of testing. On the other end of the spectrum, teachers working in other contexts such as cram schools (*gakushu juku*) are primarily focused on preparing their students for school and university entrance examinations. Teachers at the secondary level need to adhere to the MEXT Course of Study and, if teaching in a public institution, the English Education Improvement Plan of their Boards of Education should be followed.

Based on feedback from the previous study, an item about the school tier (*hensachi*) was added. Before conducting the survey, our hypothesis was that stakeholders (i.e., students, legal guardians) at competitive institutions have higher language expectations. Instead of merely obtaining compulsory language credits to meet curriculum requirements, there is a perceived future need. In the case of students, they may be thinking about opportunities such as study abroad, scholarships, graduate school, and employment. Teachers at above-average institutions (50-59) had an average score of 3.75, showing relatively strong agreement. However, teachers at the highest tier (60 or higher) had the lowest score (2.8) while teachers at below-average institutions had the highest score of 4.43. This suggests that the relationship is more complicated than we anticipated. One possible reason for these results was the substantial number of

respondents (28% / 8) who did not report *hensachi* scores. Also, only a few secondary teachers responded to the questionnaire.

Finally, there is a strong connection between teaching experience and beliefs. All the veteran teachers (21 or more years of experience) strongly agreed that teachers should understand and be able to use tests (B1 average 5 / 5). They know the complexity and uses of effective assessment and they also strongly agreed on the importance of multiple assessments (B2 average 5 / 5). As teachers gain practical experience, they become more comfortable and capable of making various assessments such as performance-based in-class tasks, formative quizzes, and summative end-of-semester final exams. They also understand the need for multiple assessments, especially formative ones, to inform instructional decisions.

Practice

The two closely related factors of age and experience indicate that older and more experienced teachers provide superior feedback (timely, descriptive, actionable) more frequently than their younger and less experienced counterparts. As a teacher gains hands-on experience in the classroom, they are more able to identify a student's strengths and weaknesses and can usually give better feedback. Also, over time, a teacher often learns how to design better assessment tasks which result in better feedback.

Experienced teachers also use assessment results more than less experienced ones to make instructional decisions. For example, a formative assessment task such as a weekly quiz can provide a teacher with insights about the students' understanding of the current unit of instruction. If many students are struggling with a specific concept, further instruction or additional activities can be implemented to address the situation.

Conclusion

One objective of this pilot study was to further refine our research instrument before distributing the questionnaire to a larger group of teachers. The high levels of agreement for all the belief statements suggest that they are appropriate for JTEs. The new items, such as institutional tier and type, had unexpected results. The relationship between tier and real-life language was not as linear as we anticipated. Also, there was no difference between public and private institutions. However, these results might be quite different if many secondary school teachers respond to the questionnaire.

The second objective was to better understand the tendencies and characteristics of JTEs. Generally speaking, the teachers have strong beliefs about classroom-based assessment. There are factors, such as educational setting and experience, that have correlations with their beliefs. As previously cautioned, these results cannot be generalized because of the small sample size along with the majority of the respondents teaching at universities. The comment by one of the participants also shows that JTEs need higher levels of assessment literacy too.

This pilot study provided us with invaluable insights about how JTEs are implementing classroom-based assessment in their classrooms. The feedback from the respondents and the audience at our presentation will help us further refine the questionnaire. For future research, we will continue to fine-tune the questionnaire items and replicate this study on a large scale. We may also administer the questionnaire to pre-service teachers (i.e., students enrolled in teacher education courses) and at professional development workshops. Also, in addition to the questionnaire, interviews and observations would help us better understand the current classroom practices of JTEs and their assessment literacy needs. We hope our future research will provide actionable suggestions on how to help JTEs more effectively implement formative and classroom-based assessment.

Bio Data

Adam Murray has been an active member of JALT since 2003. He teaches at the University of the Ryukyus (Okinawa). His research interests are listening instruction, materials development, and assessment. <murray@gec.lab.u-ryukyu.ac.jp>

Taiko Tsuchihira is an associate professor at Seitoku University. Her main interest is language testing, especially testing listening and speaking. <ttaiko@wa.seitoku.ac.jp>

References

Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing, 25*(3), 327-347. https://doi.org/10.1177/0265532208090156

Davison C., & Leung C. (2019). Current issues in English language teacher-based assessment. *TESOL Quarterly, 43*(3), 393–415. https://doi: 10.1002/j.1545-7249.2009.tb00242.x

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly, 9*(2), 113-132. https://doi.org/10.1080/15434303.2011.642041

Goodman, R., & Oka, C. (2018). The invention, gaming, and persistence of the hensachi ('standardised rank score') in Japanese education, *Oxford Review of Education, 44*(5), 581-598. https://doi.org/10.1080/03054985.2018.1492375

Hill, K., & McNamara, T. (2012). Developing a comprehensive research framework for classroom-based assessment – an empirical study. *Language Testing, 29*(3), 395-420. https://doi.org/10.1177/0265532211428317

IGI Global. (n.d). What is Teaching-Learning Process. Retrieved from https://www.igi-global.com/dictionary/teaching-learning-process/48941

Ito, Y., Nakamura, Y., Kimura, K., Tsuchihira, T., Murray, A., Okada, A., & Matsumoto, K. (2010). An analysis of English teacher education textbooks published in Japan from a testing perspective, *JACET-Kanto Journal, 7*, 27-33.

MEXT. (2016). 教職課程コアカリキュラム [*Core curriculum for teacher education courses*]. Retrieved from https://www.mext.go.jp/b_menu/shingi/chousa/shotou/126/houkoku/1398442.htm

MEXT. (2022). 令和4年度英語教育改善プラン [*English education improvement plan for 2022*]. Retrieved from https://www.mext.go.jp/a_menu/kokusai/gaikokugo/1418086_00006.htm

Michigan Assessment Consortium. (2017). *Assessment literacy standards*. Mason, MI: Michigan Assessment Consortium.

Murray, A., Nakamura, Y., Matsumoto, K., Ito, Y. & Tsuchihira, T. (2020). Teachers' beliefs about classroom-based assessment. In P. Clements, A. Krause, & R. Gentry (Eds.), *Teacher efficacy, learner agency*. Tokyo: JALT. https://doi.org/10.37546/JALTPCP2019-62

Nishino, T. (2012). Modeling teacher beliefs and practices in context: A multimethods approach. *The Modern Language Journal, 96*(3), 380-399. https://doi.org/10.1111/j.15404781.2012.01364.x

Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing, 30*(3), 403-412. https://doi.org/10.1177/0265532213480338

Thompson, G., & Woodman, K. (2018). Exploring Japanese high school English teachers' foreign language teacher efficacy beliefs, *Asia-Pacific Journal of Teacher Education*. https://doi.org/10.1080/1359866X.2018.1498062

Wicking, P. (2017). The assessment beliefs and practices of English teachers in Japanese universities, *JLTA Journal, 20*, 76-89. https://doi.org/10.20622/jltajournal.20.0_76

Wyatt, M. (2018). Language teachers' self-efficacy beliefs: A review of the literature (2005-2016), *Australian Journal of Teacher Education, 43*(4), 92-120. https://doi.org/10.14221/ajte.2018v43n4.6

Yada, A., Tolvanen, A., Malinen, O., Imai-Matsumura, K., Shimada, H., Koike, R., & Savolainen, H. (2019). Teachers' self-efficacy and the sources of efficacy: A cross-cultural investigation in Japan and Finland. *Teaching and Teacher Education, 81*, 13-24. https://doi.org/10.1016/j.tate.2019.01.014

Appendix A

Questionnaire Items (Classroom-based Assessment Self-efficacy Scale)

Beliefs

- B1. Teachers should understand and be able to use tests.
  教師はストについて理解し、使えるべきである。
- B2. Multiple assessments provide a more balanced picture of a student.
  複数の評価をすることで、学生（生徒）の学力についてよりバランスのとれた把握ができる。
- B3. When done correctly, the resulting data can be used to make sound educational decisions.
  正しい評価をすれば、より確実な教育的決定をするためにその結果を使うことができる。
- B4. Quality assessments are a critical attribute of effective teaching and learning.
  質の高い評価は、効果的な指導・学習を行う上で重要な要素である。
- B5. Assessment results should be used to make instructional decisions to improve student learning.
  評価の結果は、学生（生徒）の学習を向上させるための指導上の決定をする際に使うべきである。
- B6. Clear learning targets are necessary for learning and assessment.
  学習と評価には明確な目標設定が必要である。
- B7. Effective feedback is necessary for learning.
  学習のためには、効果的なフィードバックが必須である。
- B8. Students should learn how to use assessment results to improve their learning.
  学生（生徒）は自身の学習の向上のためにどのように評価結果を使うかを学ぶべきである。
- B9. Good classroom assessment and quality instruction are intricately linked to each other.
  良質の教室内評価と質の高い指導は密接に結びついている。
- B10. Grading is an exercise in professional judgment, not just a numerical, mechanical exercise.
  評点を与えることは、単に数量的な操作ではなく、専門的判断をするということである。
- B11. Language tests should resemble real-life language use.
  言語テストは現実生活における言語使用に似たものであるべきである。

Efficacy

- E1. I can select and use various assessment methods appropriate to assessment purposes and learning targets.
  私は、評価目的と学習指標に対して適切な様々な評価方法の選択・使用ができる。
- E2. I can implement the 5-step process (plan, develop, review, field test, review & revise) for assessment development.
  私は評価方法を開発する際に、5段階のプロセス（計画、開発、チェック、実施、改良）を行うことできる。
- E3. I can use learning targets aligned to the standards to guide instruction.
  私は指導をする際に、客観的基準に基づく学習目標に沿って行うことができる。
- E4. I can use assessment results to make appropriate instructional decisions for individual students and groups of students.
  私は、個々の学生（生徒）と集団の両方に対して、評価結果を使った適切な指導上の決定をすることができる。
- E5. I can provide timely, descriptive, and actionable feedback to students based on assessment results.
  私は、評価結果に基づいて、記述によって行動指針を示すようなフィードバックを、適切なタイミングで与えることができる。
- E6. I can use grading practices that result in grades that are accurate, consistent, meaningful and supportive of learning.
  私は確かつ一貫性があり、学習に対して意義がありそれを助けるような評点の与え方ができる。
- E7. I can use assessment results appropriately to modify instruction to improve student achievement.
  私は、学生（生徒）の学習成果を向上させるように指導方法を修正するために、評価結果を適切に使うことができる。
- E8. I can use multiple sources of data over time to identify trends in learning.
  私は学習状況の傾向を認識するために、長期間にわたって様々な情報源を使用することができる。
- E9. I can communicate effectively with students, parents/guardians, other teachers, administrators and community stakeholders about student learning.
  私は、学生（生徒）の学習について、彼らだけでなく、親や保護者、他の教師たち、学校の管理職、そして地域の利害関係者と効果的にコミュニケーションができる。
- E10. I can make tests that resemble real-life use of English.
  私は現実生活における言語使用に似せたテストを作ることができる。

Practice

- P1. I use various assessment methods.
  私は様々な評価方法を使っている。
- P2. I use the 5-step process (plan, develop, review, field test, review & revise) for assessment development.
  私は価方法を開発する際に、5段階のプロセス（計画、開発、チェック、実施、改良）を行っている。
- I use formative assessment.
  私は形成的評価を使っている。
- I use summative assessment.
  私は総括的評価を使っている。
- P3. I use learning targets to guide instruction.
  私は学習目標に沿った指導を行っている。
- P4. I use assessment results to make appropriate instructional decisions.
  私は、評価結果を使って適切な指導上の決定をしている。
- P5. I provide timely, descriptive, and actionable feedback to students.
  私は、記述によって行動指針を示すようなフィードバックを、適切なタイミングで与えている。
- P6. I use grading practices that result in grades that are accurate, consistent, meaningful and supportive of learning.
  私は確かつ一貫性があり、学習に対して意義がありそれを助けるような評点の与え方をしている。
- P7. I use assessment results to improve instruction.
  私は、指導方法を改善するために評価結果を適切に使っている。
- P8. I use multiple sources of data over time to identify trends in learning.
  私は学習状況の傾向を認識するために、長期間にわたって様々な情報源を使用している。
- P9. I communicate effectively with students, parents/guardians, other teachers, administrators and community stakeholders about student learning.
  私は、学生（生徒）の学習について、彼らだけでなく、親や保護者、他の教師たち、学校の管理職、そして地域の利害関係者と効果的にコニュニケーションを行っている。
- P10. I make tests that resemble real-life language use.
  私は現実生活における言語使用に似せたテストを作っている。