



#JALT2019 • NAGOYA 11.1-11.4

Incorporating an External Online Test Into a University Language Program

Steven Pattison

Ritsumeikan Asia Pacific University

Anthony Diaz

Ritsumeikan Asia Pacific University

Tomoko Eto

Ritsumeikan Asia Pacific University

Patrick Johnston

Ritsumeikan Asia Pacific University

Kent Jones

Ritsumeikan Asia Pacific University

Malcolm Larking

Ritsumeikan Asia Pacific University

Michael Phillips

Ritsumeikan Asia Pacific University

Reference Data:

Pattison, S., Diaz, A., Eto, T., Johnston, P., Jones, K., Larking, M., & Phillips, M. (2020).

Incorporating an external online test into a university language program. In P. Clements, A. Krause, & R. Gentry (Eds.), *Teacher efficacy, learner agency*. Tokyo: <https://doi.org/10.37546/JALTPCP2019-61>

This paper presents a discussion of the adoption of an online test at the lowest levels of an English program at an international university in Japan. We first provide some background on the project, including an explanation of the test format, before describing its implementation at the university and the subsequent materials created to support student learning. We then discuss how we evaluated the impact and efficacy of the test and related materials using four different measures. An analysis and interpretation of these results follows, where we consider aspects such as washback, agency and efficacy. We conclude by stating that while these aspects are useful as guiding thematic concepts, further systematic research is needed to continue the project beyond the summary of the implementation and evaluation stages of the test provided here.

本稿では、日本の一国際大学の英語プログラム下位レベルコースへのオンライン・テストの導入について考察する。具体的には、まず本プロジェクトの背景やテスト形式を示し、学内でのテスト実施状況と学生サポートのために作成された補助教材について言及する。続いて、4つの異なる尺度から、テストと補助教材の有効性を評価した後、波及効果、学習者主体、自己効力感の観点から、評価結果の検証を行う。本稿の検証結果は、当該テストの実施案内としては有用であると言えるが、今後は、実施総括や評価という段階を超えたプロジェクトの継続のために、さらなる体系的なリサーチが必要であると結論づける。

In this paper, we discuss the motivation for and incorporation of a commercial online test – Pearson’s *Progress* test (hereafter referred to as *Progress*) – into the English language program at an international university in Japan. There were two perceived benefits of the test: it provides stakeholders with detailed feedback, and it encourages the development of language skills relevant to both general and academic English contexts. Following a period of investigation and trialling in 2016, *Progress* was then introduced into the lower two levels of the program in 2017.

Theoretical factors influencing this decision include the notions of self-efficacy (Bandura, 1997) and agency (Vitanova, Miller, Gao, & Deters, 2015) and we were also



#JALT2019 • NAGOYA 11.1-11.4

concerned with factors such as student achievement, motivation and autonomy (Spratt, Humphreys, & Chan, 2002), the impacts of success and failure, the achievements of peers, and positive feedback from teachers (Bandura, 1986). Margolis and McCabe (2006) also provide teaching strategies which can be applied to further enhance learners' self-efficacy, namely: assign moderately challenging tasks; employ peer models; train students in particular learning strategies; incorporate students' choices and interests into tasks; and provide students with frequent, focused feedback and encouragement. These concepts underpin the pedagogical practices adopted in tandem with *Progress* and also serve as criteria for evaluating them. Positive washback (McKinley & Thompson, 2018) was another important reason for adopting the test, particularly given its potential to reinforce those language skills that are central to the university's English program.

Despite these apparent theoretical benefits, there were also a number of practical challenges that arose with the implementation of the new test. These initially included increasing teacher and student familiarity with *Progress*, as well as helping teachers understand and effectively utilise the complex feedback provided to students by the test. In response, a range of materials and instructional approaches were developed to support students' in-class development of the language skills expected by test items. A student needs survey was also developed, as were extra-curricular support through a series of workshops and additional self-study materials. Finally, various measures were used to evaluate *Progress*: a review of previous *Progress* data; a student survey; visits to other universities using the test; and the compilation of can-do data from students who had taken the test.

Background

Progress is an online English proficiency test developed by Pearson Education (2015). It is an integrated test, assessing the four main language skills as well as two *enabling skills*: vocabulary and grammar. The test provides students with detailed feedback in the form of an overall Global Scale of English (GSE) score from which equivalency to CEFR bands, and other well-known tests such as IELTS and TOEFL iBT, can be derived, while also providing GSE scores for individual skills such as reading or grammar (Appendix A).

Progress is divided by GSE ranges into six different tests, all of which share a common scale (GSE 10-90), so it is possible to select an appropriate version of the test for different levels of students. In addition to a GSE score for each of the main language and enabling skills, test-takers receive detailed feedback in the form of can-do statements. While these statements usually describe a language learner's proficiency level in terms of what they are currently able to do in the language, in *Progress* they indicate those skills that the learner has not yet mastered [or acquired].

At our university, the main objective of the English program, after completion of the pre-intermediate level, is to develop communicative and academic language skills in preparation for mandatory English-medium lectures. The language skills assessed by *Progress* were judged to be in alignment with these objectives, and as a result, the use of *Progress* began in 2017. The university's elementary English course adopted the GSE 15-30 test whereas pre-intermediate adopted the 25-40 one, with an expected target score of 35 (CEFR A2) in elementary and 39 (CEFR A2+) in pre-intermediate (note that the test is able to provide a score up to 35 in the 15-30 test and up to 45 in the 25-40 test). At the end of the semester, after taking multiple tests, the students' best score from *Progress* then accounts for up to 25% of their total grade.

As the test was being considered, consultations with the developers were also carried out to further understand the test and how different question types were designed to evaluate students' English proficiency. This was done in order to refine the guidelines and materials that were developed to help students become familiar with the test format and to foster the relevant language skills.

Implementation

Three main student needs were recognised by teachers shortly after the implementation of the test. First, students were expected to improve their general English proficiency in order to be able to improve their *Progress* scores, even though they did not necessarily understand the expectation or the test itself. Further, because the test is online, they needed a similar online platform to practice and reinforce these skills while studying English. Finally, students were also expected to become learners who were independently able to plan, monitor, and self-regulate their own language study regime.

One further concern in choosing *Progress* was that, unlike older tests such as IELTS and TOEIC, it is not supported by currently published practice materials. Hence, beyond familiarising students with the format of the test and providing them with advice about how to deal with the various question types, teachers were faced with the additional challenge of having to develop a range of materials to compliment the skills assessed in the test (such as describing, paraphrasing and dictation). It was expected that these materials, once developed, would help students to incrementally improve both their generic English language proficiency and specific *Progress* competence. The materials, as described below, also needed to be accessible to students in both physical and online formats and for in-class and independent study.



#JALT2019 • NAGOYA 11.1-11.4

Development of Progress Teaching Resources

In response to these challenges, four types of supplementary teaching resources were developed.

Practice Websites

These were designed to link the test questions with overall proficiency development. In order to design these sites, *Progress* questions were categorised into eight types by language skill, including reading, writing (dictation, rewriting-a-passage, timed-writing), speaking (describing-a-picture, repeating-a-sentence, retelling-a-story), and enabling skills (vocabulary and grammar). Although some question types required multiple skills, such as rewriting-a-passage questions, this categorisation was made for the sake of simplicity.

We then designed exercises based on these categories, such as timed-writing, while matching them with existing coursebook topics in the two levels of the English program. For example, students wrote short essays of 50 to 75 words in 10 minutes, as per the *Progress* timed-writing format, about topics such as business, festivals and sports. The aim of this approach was for students to consolidate their understanding of the language learned in the coursebooks while also studying for *Progress*.

To share these exercises, two websites were designed, one for each of the two English courses. Each site included questions on a target skill, a countdown timer, and links to helpful external resources, such as voice recording, writing, and dictation websites. These websites were then shared via the course's learning management system (LMS), and became available for instructor's and student's to use in class and independently.

Classroom Workshops

A series of *Progress* skill PowerPoint workshops were also designed for use by teachers in class. Each workshop targeted a specific question type, such as describing-a-picture, with students first receiving instruction on the target question type followed by some strategies they could use to answer it. They then practised by answering a few sample questions, were divided into pairs or groups for practice and mutual feedback, and were finally given advice on how to continue practising by themselves.

Extracurricular Workshops

Three 90-minute workshops were held outside of class-hours primarily for students who were struggling with specific skills, such as describing, oral summaries, and written

paraphrases. As with the classroom workshops, these introduced materials that students could access for independent language practice while also aiming to provide instruction on how to use those materials effectively (Appendix B).

Test Review and Study Guide

As a tool for students to better monitor and plan their studies, a *Progress* review and study guide was created. Students were asked to use this guide to compile precise information and to create a language profile, including the can-do statements identified by *Progress* (the skills they needed to work on), their overall GSE scores, and their test scores in each skill area. After analysing this data, students were tasked to assume responsibility for self-studying by setting target GSE scores and planning tasks focusing on specific can-do statements in their guide. In addition, the guide was also expected to help teachers to understand their students' progress and give focused advice on specific skill areas.

Evaluation

Once the test was in use, efforts were then made to assess the usefulness of the test across four discrete measures.

Progress Data

GSE score data for the overall test and individual skills from *Progress* were collected and analysed to determine students' strengths and weaknesses in regard to the language skills needed for success on the test. Data were collected in both 2017 and 2018 and the 2018 results (Appendix C) represent the scores by skill achieved by students for all test administrations in spring and fall 2018.

The data show that, according to the test, speaking was the strongest of the four main skills in the pre-intermediate class. In the elementary class, speaking was also the strongest of the four principle language skills in spring 2018 when the majority of students in the level were new to the university. In fall 2018, the scores for each skill were lower than for spring 2018. This was probably due to the smaller number of test takers (N=39) and the fact that the majority of these students were repeating the course. Reading was the weakest skill in both the elementary and pre-intermediate levels. In pre-intermediate, the average reading score was higher than the average in elementary for both spring and fall 2018, increasing by 5.7 points from spring 2018. Moreover, the average scores for all of the skills were higher in pre-intermediate than in elementary.

Nevertheless, since approximately 180 students moved from elementary to pre-intermediate between the spring and fall semesters, it can be suggested that elementary students were still performing satisfactorily. It is probable that familiarity with the test was a factor in this, but so too was the probable development of language skills (such as paraphrasing, pronunciation, taking dictation and description). Overall, since the majority of students in pre-intermediate in fall 2018 were elementary students in the preceding spring semester, the test scores suggest that students were developing overall competence by either making improvements in their language and enabling skills, becoming more familiar with the test, enhancing their question-answering techniques, or possibly a combination of all three.

Student Survey

At the end of the Spring 2018 semester, an online survey was administered to 364 students about their impressions of *Progress*, the practice materials, and their preparedness for the test. The survey included 21 questions about what skills and questions the students found challenging, their study methods, and their self-rating of their ability to answer various question types (Appendix D). Of the respondents, 46.8% were in the elementary level of the program and 53.2% were in the pre-intermediate level, with all students in either the International Business (APM) or Asia Pacific Studies (APS) colleges.

Students' Perceptions of the Test and Test Preparation

The results of the survey yielded a number of important findings about how the students prepared, or failed to prepare, for the test. The survey showed that the students needed to increase the amount of time they self-study and how often they make use of the practice materials. Students reported that they used the practice materials "quite rarely," with more than half of the respondents claiming that they only used them between 1 and 5 times.

The survey also asked which question types were the most challenging by skill area and responses indicated that the questions that required more productive language skills were the most challenging. From the students' perspective, speaking was most problematic, though there seemed to be a general belief that they needed to develop their proficiency in all skill areas. Responses to questions about the most difficult grammar and vocabulary questions showed that fix-the-mistake items were perceived as the most difficult. For the speaking questions, the respondents thought the retelling-a-story item was the most difficult and for writing, the type-what-you-hear item.

Students' Perceptions of the Study Materials

In terms of how easy the students felt the online practice materials were to use, the main response was neutral, though some respondents felt the ease of use could be improved. Nevertheless, students indicated that they found the online materials more useful than other study options such as grammar books and online quizzes. When students were asked about whether they would consider attending workshops that showed them more effective ways of studying for the test, 64% answered that they would. This apparent desire for extra practice was a positive result and suggests that the students did indeed desire more practice but did not yet have the necessary know-how or initiative to study well by themselves.

For the practice materials, the majority of respondents thought that the picture description tasks were the most useful. They also reported that the reading, sentence-dictation, and the repeating-a-sentence activities were fairly useful. In contrast, the retelling-a-story, grammar-and-vocabulary, and short-essay activities were considered less useful. Students thought the retelling-a-story and rewriting-a-passage questions were quite challenging to study by themselves, and they wanted more practice with retelling-a-story, sentence-dictation, and repeating-a-sentence questions. In contrast, reading questions were considered to be the easiest. As a result of the data, understanding which specific questions were difficult for the students to master allowed the teachers to create more targeted practice materials.

University Visits

In 2018, as part of this research project, two universities that currently use *Progress* were visited in order to comparatively evaluate their approaches. The aim of the talks was to share common experiences of administering the test (i.e. with practice and study materials, test preparation, test results, teacher attitudes, student feedback surveys and the test's overall impact on their English program's curriculum).

At the first university, whose experiences proved most similar to ours, analysis of student performance over a 1-year period from 2015-2016 indicated a positive trend. In particular, the students' speaking scores improved significantly, with an 8-point average increase. The second most improved area was vocabulary (7 points), followed by grammar (6 points). Some successful students were actually able to improve their CEFR score from A1 to A2+ with significant GSE increases reported for speaking (13 points), grammar (18 points), and overall score (13 points) between the start- and end-of-course tests.

A survey conducted by the same university to determine how the test influenced their students' study methods and motivation showed 69% of their students agreed that the tests were useful in helping them understand their strengths and which skills needed more practice. Forty-six per cent of their respondents agreed that they had found the test score data helpful in determining how to study. This indicates that the test diagnostics were being used by their students to help streamline their study methods and attend to their specific needs. The survey also showed that their data on students' improvements between tests increased motivation to study, with 63% reporting improved motivation.

Follow-up interviews conducted with teachers and students showed that students appreciated receiving detailed feedback about their weaknesses from the test, allowing them to adapt their study methods accordingly. Teachers also reported that they had adapted their teaching methods to address the skill areas needing the most attention, for example by introducing more speaking activities in class. Their students responded positively to the increased amount of speaking and listening practice and the teachers noted that their motivation to speak in class had improved significantly.

Overall, when compared with our test and survey data, the meetings with these universities elucidated several common challenges, but it was evident that *Progress* had had a positive impact on their curricula by enhancing productive skills training, student performance, and student motivation to study. The professors at the other institutions observed several potential areas of improvement while preparing their students for *Progress*. For example, they acknowledged the need to create more targeted training resources to increase students' self-study time for the test and also that increasing motivation for low-achieving students was an ongoing challenge, as was the issue of how to increase the students' general IT skills.

The Collection of Can-Do Statements

After a test has been administered, *Progress* generates a list of CEFR-based can-do statements for individual students. These are the GSE learning objectives that, according to the test, they have failed to achieve. The rationale for the analysis of this data was to utilise it as a basis for the creation of materials which would help students by identifying and addressing their areas of greatest need and assist them with acquiring relevant language skills (as explained earlier). Furthermore, since there is an overlap in the language skills developed in courses at each level of our English program (such as note-taking, summarising, paraphrasing, and the production of well-structured written texts) this data provides potential insights into which aspects of the English curricula presented the most difficulty for our students.

Data, collected in 2018 from 200 students across the elementary and pre-intermediate levels, yielded 308 distinct can-do statements which occurred with varying frequency. The most commonly occurring statement data (Appendix E) suggests that more listening practice material is required. For example, designing listening materials that focus on the various ways cause-effect relationships are expressed in spoken English could assist students in recognising the connections among different elements of speech. A further implication is that students are not competent in listening to informal conversations by native speakers at a natural speed. Thus, exposure to listening materials recorded at natural speeds would likely be beneficial to them.

Discussion

In this paper, we have described the implementation of an online test in the English language curriculum of an international university in Japan. The overarching concern has been to identify the apparent advantages of incorporating *Progress* into our program, primarily in regard to the promotion of effective language learning practices from the standpoint of both the instructor and learner.

This adoption was also accompanied by the introduction of materials and approaches aimed to provide instruction to students on discrete language skills the test assesses, including describing, paraphrasing, and the production of short well-structured written responses. This then allowed for the integration of new materials within the existing content of the target English courses. This had the dual advantage of providing reinforcement of the current language skills and content being learnt and extending the use of existing materials by focusing on the language skills also assessed by *Progress*. Based primarily on the analysis of test scores, accompanying feedback, and responses from student surveys, we then considered what could be improved, where stakeholders' energies should be focused, how to evaluate what appears to have been effective, and what was learnt from the parallel experiences of other institutions.

From this work, we believe that *Progress* teaching resources can work to raise aspects of language learning, such as students' agency and self-efficacy. However, in order to make the teaching resources function more effectively, further development is needed, especially in regards to learnability, feedback and user rate. For example, we understand that students' experience of learning on the *Progress* websites can be improved. Currently, the websites do not provide students with correct or model answers, so it is difficult for them to understand how well they have completed the exercises. If criteria and model answers for exercises, including number of words or sentences and target grammar and

vocabulary, are shared, students themselves can more easily distinguish what they can and cannot do and attempt to set their own focused learning goals as a result.

Additionally, although the *Progress* review and study guide presents detailed information on student learning, it has not been utilised systematically by teachers when providing regular guidance and feedback to students. Thus, establishing a system to incorporate teachers' feedback into coursework is needed. A first step towards this would be for teachers to collect the study guide at regular intervals, review it, and return it to students with some explanatory feedback.

Also, when looking at these *Progress* teaching resources from the perspective of learners' self-efficacy, it turns out that certain features of these resources can also help teachers develop students' language proficiency while also implementing Margolis and McCabe's (2006) general strategies for raising learners' self-efficacy (Table 1).

Table 1. Teaching Mediums, Features, and General Strategies

Teaching medium	Feature	General strategy
<i>Progress</i> practice websites	1. Students are able to select the question types, topics and task difficulties of the exercises that they would practice	1. Assign moderately challenging tasks 2. Incorporate students' choices and interests into tasks
Classroom workshops	1. Specific strategy training is given 2. Peer activities for practice and advising are included	1. Train particular learning strategies 2. Employ peer models
Extracurricular Workshops	1. Advice on how to access online materials for self-study is given 2. Training on how to use the materials effectively	1. Assign moderately challenging tasks 2. Train particular learning strategies
<i>Progress</i> Review and Study Guide	1. Detailed information about students' progress is shared with teachers	1. Provide students with frequent, focused feedback and encouragement

Note. General strategies are taken from Margolis and McCabe (2006).

Next, in terms of evaluation of *Progress* at the university, the results of the student survey (Appendix D) revealed that the websites were rarely used by students. Consequently, further activities linked with these websites should be introduced in classroom workshops. Training in how to access online resources and employ effective self-study practices also needs to be made available outside of regular class time, especially to students who are struggling with doing this themselves.

In addition, the assessment of student needs highlights the importance of creating materials understandable by the end-user, so that learners can both navigate and make effective use of them. The in-class and supplementary workshops were designed to provide students with compatible methods to practise various language skills and to enhance agency, but their effectiveness is somewhat contingent on learners' ability and willingness to access them. Further, based on student responses, the current organisation and delivery mode of the materials is also inhibiting the practice of language skills. Indeed, this is of concern given that the goal of introducing independent learning strategies to the target learners was based on agency (Gao 2010; Margolis & McCabe, 2006).

Also pertinent to the goal of increasing learner agency is the ability of students to navigate the *Progress* site in order to comprehend and use the can-do-statement feedback. However, the volume and complexity of the feedback restricts learners' ability to process and apply it to their individual language learning context. The independent learning guide moderates this process to some extent, but more can be done to ensure that it is introduced in a systematic way in all classes and that teachers make use of the students' learner profiles to offer better support and encouragement.

According to Bandura (1997), self-efficacy is the foundation upon which agency is developed. As observed above, the test can have a detrimental effect on both of these, not only when an apparently disappointing score is received, but also when an unmanageable amount of feedback is generated. Having students filter this feedback to help them tailor their learning to their own needs is an important step in reducing the harmful effects of negative washback, where students simply practise in order to master the test without acquiring the skills targeted by the test. As was noted in the survey results, students' perceptions of their own weakest skills are not borne out by the test results, so reference to the test data is important for raising the students' awareness of their strengths. However, washback from the test can also have a positive effect on students in terms of self-efficacy since it provides confirmation of the language skills in which they are most proficient. Moreover, by identifying areas of relative weakness in a granular form, the test provides specific guidance to learners on which language skills require greatest attention, thus linking test preparation more directly with general proficiency.



#JALT2019 • NAGOYA 11.1-11.4

Pattison, Diaz, Eto, Johnston, Jones, Larking, & Phillips: *Incorporating an External Online Test Into a University Language Program*

Arguably the greatest benefit of implementing *Progress* in the English program is that it has the potential to raise individual learners' awareness of their language proficiency and offers them precise feedback on where to direct their learning. Furthermore, it encourages the development of language skills that are useful not only in the context of the test, but more generally in their academic and social contexts. It is still too early to gauge whether the implementation of *Progress* has had a positive effect on students' self-efficacy as language learners, but many of the actions taken by instructors were motivated by consideration for increasing student agency, itself an important factor for increasing self-efficacy.

Conclusion

The opportunity to focus on different aspects of using and understanding *Progress* through this study has benefited our English program in terms of evaluating the development of materials and methods for practising relevant language skills. Although much of this work remains ongoing, we believe that we have taken sizeable strides towards understanding the needs of students, developing materials and ideas to help students both inside and outside the classroom, understanding the results of tests and where we should focus our energies, and evaluating our performance relative to other institutional users.

That said, it is probably too early to make sure-footed claims about the effect of the implementation of the test on students in their language learning and general proficiency. Although the university provides students with support systems to familiarise them with the test and develop their language skills in a focused way, more can be done to improve methods for helping students to make sense of the test feedback and to use it more effectively. Moreover, students have expressed a need for better access to and clearer instruction on how to independently use materials for practising the language skills evaluated by the test. In addition, beyond the feedback provided by the test, more can be done to provide students with comprehensible, actionable feedback from practice activities, which could also positively influence learner agency and self-efficacy. While these concepts serve as useful criteria for guiding and evaluating the development of teaching materials, methods of learning, and use of feedback, clearly more work is needed to make systematic and rigorous use of them in evaluating our pedagogical practices and their outcomes.

To more confidently determine what the effects on agency and self-efficacy of the test are, a longitudinal study following learners through the entirety of their English language studies at the university should be conducted. Such a study could reveal, for example,

variation in the factors that influence self-efficacy at different levels of the test. The results of such a study could also help us to further refine the methods of instruction, practice materials, and use of feedback to more effectively address the students' needs with particular regards to learner agency and self-efficacy.

Bio Data

Steven Pattison is an associate professor in the Center for Language Education, Ritsumeikan Asia Pacific University in Oita, Japan. His research and teaching interests include L2 reading; pragmatics, particularly Gricean and neo-Gricean theory; and stylistics. He is also interested in the intersection between stylistics and cultural analysis of literary texts as a medium for language learning and teaching. <steven@apu.ac.jp>

Anthony Diaz is an English lecturer at the Center for Language Education, Ritsumeikan Asia Pacific University in Beppu, Oita Prefecture. He has taught a variety of students including academic, vocational, and adult learners in the United States and Japan. His research focuses on pronunciation, speech perception, and corpus linguistics. <antdiaz@ucdavis.edu>

Tomoko Eto is a tenured senior English lecturer at the Center for Language Education, Ritsumeikan Asia Pacific University, in Oita, Japan. Her research interests include language learning strategies, learner beliefs, and needs analysis. <etomoko@apu.ac.jp>

Patrick Johnston is an English lecturer at the Center for Language Education, Ritsumeikan Asia Pacific University in Oita, Japan. He has taught English in Japan for over 12 years and his research interests are primarily related to computer-based testing. <johnston@apu.ac.jp>

Kent Jones is a tenured senior lecturer at Ritsumeikan Asia Pacific University. He has taught ESL at universities and colleges in Canada as well as EFL in universities in Japan. His research is largely related to writing, particularly feedback. <joneske@apu.ac.jp>

Malcolm Larking is a tenured senior English lecturer at the Center for Language Education, Ritsumeikan Asia Pacific University, in Oita, Japan. His research interests include sociocultural theory, conversation analysis, and critical reading strategies. <larkingm@apu.ac.jp>

Michael Phillips is an English lecturer at the Center for Language Education, Ritsumeikan Asia Pacific University in Oita, Japan. He is the current president of the JALT Oita Chapter and publications officer of the JALT BizCom SIG. His research

interests include presentation skills, intercultural business competence, and corpus linguistics. <michaelp@apu.ac.jp>

References

- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. New Jersey: Prentice-Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman and Company.
- Gao, X. (2010). *Strategic language learning: The roles of agency and context*. Bristol, UK: Multilingual Matters. <https://doi.org/10.21832/9781847692450>
- Margolis, H. & McCabe, P. P. (2006). Improving self-efficacy and motivation: What to Do, What to Say. *Intervention in School and Clinic*, 41(4), 218-227. <https://doi.org/10.1177/10534512060410040401>
- McKinley, J. & Thompson, G. (2018). Washback effect in teaching English as an international language. In J. Lontas, M. DeliCarpini & S. Abrar-ul-Hassan (eds) *TESOL Encyclopedia of English Language Teaching* (pp. 1-12). Hoboken, NJ: Wiley-Blackwell. <https://doi.org/10.1002/9781118784235.eelt0656>
- Pearson. (2015). *Progress: Test information booklet*. Retrieved on January 2, 2020, from: <https://www.pearson.com/content/dam/one-dot-com/one-dot-com/english/SupportingDocs/Progress-test-information.pdf>
- Spratt, M., Humphreys, G., & Chan, V. (2002). Autonomy and motivation: Which comes first? *Language Teaching Research*, 6, 245-266. <https://doi.org/10.1191/1362168802lr1060a>
- Vitanova, G., Miller, E. R., Gao, X., Deters, P. (2015). Introduction to theorizing and analyzing agency in second language learning: Interdisciplinary approaches. In Deters, P., Gao, X., Miller, E. R., and Vitanova, G. (Eds.). *Theorizing and analyzing agency in second language learning: Interdisciplinary approaches*. Bristol, UK: Multilingual Matters.

Appendix A The Global Scale of English

Progress provides scores using its own scale, called the Global Scale of English (GSE), which works alongside CEFR. The GSE is a numerical scale, aligned to a series of granular descriptors which provide detailed descriptions of performance within a particular CEFR band. The GSE ranges for each CEFR band covered by *Progress* taken by students in elementary and pre-intermediate English are shown below:

CEFR LEVEL	GSE Point Ranges
<A1	10-21
A1	22-30
A2	31-35
A2+	36-42
B1	43-50

Appendix B Workshop Examples

Workshop 1 - Grammar and Vocabulary - Picture-Description Task Progression

Workshop 2 - Listening and Speaking - Retelling-a-Story Task Progression

Workshop 3 - Reading and Writing - Rewriting-a-Story Task Progression

Appendix C

Can-do Statement Data

Note that scores marked as <10 fall below the assessed range and that scores of >45 fall above it, and that scores of <10 were treated as zero while scores of >45 were capped at 45.

Table 1. Elementary GSE Scores by Skill (Spring, 2018)

Skill	Mean	S.D.	CEFR	Lowest	CEFR	Highest	CEFR
Reading	20.65	6.5	<A1	<10	<A1	35	A2
Listening	27.8	6.83	A1	<10	<A1	35	A2
Speaking	29.7	7.1	A1	<10	<A1	35	A2
Writing	21.02	6.7	<A1	<10	<A1	35	A2
Vocabulary	25.3	7.33	A1	<10	<A1	35	A2
Grammar	23.7	7.18	A1	<10	<A1	35	A2

Table 2. Elementary GSE Scores by Skill (Fall, 2018)

Skill	Mean	S.D.	CEFR	Lowest	CEFR	Highest	CEFR
Reading	15.8	10.2	<A1	<10	<A1	35	A2
Listening	18.5	13.8	A1	<10	<A1	35	A2
Speaking	17.1	13.91	A1	<10	<A1	35	A2
Writing	16.92	9.24	<A1	<10	<A1	34	A2
Vocabulary	18.69	10.61	A1	<10	<A1	35	A2
Grammar	19	10.32	A1	<10	<A1	34	A2

Table 3. Pre-Intermediate GSE Scores by Skill (Spring, 2018)

Skill	Mean	S.D.	CEFR	Lowest	CEFR	Highest	CEFR
Reading	22.7	13.01	<A1	<20	<A1	44	B1
Listening	33	7.32	A2	<20	<A1	45>	B1
Speaking	38	6.8	A2	<20	<A1	45>	B1
Writing	28	9.2	A1	<20	<A1	45>	B1
Vocabulary	32	9.84	A1	<20	<A1	45>	B1
Grammar	31.5	11.14	A1	<20	<A1	45>	B1

Table 4. Pre-Intermediate GSE Scores by Skill (Fall, 2018)

Skill	Mean	S.D.	CEFR	Lowest	CEFR	Highest	CEFR
Reading	28.4	13.44	A1	<20	<A1	45>	B1
Listening	29.7	8.32	A1	<20	<A1	45>	B1
Speaking	37	9.79	A2	<20	<A1	45>	B1
Writing	28.6	9.1	A1	<20	<A1	45>	B1
Vocabulary	35.1	10.24	A2	<20	<A1	45>	B1
Grammar	34.7	11.7	A2	<20	<A1	45>	B1

Appendix D Progress Survey Results

1. Which English course are you taking?

Elementary	46.8%
Pre-Intermediate	53.2%

2. Which *Progress* skill do you need the most practice at?

Listening	20.7%
Reading	12.7%
Writing	10.5%
Speaking	28.7%
Grammar and vocabulary	27.5%

3. Which grammar and vocabulary questions are the most difficult for you?

Reorder the sentence	17.8%
Fill in the gaps	36.7%
Fix the mistakes	45.6%

4. Which speaking questions are the most difficult for you?

Describe a picture	4.4%
Repeat the sentence	36.1%
Retell the story	59.4%

5. Which writing questions are the most difficult for you?

Type the sentence	49.6%
Rewrite the passage	39.3%
Short essay	11.1%

6. How often have you used the online-practice site (Google site) for your study (both inside and outside the classroom)?

Never	13.6%
About 1 - 5 times	54.9%
About 6 - 10 times	18.9%
About 11 - 20 times	8.4%
More than 21 times	4.2%

7. Which online-practice questions were the most useful?

Reading questions	30.4%
Describe a picture	42.2%
Repeat the sentence	24.7%
Retell a story	20.8%
Sentence dictation	26.8%
Grammar and vocabulary	17.2%
Rewrite the passage	12.3%
Short essay	16.6%

8. How easy to use was the online-practice site?

Very easy	10.1%
Easy	25.1%
Neutral	43.8%
Difficult	18%
Very difficult	3%

9. What practice materials were helpful for your self-study?

Online practice materials	47.5%
Online workbook	37.5%
Teacher made materials	43.7%

10. What other ways of practicing for *Progress* did you use?

Various answers, including-

- ESL Fast
- Self Access Learning Center
- Word Engine
- Grammar book
- Quizlet

11. How often did you spend practicing for *Progress* outside of class?

Never	24.4%
Rarely	29.4%
Sometimes	30.5%
Often	10.6%
Very often	5%

12. How well do you understand the test instructions?

Not very well	4.2%
Not well	8.4%
Neutral	30.6%
Well	34.8%
Very well	22%

13. What are some technical problems you have faced when taking the test?

Recording your voice	19.2%
Freezing screen	18.1%
Microphone not working	17.6%
Skipping a question	58.8%

14. Which response is closest to yours when you answer the describe-a-picture questions?

I do not know what to describe	7.6%
I know what to describe but cannot make sentences	38.9%
I can make sentences but cannot speak fluently	36.4%
I can fluently describe the picture	17.1%

15. Which response is closest to yours when you answer retell a story questions?

I cannot catch what I heard	24%
I can catch what I heard but cannot identify which point I have to talk about	32.6%
I can catch what I heard and identify which point I have to talk about but cannot make sentences	20.9%
I can catch what I heard, identify which point I have to talk about and make sentences, but cannot speak fluently	15.9%
I am confident in both my listening and speaking	6.7%

16. Which response is closest to yours when you answer rewrite-a-passage questions?

I cannot understand the passage	14.5%
I can understand the passage, but do not know what to write about	32.1%
I can understand the passage, know what to write about but cannot write sentences	43%
I am confident in my reading and writing ability	10.3%

17. Which response is closest to yours when you answer a short-essay question?

I cannot understand the writing prompt	8.1%
I can understand the writing prompt but do not know what to write about	20.8%
I can understand the writing prompt, have some ideas to write about but cannot make sentences	33.1%
The time for writing is too short	23.6%
I am confident with my writing ability	14.3%

18. Do you feel your *Progress* score accurately reflects your ability level?

No, not at all	17.7%
No, not really	16.9%
Neutral	41.9%
Yes, somewhat	17.4%
Yes, definitely	6.2%

19. Which question types are the easiest for you to practice by yourself?

Reading questions	46.8%
Describe a picture	33.8%
Repeat a sentence	32.4%
Retell a story	18.6%
Sentence dictation	35.2%
Rewrite a passage	22%
Short essay	33.8%

20. Which question types do you want more help with in your English class?

Reading questions	24.4%
Describe a picture	15.8%
Repeat a sentence	33.5%
Retell a story	37.2%

Sentence dictation	34.4%
Rewrite a passage	22.3%
Short essay	18.6%

21. If workshops were offered which showed you effective ways to practice for the test, would you attend them?

Yes	64%
No	36%

Appendix E Top Ten Can-do Statements Elementary

Rank	Frequency	Can-do Statement
1	126	Can understand cause and effect relationships in informal conversation at a natural speed.
2	126	Can converse naturally, fluently and effectively.
3	110	Can follow much of everyday conversation if speakers avoid very idiomatic usage.
4	101	Can follow speech which is very slow and carefully articulated, with long pauses.
5	101	Can follow an animated conversation between two native speakers.
6	98	Can understand main points of standard speech on familiar topics (e.g. work, leisure).
7	97	Can produce simple, mainly isolated phrases about people and places.
8	92	Can understand the majority of broadcast material on familiar topics in clear standard speech.
9	86	Can extract relevant details in everyday letters, brochures and short official documents.
10	83	Can understand short written notices, signs and instructions with visual support.



#JALT2019 • NAGOYA 11.1-11.4

Pattison, Diaz, Eto, Johnston, Jones, Larking, & Phillips: *Incorporating an External Online Test Into a University Language Program*

Pre-Intermediate

Rank	Frequency	Can-do Statement
1	191	Can understand scripted speech delivered quickly, if the accent is familiar.
2	159	Can follow most of an everyday conversation if speakers avoid very idiomatic usage.
3	158	Can develop a clear written description or narrative with relevant supporting detail and examples.
4	156	Can understand cause and effect relationships in informal conversation at natural speed.
5	140	Can follow most of a clearly structured presentation within their own field.
6	130	Can paraphrase in simpler terms what someone else has said.
7	127	Can quickly scan long, complex texts for key information.
8	118	Can understand the key points about a radio programme on a familiar topic.
9	117	Can follow familiar topics if the speaker is clear and avoids idiomatic usage.
10	116	Can understand the speaker's point of view on most topics delivered at natural speed and in standard language.