Establishing Oral Reading Fluency Norms for Japanese English Language Learners

by
Elton LaClare*

Abstract

The present study aims to establish Oral Reading Fluency (ORF) norms for various CEFR proficiency levels as reported by the Oxford Online Placement Test (OOPT). The participant population for the study consisted of a sample of incoming SILC English Communication 1 students whose OOPT results indicated that their English abilities fell within the A1 or A2 bands of the CEFR scale. To further refine fluency norms, the A2 band was subdivided into *low* and *high* according to raw scores. The researcher administered the ORF measure and recorded both the reading rate, Words Read Correctly per Minute (WCPM), and accuracy, WCPM divided by the total number of words read, of each participant. Scores were grouped according to CEFR proficiency level then analyzed to determine the mean speed and accuracy scores as well as the as the standard deviation for each level. Although reading rate and accuracy means were calculated for the A1, A2 (low), and A2 (high) levels, only the A2 (low) level contained a sufficiently large sample size (n = 43) to enable the establishment of an ORF norm. Data from the remaining groups will be retained and augmented with further ORF testing of future intakes of SILC English Communication 1 students.

Key Words：　Oral Reading Fluency,　L2 reading,　curriculum-based measurement

1. Introduction

The Oral Reading Fluency (hereafter ORF) measure is an example of Curriculum-Based Measurement (CBM), which has been used extensively in educational environments throughout North America. The purpose of CBM is threefold: to monitor the progress of students, to assess the effectiveness of classroom interventions, and to identify students who require special attention (Fuchs, L. S. & Fuchs, D., 2007). The history of CBM spans several decades, and extensive norming has been accomplished. This norming has resulted in the establishment of performance benchmarks in core skills that assist in individualized decision making with respect to academic skill development (Deno, Fuchs, Marston & Shin, 2001). The principal benefits of interpreting progress-monitoring measures such as the ORF in the context of established performance

＊Lecturer, Sojo International Learning Center

norms are:

1. It is possible to estimate rates of improvement.
2. Students not demonstrating adequate progress can be identified and offered alternative forms of instruction.
3. Different forms of instruction can be compared to determine their efficacy.

Rather than functioning as a part of the assessment framework, the ORF may be used to inform pedagogical decisions with regard to the development of reading skills (Fuchs, L. S. & Fuchs, D., 1996; Gersten & Dimino, 2001). Unlike traditional assessments, which are often conducted at the end of a term of study, the ORF is administered at the outset, enabling teachers to identify low achievers and enact measures to address the performance gap concurrent with the course of study.

As it is simple and easy to administer, the ORF provides a quick method of obtaining empirical information on the progress of students. This information can be particularly valuable in contexts in which teachers and administrators are expected to demonstrate proficiency gains in the student population. External assessment tools in the field of ELT are most often commercially prepared and administered at a single point in time (e.g., TOEIC, TOEFL, and IELTS). Because such tests are designed to measure the full spectrum of English proficiency from beginner to advanced, they lack the sensitivity to demonstrate progress over the short term. Moreover, they provide no diagnostic information concerning the nature of the test-taker's deficiencies. A further problem of using external tests for benchmarking and progress-monitoring purposes is that scores are interpreted on a standard scale of scores and averages. These averages often differ substantially from those of an institutional sample of students. The ORF allows teachers to compare individual student data to others in the same classroom or academic program. An institution may also collect normative data on the student population in order to provide a local normative framework for interpreting scores (Fuchs & Fuchs, 2007).

In the context of the SILC, there are numerous reasons to enact forms of Curriculum-Based Measurement such as the ORF. Firstly, there is a strong institutional imperative to demonstrate improvement in core English language skills among the student population. At present, the preferred index of English proficiency is the TOEIC. As mentioned previously, a standardized assessment tool such as the TOEIC is inappropriate in the context of the SILC. Therefore, the ORF could function as part of a suite of alternative assessment tools capable of demonstrating improvement in core skills during the course of mandatory English language study.

Secondly, with the introduction of a fully-fledged extensive reading program in the second year of the curriculum, it is becoming increasingly important to monitor students' reading performance. Administering the ORF at the beginning of the English curriculum allows teachers to evaluate the students' reading level and establish appropriate goals for the course of study. Without such testing, it is very difficult for teachers to estimate the reading competencies of their students much less make informed decisions on the most appropriate forms of intervention. Considerable curriculum time and resources have been dedicated to the extensive reading program, yet at present there is no mechanism for establishing its efficacy. Although existing management tools, such as Moodle Reader, enable teachers and administrators to make deductions about the reading habits of the students, they do not provide evidence of gains in reading proficiency. The ORF, if properly implemented, could provide meaningful data on the overall effectiveness of the extensive reading program.

Finally, the second year curriculum also features a speed reading program in which students are encouraged to read simple texts as fast as possible

in order to assist with the development of automatic processing. Although data related to a learner's reading rate is collected throughout the program, this data is self-reported and lacks the credibility of more objective measures. In general, speed reading programs seek to measure a learner's silent reading rate. However, it is reasonable to hypothesize that improvements in the silent reading rate would correlate with gains in oral reading fluency. As such, collecting ORF measures both before and after the completion of a speed reading program should allow teachers and administrators to assess the effectiveness of such programs within a given context.

2. Design and Research Questions

The present study is the continuation of a pilot study which was first initiated in June of 2013. Both the original and subsequent investigations attempted to address a number of practical and theoretical questions.

1. Can the ORF be implemented quickly and effectively by classroom teachers with limited training?
2. Are ORF scores diagnostic of reading deficiencies?
3. Are ORF scores an indicator of general English proficiency?
4. Can OFR performance norms be established according to CEFR proficiency scales?

The aim of the researcher was to establish appropriate protocols for administering large scale ORF testing that may occur in the future as well as explore hypothesized correlations between ORF scores and general English proficiency (as measured by the OOPT). The existing literature pertaining to Curriculum-Based Measurement and the ORF suggest that ORF testing could serve as a useful diagnostic of chronic reading deficiencies. As such, the investigations sought to explore the extent to which the ORF could illuminate the nature of the difficulties faced by second language readers undertaking the SILC curriculum. Finally, the pilot and subsequent studies would initiate data collection leading to the establishment of ORF performance norms for the population of SILC students.

3. Methodology

The researcher administered the ORF to three classes of incoming English Communication 1 students during the third week of the Spring semester of 2014. The classes represented the following departments: Computer and Information Sciences, Life Science, Aerospace Systems Engineering, Nanoscience, and Biotechnology. Students were recorded reading a level-appropriate passage aloud for one minute. Reading rate was calculated by counting the number of words read correctly per minute (WCPM). Accuracy was calculated by dividing the WCPM by the total number of words read.

Prior to the test, students were given a short orientation by their classroom teacher. The orientation explained the nature of the test and what was expected of the participants. An instruction card written in Japanese was also prepared and shown to each participant prior to the commencement of the test. Consent to use the results of the ORF for the present research was obtained using a form located in the SILC's learning management system (Moodle). In addition to explaining the purpose of the research, the consent form also made clear to the participants that results of the ORF would not be used in calculating grades for English Communication 1.

The ORF test sheet and scoring papers were designed using a template available at www.interventioncentral.org/teacher-resources/oral-reading-fluency-passages-generator. Training in administering the ORF was also obtained at www.interventioncentral.org. The text used for the test was taken from the University of Oregon's

DIBELS 6th Edition benchmark assessment materials for 3rd grade (Good & Kaminski, 2007).

Scoring of the ORF was completed by the researcher during the administration of the test and confirmed at a later date using the recordings made at the time of testing. Students were grouped into CEFR proficiency levels according to results of the OOPT. Mean reading rate and accuracy were calculated for each group. Standard deviations were also calculated for each of the group means. Results were compared to the normative framework established for American school children to enable reporting of results by grade level.

4. Results

The ORF was administered to a sample of 79 English Communication 1 students for the purpose of the present study. Participants were divided into groups according CEFR proficiency levels as reported by the OOPT. The mean reading rate and the standard deviation were calculated for each of the groups along with the mean accuracy figures. The results are displayed in Table 1.

Table 1
Mean reading rate and accuracy as well as standard deviations for each. Participant population grouped according to CEFR Proficiency Scale.

	A1	A2 (low)	A2 (high)
n	25	43	11
WCPM (mean)	86.36	97.77	111.81
Accuracy (mean)	96.16%	96.72%	98.45%
STD (WCPM)	19.27	20.25	21.94
STD (Accuracy)	4.29	2.94	1.21

Scores from the OOPT and the ORF test were compared to see how well they correlated. Results indicate a moderate positive correlation between the two variables — r(79) = 0.39 — according to criteria for correlations in the behavioural sciences (Cohen 1988). Figure 1 summarizes the results.

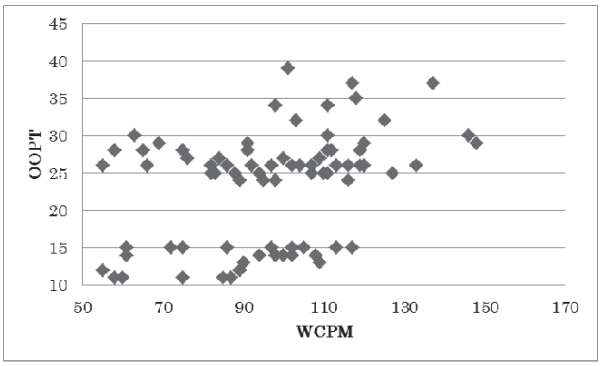

Figure 1. Scatter Plot of Student Scores from the Oxford Online Placement Test and ORF Test.

5. Discussion

The present study was undertaken with the aim of establishing ORF performance norms according to CEFR proficiency levels. The ORF measure was administered to a sample of 79 incoming SILC English Communication 1 students. Results were analyzed with reference to OOPT scores obtained at the commencement of the academic year. Although the present study resulted in the establishment of a performance norm for only one CEFR proficiency level (A2 *low*), a number of important questions concerning the feasibility and relevance of large scale ORF assessment were answered.

Considering the relative prominence of reading skills development in the SILC curriculum, it is important to identify and implement an instrument capable of measuring gains in reading proficiency. Concerning the first research question posed by the present study, it has been demonstrated that ORF testing can be administered quickly and easily with minimal training and only modest expenditures of valuable curriculum time. The related metrics are easy to calculate and meaningful without complicated conversions to alternative indices.

Concerning the second research question, it is clear to the researcher that conducting ORF testing

illuminates the reading challenges of individual learners while exposing broader trends among the population. Perhaps the greatest advantage of the ORF is the wealth of diagnostic data it provides to the classroom teacher about the L2 reading skills of their students. Low reading rates indicate a lack of automaticity and poor sight word recognition skills while low accuracy reflects poor decoding skills or a tendency to read orthographically. Administering the ORF measure at the beginning of a course of study provides teachers with the information required to determine appropriate remediation for struggling readers.

With reference to the third and fourth research questions posed by the present study, the data collected thus far indicates a moderate correlation between OOPT and ORF scores. In other words, higher overall language proficiency is a moderate predictor of higher measures of reading fluency. Whether this correlation is strong enough to warrant the establishment of ORF performance norms according to CEFR proficiency scales is yet to be determined. Additional data are required in order to reveal the efficacy of establishing norms in this way. Considering the scope and purpose of the OOPT, it is unreasonable to expect an overly strong correlation with a reading skills measure such as the ORF. Indeed, considering that the OOPT purports to provide an overall measure of English language ability, a correlation coefficient above 0.5 could raise questions of construct validity.

In general terms, the purpose of undertaking Curriculum-Based Measurement such as the ORF may be summarized as follows:

1) establish a baseline of proficiency
2) measure potential gains
3) assess the efficacy of classroom interventions

Further studies will explore the potential of using the ORF to accomplish these objectives within the context of the SILC.

6. Conclusion

The present study set out to establish ORF performance norms according to CEFR level as reported by the OOPT. Although a performance norm was established for only one proficiency level, the study served as a valuable pilot for future exploration of the applications of the ORF in the context of the SILC. There were a number of limitations to the present study, the most significant of which was the small sample size. As only 79 students were administered the ORF, the extent to which the results can be presumed to apply to the larger constituency of SILC students is modest at best. Future studies should continue to explore the correlation between OOPT and ORF scores for estimating the reading ability of incoming students. One benefit of the present study is that any subsequent use of the ORF (provided it is administered shortly after the OOPT) can contribute to the data set required to fulfil the mandate of the present study.

It is too early to assess the substantive impact of implementing ORF testing at the institutional level. Further studies are needed to determine whether or not the ORF is a useful metric for demonstrating student gain or assessing the efficacy of classroom reading interventions. However, the simplicity of the ORF to use, understand and administer make it an attractive alternative to commercial tests such as the TOEIC, TOEFL, and IELTS.

References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30*, 507-526.

Fuchs, L. S. & Fuchs, D. (1996). Combining performance assessment and curriculum-based measurement to strengthen instructional planning.

Learning Disabilities Research and Practice, 11, 183-192.

Fuchs, L. S. & Fuchs, D. (2007). Using curriculum-based measurement for progress monitoring in reading. *National Center on Response to Intervention*.
Retrieved from http://www.rti4success.org.

Gersten, R. & Dimino, J. A. (2001). The realities of translating research into classroom practice. *Learning Disabilities Research and Practice, 16*, 120-130.

Good, R. H., & Kaminski, R. A. (2007). Dynamic Indicators of Basic Early Literacy Skills (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.
Retrieved from http://dibels.uoregon.edu.