

Balancing Real-Time vs. Postperformance Feedback for EFL Presentation Classes

Don Hinkelman

Sapporo Gakuin University

Matthew Cotter

Hokusei Gakuen Junior College

Reference Data:

Hinkelman, D. & Cotter, M. (2018). Balancing real-time vs. postperformance feedback for EFL presentation classes. In P. Clements, A. Krause, & P. Bennett (Eds.), *Language teaching in a global age: Shaping the classroom, shaping the world*. Tokyo: JALT.

During an 8-year action research project investigating self-, peer, and teacher assessment in EFL presentation classes, a teaching team at a Japanese university built an online video recording process to optimize teacher and student ease of use. This resulted in greater student understanding of rubrics, student scoring that was more consistent with teacher scoring, and improvement in subsequent student performances. Based on the principles of *assessment for learning* (Gardner, 2012) and *embedded assessment* (Wiliam, 2011), assessments were designed within learning activities. However, as self-, peer, and whole class assessments increased in time and number, the assessment load increased for both students and teachers. Using a collaborative action research approach, teachers decided to increase real-time coaching and reduce postperformance feedback. This paper is a report on 2017 classroom trials to balance a variety of feedback formats, roles, and timings. Results include reports of student satisfaction, teacher experimentation with mixed-modal assessments, and balanced types of learning feedback.

大学の教員チームは、プレゼンテーションの授業で自己評価、ピア評価、教員による評価をする8年間の行動研究プロジェクトを行った。その間に、評価を容易にするためにビデオレコーディング利用の評価方法を構築した。これにより、ループリックに対する学生の理解度が高まり、学生の自己採点と教師の採点結果の整合性が向上し、その後の学生の成績が向上した。「学習評価」(Gardner, 2012)と「組み込み評価」(Wiliam, 2011)の原則に基づき、評価は学習活動の中で設計された。しかし、自己評価、ピア評価、全体評価の時間と回数の増加により、評価をする際の負荷が学生と教員共に増加した。協働行動研究のアプローチを利用して、教員はリアルタイムコーチングを増やし、実績後のフィードバックを減らすことに決めた。本論では、様々な

フィードバック形式、役割、タイミングのバランスをとるために、2017年の授業内試験について報告する。結果には、学生の満足度の高さ、混合様式評価の実験、および学習フィードバックのバランス型が含まれる。

As second language education increasingly moves away from testing of knowledge retention to assessment of task performance (McNamara, 1996; Wigglesworth & Frost, 2017), important questions are raised as to how to assess student performance, who should do the assessment, and how to give individual feedback to improve performance. In this study we examined these questions in the context of an EFL oral communication class on presentation skills for university English majors in Japan. As a framework for this research, four principles from classroom practice and educational theory were used: (a) assessment for learning; (b) embedded feedback; (c) multimodal, blended technologies; and (d) transparent learner analytics.

Principle 1: Assessment for Learning, Not Grading

In contrast to summative assessment, or *assessment for grades*, formative assessment is concerned with *assessment for learning*. Formative assessment enables students to learn through the assessment itself (Gardner, 2012; Wiliam, 2011) through qualitative feedback given back to learners. This study was focused exclusively on formative assessment, although some of the assessment results were also included in the grading process.

Principle 2: Embedded, Unlimited Feedback

Wiggins (2012) defined feedback as actionable advice to students; upon retrying a task, performance is directly improved. Thus, a presentation skills class, which is almost entirely devoted to *performance* rather than knowledge, must focus on feedback (given synchronously face-to-face and asynchronously on paper or online). However, the amount of feedback given to learners is an important issue. Wiggins (2012, August 27) suggested that there is no limit to the amount of feedback that teachers can give: “Learning is caused by learners attempting to

do something and getting feedback on the attempt. So learners need endless feedback more than they need endless teaching” (para. 4). This study was focused on the amount and media format of feedback used in an EFL presentation skills class.

Principle 3: Multimodal, Blended Technologies—Online, Paper, and Voice

Since the 1970s, video recording technology has been used persistently for speech performance assessment. However, due to the complexity and cost of file handling, this technology has been limited to professional training in small classes. Advances in digital recording and the ease of programming open source learning management systems mean that this is now easier than expected, and large class use of video recording in EFL programs is now possible. In second language learning, Purpura (2016) noted that as technologies are blended, comprehensive change in assessment design and development is more possible, allowing automated scoring, corpus analyses, feedback delivery, and validation. However, blind adoption of new technologies can lead to unintended consequences, as the learning may be affected in the interplay of complex ecologies of pedagogic action, timings, groupings, spaces, texts, and tools (Gruba & Hinkelman, 2012). In a formative assessment process, the online, paper, and face-to-face technologies can be blended locally to achieve values of classroom purpose, appropriateness, multimodality, and sustainability. In some cases, low-tech, locally configured ecologies are constructed rather than high-tech, globally published packages (Hinkelman & Gruba, 2012), which is an example of teachers gaining greater power and autonomy.

Principle 4: Learner Analytics for Transparency

In addition, formative assessment can also be enhanced by learner analytics. By mining data from networked student data, teachers can understand what students are learning, how well they have learned, and where students are falling behind or moving ahead. Dufour (2004) recommended learner analytics for collaborative teaching teams in order to “quickly learn when a teammate has been particularly effective in teaching a certain skill, . . . attempt to replicate it in their own practice, [and] . . . identify areas of the curriculum that need more attention” (p. 10). Furthermore, learner data assists students in self-evaluation and aids teachers in giving individual attention: “Educators must begin to embrace data as a useful indicator of progress. They must . . . focus on the success of each student” (p. 11). Learner analytics forces transparency when the numbers reveal inconsistencies and outlying successes and failures. Yet the cost of transparency can be impractical and become an unsustainable practice if teacher workload to provide analytics proves burdensome.

Research Questions

Based on these principles, this 8-year study examined the primary research question of “How can feedback and assessment improve EFL presentation performance?” Each year, a more specific research question or questions were chosen based on reflection by the teaching team. The research questions for the present cycle of study were created in face-to-face focus group sessions and online forums by the team of three teachers. This same team developed rubrics, activities, and online tools to teach, coach, and assess EFL oral presentation skills in the class. In the months before the 2017 cycle, this team identified two problems on which to focus research: (a) time-consuming, overassessment and (b) possible mistimed feedback. Thus, the specific questions selected for this cycle were articulated as follows:

- RQ1. How did teachers adjust the balance of real-time and out-of-class feedback in class? (timing)
- RQ2. Who should provide the feedback? What types of real-time and out-of-class feedback do students and teachers feel is important for learning and improvement? (teaching)
- RQ3. What is an effective and manageable balance of online, paper, and voice tools for these kinds of feedback? (tools)

These questions are not “either/or” questions, but rather questions of balance, in which complex relations in a learning ecology are continually improvised and improved. Figure 1 illustrates these relations.

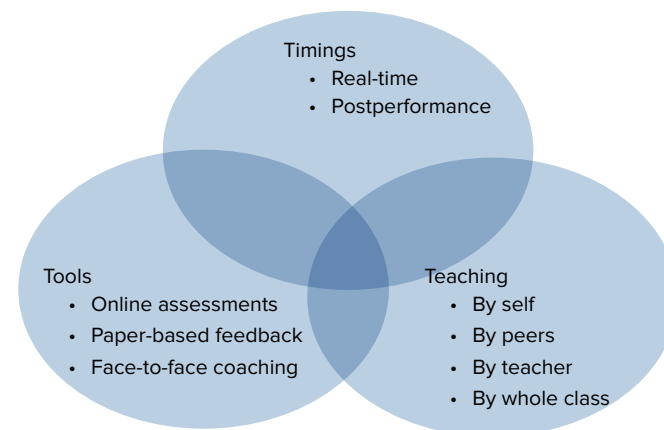


Figure 1. A learning ecology for feedback in EFL performances.

Research Design, Participants, and Course Background

Over the course of 8 years, collaborative teaching teams at Sapporo Gakuin University developed a curriculum for the class titled “Oral Communication C,” which was aimed at improving performance of oral presentation skills of 2nd-year English department majors. The curriculum was focused on intermediate presentation skills for professional use (teacher training, business promotion, seminar leadership). Curriculum themes were nonverbal skills (30%), visual design and explanation skills (30%), and verbal structure and persuasive language skills (40%).

Participants in this longitudinal study involved three course administrators/teachers, who designed the syllabus, wrote and illustrated the teaching materials, and also taught the class sessions (collaborative aims, independent lesson plans), as well as the students themselves. Each year, 50-60 students were enrolled in the compulsory class. The students were divided into three streamed classes of 15-20 students and taught for 90 minutes per week over 15 weeks in a single semester.

Action research was chosen as the research methodology as it involves cycles of reflection and intervention, which lend themselves well to longitudinal study (Nunan & Bailey, 2009). Action research is significantly different from *reflective practice* in that data is collected, analyzed, and interpreted (Burns, 2017). A research timeline of the development of the 8 yearly cycles, including three published reports, is summarized in Table 1.

Table 1. Action Research Timeline of Assessment of Student Presentations

Cycle	Year	Description of interventions
One	2010	Adapted syllabus design: Physical/visual/story messages based on Harrington and Lebeau (2009). Unified syllabus across all classes. Agreed on types of and criteria for presentation assessments.
Two	2011	Added video recording of presentations in all classes for more accurate assessment. Used unlisted YouTube accounts to store and assess videos. Linked videos to class website. Postperformance viewing of the videos allowed teachers the time to assess rubric criteria, tabulate the data, and provide feedback either to individuals or the class as a whole (reported in Rian, Hinkelman, and McGarty, 2012).

Cycle	Year	Description of interventions
Three	2012	Designed new LMS (Moodle) video assessment module to allow (a) direct video upload and storage, (b) playback within class website alongside assessment rubrics, (c) teacher <i>one-click</i> grading on rubric criteria and comment boxes, (d) automatic grading of rubric criteria and rubric display to students. Module design described in Hinkelman (2014).
Four	2013	Added self- and peer assessment weightings for video assessment module. Weightings added to allot 10% of grading to self-assessment, 10% to peer assessment, and 80% to teacher assessment.
Five	2014	Teacher, self-, and peer postperformance assessments were compared and reported in Rian, Hinkelman, and Cotter (2015). Students' self-assessments were lower compared to their teachers' assessments. Students scoring of themselves (self-evaluation) and others (peer evaluation) had modest variance to teacher scoring.
Six	2015	Tried whole class synchronous assessment option for video assessment module on student mobile phones. Real-time (synchronous) assessment online replaced paper forms to achieve similar results but allow tabulation of results to be included in grading.
Seven	2016	Added rubric precalibration option for video assessment module. Teachers uploaded a demonstration presentation video after already inputting model scores for each of the criterion on the rubric. The students then watched and scored the presentation and on completion, saw if their scores matched teacher scores. Tests of the pre-calibration feature were deemed unsuccessful and excessive and therefore discontinued.
Eight	2017	Focused on synchronous, in-class teacher coaching. To balance the growing assessment load, in-class teacher assessment (real-time coaching) was increased and out-of-class self- and peer assessment was reduced.

Table 1 shows the evolution of assessment types, tools, roles, and rubrics in a single class repeated over 8 years. During the first two cycles of action research, the teachers

Hinkelman & Cotter: *Balancing Real-Time vs. Postperformance Feedback for EFL Presentation Classes*

experimented with types of assessment formats and a variety of assessors (self, peer, and teacher), which was reported in Rian, Hinkelman and McGarty (2012). The video assessment module was later re-engineered to include three types of postperformance assessment—teacher, self, and peer—as described in Hinkelman (2014). Then the scores of teacher, self-, and peer postperformance assessments were compared in Rian, Hinkelman, and Cotter (2015). Overall, the results showed greater student understanding of rubrics, more consistent student scoring with teacher scoring, and also improvement in subsequent student performances.

Data Collection and Analysis

The steps of the 2017 research cycle involved the following procedures:

1. End-of-term survey: A summarized chart of student evaluations of the class (Appendices A & B).
2. End-of-term reflection: A recorded focus group discussion with the teachers (Appendix D).
3. Intervention selection: Meeting to review survey, reflection, and previous cycle data analysis. Carried out before syllabuses were submitted in February of each year.
4. Implementation during 15 classes with weekly teacher logs—from April through July.
5. Data analysis of current cycle: downloaded assessment scores from website. Summarized results and compared self-, peer, teacher, class assessment (Table 3). Transcribed student coaching (Appendix C).
6. Reported results in teacher conferences (regional JALT, international JALT, and in-house faculty meetings) in fall/winter of each year.
7. Publishing results in peer-reviewed journals or conference proceedings.
8. Data summaries and samples are shown in Appendices A-D and Tables 2-4.

Results and Discussion

In 2017 (Cycle 8), a qualitative approach analyzed four types of data:

1. Student satisfaction surveys (averages tabulated),
2. Learner performance assessment scores and comments (numerical data online),
3. Teacher weekly teaching reflection logs (text in online forum), and
4. End-of-term focus group discussion of three teachers (video-recorded).

1. With regard to the student survey, data summaries are displayed in Appendix B and C, showing that students found value in both real-time and postperformance self- and peer assessment. However, students in general found more value in feedback from their teacher than from other sources. Therefore, not only using the online asynchronous assessment tools, but also real-time teacher coaching was deemed important by the students (Appendix D). How to provide the best balance, timing, and workload for these assessment methods is the major question of this study.

2. Results from 2017 online learner assessment scores, highlighted in Table 2, draw some consistencies with previous years. Students continued to score themselves lower on postperformance assessment tasks than teachers on all presentations. An average total over the five presentations saw self-assessments 9% lower than teacher assessments. It is encouraging that students did not take advantage of an opportunity to raise their scores, but graded themselves more severely than did their teachers. This is consistent with the general tendency of Japanese students to rate themselves modestly. Due to this high variance between teacher and self-ratings, a lower weighting of 20% was assigned to self-assessment scores than to teacher scores (80%).

Table 2. Student and Teacher Scoring of Presentations in 2017

Type of assessment	Criteria	Presentation number (number of assessments)					Average
		1 (n = 40)	2 (n = 48)	3 (n = 36)	4 (n = 34)	5 (n = 49)	
Class (Real-time)	Number of assessments	38	46	36	33	-	
	Score (out of 100)	81.4	70.9	79.0	79.1	-	77.6
	Variance with teacher assessments	-6.5	-17.2	+4.6	-9.3		-7.2
Self (Post-performance)	Number of assessments	12	42	35	14	45	
	Score (out of 100)	81.2	69.8	69.2	82.9	76.2	75.9
	Variance with teacher assessments	-6.7	-18.3	-5.3	-5.6	-9.1	-9.0

Hinkelman & Cotter: *Balancing Real-Time vs. Postperformance Feedback for EFL Presentation Classes*

Type of assessment	Criteria	Presentation number (number of assessments)					Average
		1 (n = 40)	2 (n = 48)	3 (n = 36)	4 (n = 34)	5 (n = 49)	
Peer (Post-performance)	Number of assessments	0	0	0	0	0	
	Score (out of 100)	-	-	-	-	-	-
	Variance with teacher assessments	-	-	-	-	-	-
Teacher (Post-performance)	Number of assessments	37	48	36	34	49	
	Score (out of 100)	87.9	88.0	74.4	88.4	85.3	84.8

The most obvious difference in 2017 is that teachers deemed it necessary to drop postperformance peer assessment (watching recorded videos of classmate presentations) in order to reduce the workload for students and increase presentation preparation time. However, class, self-, and teacher assessment continued. Previous findings (Rian, Hinkelman, & Cotter, 2015) showed that postperformance peer assessment closely mirrored teacher scores, but in-class, real-time assessment of peers' presentations yielded scores that were much higher than postperformance scores. Possible reasons could be inadequate time for students to contemplate rubric criteria in detail or students being distracted from scoring due to watching presentations. With postperformance scoring, students received a second chance to see the presentation while watching the video at their own pace and scoring on the rubric.

3. A weekly reflection log was carried out by teachers using an online forum on Moodle and the data obtained is shown and analyzed in Table 3. Comments were selected by purposive criteria and extract codes assigned to each one. These codes were then collated and given theme titles (Miles, Huberman, & Saldana, 2014).

Table 3. Teaching Log Extract Summary on Theme of Feedback

Theme	Extract codes	Comments by teachers
Types and timing of feedback	[Types/spectrum of feedback]	<ul style="list-style-type: none"> Students marked the presentation scripts with triangles, circles, and arrows. I checked and revised each one individually. Also, I found that grammar problems, missing slides, and missing sentences were difficult to give feedback on orally. So I decided to move to paper feedback. With paper feedback the “red pen” marks were clear and they had to fix them before doing the presentation again. I had them make a script that combined the slide with the spoken text side by side. The ideal printing format was “haifu shiryō-3” available in Word both in Mac and Windows.
	[Postperformance verbal comments]	
	[Preperformance written comments]	
Feedback roles	[New type of feedback: multimodal transcripts]	<ul style="list-style-type: none"> After each performance, I said 1-2 good points and 1-2 suggestions for improvement. I will note how often I do this, and whether this affects later presenters
	[Precalibration alternatives]	
	[Student assigned verbal commenting]	
Effectiveness of feedback	[Need for required commenting in peer assessment]	<ul style="list-style-type: none"> My verbal coaching was a waste of time. He was not prepared, many absences. Needed to show me a script. Too many details for him to catch and he was not interested in hearing my comments. Doing a good/bad presentation in front of class, with students marking me—I think was the most effective.
	[Failure of coaching]	
	[Most effective feedback—good/bad demos]	
Aims, goals, standards, rubrics of feedback	[Most effective feedback—paper scripts with slide images]	<ul style="list-style-type: none"> Not passing someone until they could meet the standard was important. Knowing the standard meant that I could coach confidently and clearly. How do you give feedback to students who are not in class? I often heard the words “I was nervous,” but often with a look of accomplishment. I told them conquering nervousness was the number 1 aim, and they will be able to claim this skill when they are job hunting.
	[Rubric minimum standards]	
	[Standards affected by absenteeism]	
	[Affective goals-confidence]	
	[Motivation by awards and badges]	

Hinkelman & Cotter: *Balancing Real-Time vs. Postperformance Feedback for EFL Presentation Classes*

Theme	Extract codes	Comments by teachers
Content of feedback	[Coaching points-home position] [Analytical thinking and explaining] [Extemporaneous topics] [Weakness of story message]	<ul style="list-style-type: none"> Students did too many gestures nervously, I taught how to do the “home position”—holding their hands together at their waist. Students need research and analytic skills, so the comparison table and graph description helps them logically summarize relevant points.
Feedback tools & activities	[Better audio recording tools] [Need for higher quality video resolution] [New activity to train slide explanations]	<ul style="list-style-type: none"> We might want to think about using a microphone too as some of the audios were poor. Students should practice explaining a pre-set slide set. In pairs, A/B, they explain using template sentences, add questions,

Table 4 shows changes in types of assessment, assessment variety timing, and tools (paper or online) during Cycles 3, 5, and 8. This table shows that teachers emphasized or reduced feedback types over time. To illustrate, an example of an intervention that was born from the teachers log was a new type of feedback piloted in this cycle—slide images with transcripts and teacher feedback and improvements (Figure 2). This format requires multiliteracies of the student—the ability to interpret and communicate combined images, voice, text, and physical motion (Brown, Lockyer, & Caputi, 2010).

Table 4. Assessment and Feedback: Progression of Timing, Tools, & Teaching

Assessment type	Cycle 3: 2012 Timing/ tools/teaching	Cycle 5: 2014 Timing/ tools/teaching	Cycle 8: 2017 Timing/ tools/teaching
1. Teacher assessment (in-class)	- paper rubric assessment with comments ✓ - oral feedback on good or weak elements (5 of 5 presentations)	- none ✗	- real-time comments - real-time demonstrations (5 of 5 presentations) ✓
2. Peer assessment (in-class)	- no formal evaluation ✗ (unrecorded paper check of inclusion or absence of 5 key presentation points)	- paper rubric assessment (4 of 5 presentations) ✓ - no comments	- real-time video assessment via smartphone on Moodle (4 of 5 presentations) ✓
3. Teacher assessment (out-of-class)	✓ - Moodle forum assessment with comments (2-3 presentations)	✓ - Moodle video assessment with comments (5 presentations)	✓ - Moodle video assessment with comments (5 presentations)
4. Peer assessment (out-of-class)	✓ - Moodle forum assessment with comments (2-3 presentations)	✓ - Moodle video assessment with comments (4 presentations)	- none ✗
5. Self-assessment (out-of-class)	✓ - Moodle forum assessment with comments (2-3 presentations)	✓ - Moodle video assessment with comments (4 presentations)	✓ - Moodle video assessment with comments (3 presentations)

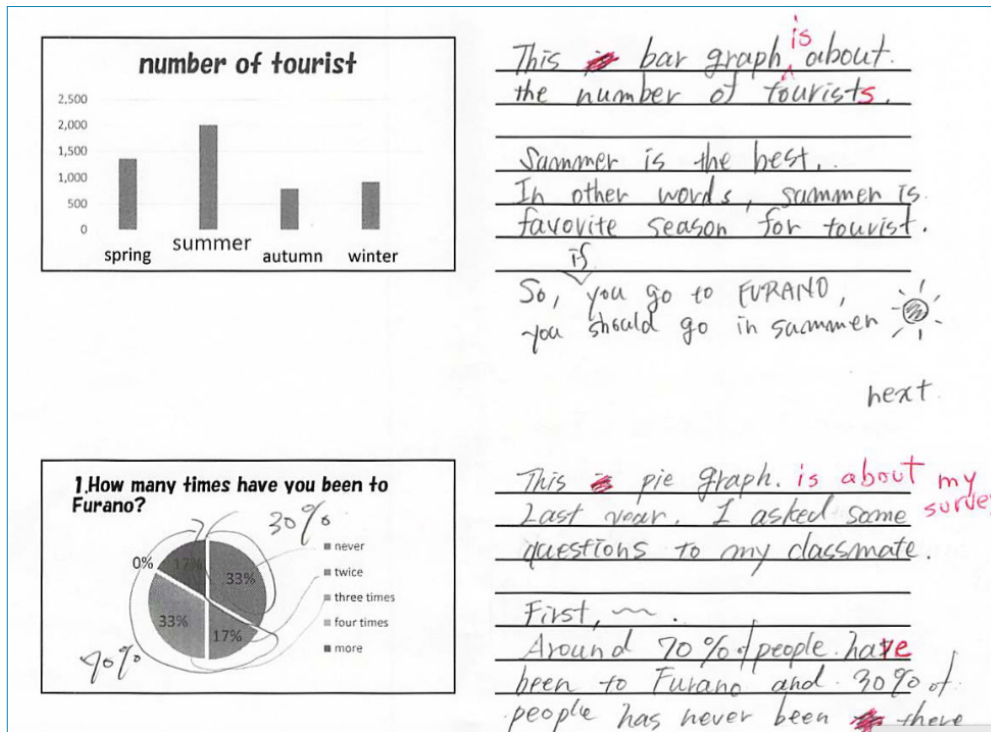


Figure 2. Presentation slide transcripts with teacher edits.

4. At the end of the semester a final group reflection and focus discussion was carried out by the three instructors. Video recording of the discussion allowed for key points to be reviewed and recorded at a later date with an aim at deciding future goals and an intervention plan for the 2018, ninth cycle in the study. Key points of the discussion were transcribed (Appendix D) and used for selecting goals, aims, and interventions for 2018.

Interventions for the 2018 Cycle

Student opinions from surveys, teacher observations and conclusions, and also the analysis of quantitative and qualitative data as detailed above were used to formulate action research interventions for 2018. Subsequently an intervention list of the key points for inclusion in 2018 was created and is shown in Table 5.

Table 5. Key Points for 2018 Interventions

Intervention	Justification
Materials (Desk resource - textbook)	<ul style="list-style-type: none"> introduce refined course content (vocabulary and assessment rubrics) for easy referral by both students and teachers
Materials (Online resource - Moodle)	<ul style="list-style-type: none"> new flipped quizzes to incorporate elements of assessment precalibration to ensure scoring proficiency further development of the video assessment module to include a possible portfolio and badges for assessment completion and achievements
Syllabus change	<ul style="list-style-type: none"> include impromptu type warm-ups and activities to prepare students for limited preparation time, real-life presenting situations

Conclusion

The need for endless feedback, proposed by Wiggins (2012, August 27), may be a useful general principle for teaching performance skills, in order to emphasize that “teaching” is less effective than “feedback.” However, taken literally, the results and evaluation of this classroom practice in 2016 found that increasing out-of-class video assessment reached a point of unsustainability. High numbers of self- and peer assessment were found to be counterproductive for student satisfaction and increased teacher management burdens. In response, our interventions in 2017 showed that balancing limited online assessment with targeted teacher coaching and demonstrations (face-to-face synchronous feedback) and written, paper script feedback was necessary.

Our findings also suggest teachers of performance skills should increase the variety of feedback and find a balance of teaching/coaching, timings, and tools/media. Our proposed intervention, real-time coaching, did not provide a sole answer to the question of appropriate feedback. Teachers experienced both failure and success with real-time coaching: They found it especially useful for nonverbal skills (voice, posture, gesture, eye contact) but less so for verbal skills. To cope with verbal skill feedback, new multimodal assessment interventions were tested, with teachers marking printed visual slides side-by-side with student transcripts of the spoken words accompanying those slides (Figure 2). Teachers could give specific, actionable suggestions that students could use in the

Hinkelman & Cotter: *Balancing Real-Time vs. Postperformance Feedback for EFL Presentation Classes*

next iteration of task performance. In this case, paper as a media evolved as a new form of assessment and superseded online assessment in the learning ecology.

Although teachers should continue to use recorded assessments and give weight to a variety of assessors, they should embrace multiple formats of online, paper, and voice feedback. In this way, we noticed that in our classroom practice, we used a variety of *blended* assessments. There were no pure *e-assessments* as categorized by Gardner (2012), because our assessments tend to be blended with a mix of online, paper, and face-to-face aspects.

This research confirms the model of learning ecologies as complex blends of timings, tools and teachers that evolve in classroom use and require balancing rather than optimizing. Therefore, generalizable claims of *best practice* or *most effective* are not relevant. Instead, the *small cultures* of innovation and change predicted by Holliday (1999) is a more powerful metaphor to describe how classroom assessment systems of student performance grow and succeed. In classroom research and task performance assessment, we discovered learning ecologies that combine teaching, groupings, spaces, timings, texts, and tools in unique configurations that can be documented and shared as insights but not solutions.

Bio Data

Don Hinkelman is a professor of English communication and intercultural communication at Sapporo Gakuin University. His research has focused on how blended language learning principles can combine online and face-to-face technologies. He has authored *Blending Technologies in Second Language Classrooms* and articles in intercultural communication. For sharing open resources, including the video assessment module code, he coordinates an open courseware website of the Moodle Association of Japan. <hinkel@sgu.ac.jp>

Matthew Cotter is a lecturer at Hokusei Gakuen University and also a part-time lecturer at Sapporo Gakuin University. His research interests include CALL- and CLIL-based approaches to learning English content, in particular intercultural communication, sport, and also indigenous language and culture. He has received Best Open Courseware awards from the Moodle Association of Japan and has developed the Essay (auto-grade) and ReactForum plugins. <m-cotter@hokusei.ac.jp>

References

- Burns, A. (2017). Envoi: From reflective practice to action research? In R. Barnard & J. Ryan (Eds.), *Reflective practice: Voices from the field* (pp. 140-155). New York, NY: Routledge. <https://doi.org/10.4324/9781315397665>
- Brown, I., Lockyer, L., & Caputi, P. (2010). Multiliteracies and assessment practice. In D. R. Cole & D. R. Pullen (Eds.), *Multiliteracies in motion* (pp. 191-206). London, England: Routledge.
- Dufour, R. (2004). What is a professional learning community? *Educational Leadership*, 61(8), 6-11.
- Gardner, J. (Ed.). (2012). *Assessment and learning* (2nd ed.). London, England: Sage. <https://doi.org/10.4135/9781446250808>
- Gruba, P., & Hinkelman, D. (2012). *Blending technologies in second language education*. London, England: Palgrave-MacMillan. <https://doi.org/10.1057/9780230356825>
- Harrington, D., & Lebeau, C. (2009). *Speaking of speech*. London, England: MacMillan Education.
- Hinkelman, D. (2014). Rubric-based assessment for video-recorded learner performances: Blending paper rubrics with a Moodle LMS module. In *MoodleMoot Japan 2014 Proceedings* (pp. 63-69). Tokyo: Moodle Association of Japan.
- Hinkelman, D., & Gruba, P. (2012). Power within blended language learning programs in Japan. *Language Learning & Technology*, 16(2), 46-64.
- Holliday, A. (1999). Small cultures. *Applied Linguistics*, 20(2), 237-264. <https://doi.org/10.1093/applin/20.2.237>
- McNamara, T. (1996). *Measuring second language performance*. London, England: Longman.
- Miles, M., Huberman, M., & Saldana, J. (2014). *Qualitative data analysis* (3rd ed.). Thousand Oaks, CA: Sage.
- Nunan, D., & Bailey, K. (2009). *Exploring second language classroom research: A comprehensive guide*. Boston, MA: Heinle.
- Purpura, J. E. (2016). Second and foreign language assessment. *The Modern Language Journal*, 100(S1), 190-208. <https://doi.org/10.1111/modl.12308>
- Rian, J. P., Hinkelman, D., & Cotter, M. (2015). Self-, peer, and teacher rubric assessments of student presentation videos. In P. Clemens, A. Krause, & H. Brown (Eds.), *JALT2014 Conference Proceedings* (pp. 688-697). Tokyo: JALT. Retrieved from <http://jalt-publications.org/proceedings/articles/4765-self-peer-and-teacher-assessments-student-presentation-videos>
- Rian, J. P., Hinkelman, D., & McGarty, G. (2012). Integrating video assessment into an oral presentation course. In A. Stewart & N. Sonda (Eds.), *JALT2011 Conference Proceedings* (pp. 416-425). Tokyo: JALT. Retrieved from <http://jalt-publications.org/proceedings/articles/1765-integrating-video-assessment-oral-presentation-course>
- Wiggins, G. (2012). Seven keys to effective feedback. *Educational Leadership*, 70(1), 10-16.

Wiggins, G. (2012, August 27). Less teaching and more feedback? [Web log message]. Retrieved from <http://inservice.ascd.org/less-teaching-and-more-feedback/>

Wigglesworth, G., & Frost, K. (2017). Task and performance-based assessment. In E. Shohamy, I. Or, & S. May (Eds.), *Language testing and assessment* (pp. 121-133). New York, NY: Springer Intl. https://doi.org/10.1007/978-3-319-02261-1_8

Wiliam, D. (2011). *Embedded formative assessment*. Bloomington, IN: Solution Tree Press.

Appendix A

Review of Student Responses to Selected Course Surveys, 2012-2017

What helped improve your presentations?	2012 (N = 49)					2013 (N = 31)					2014 (N = 53)					2017 (N = 48)					Total Agree Number on Likert >4 / n (%)
	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1	5	4	3	2	1	
Likert Scale																					
Watching videos of my own presentations.	16	23	10	0	0	16	12	3	0	0	17	28	5	0	1	20	20	5	3	0	152 /181 (84%)
Assessing my own presentations.	-	-	-	-	-	-	-	-	-	-	13	27	6	0	0	11	24	11	2	0	75/101 (74%)
Watching classmates' live presentations.	13	26	8	2	0	12	16	2	1	0	21	25	4	0	1	27	12	7	2	0	152/181 (84%)
Having classmates assess my presentations in real time.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	16	16	14	2	0	32/48 (67%)
Feedback from the teacher.	29	16	4	0	0	18	11	2	0	0	24	25	2	0	1	34	8	6	0	0	165/181 (91%)
Watching videos of classmates' presentations.	11	31	5	2	0	11	14	5	1	0	16	29	7	0	0	-	-	-	-	-	112/133 (84%)
Feedback from classmates.	10	22	16	1	0	10	13	6	2	0	14	24	10	0	2	-	-	-	-	-	93 (70%)
Postassessing my classmate's presentations.	-	-	-	-	-	-	-	-	-	-	8	26	10	0	0	-	-	-	-	-	34/53 (64%)
Having classmates postassess my presentations.	-	-	-	-	-	-	-	-	-	-	11	26	7	8	1	-	-	-	-	-	37/53 (70%)
What helped you the most? (n=48)	Watching videos of my own presentations.					Watching in-class live presentations.					Watching example presentation videos.										
	21					27					0										

Appendix B

Comment Responses to 2017 Student Survey

10. What else helped you improve? Please explain. そのほかに、ためになった・助かったことあれば是非、具体的に教えてください。

- 他人のプレゼンを見て、自分をもっと改善しなければいけないところは何かをしっかりと学び、7月10日の時に成果が少しでも出たこと。

- The teacher taught to me advice. It was very good. Thank you very much!!

Teacher

- I think that presentation skills are important for job.

I recommend keeping presentation in this class.

Tool

- My teacher checked our presentation by one by that is one of the good thing for me.

I can see my improve points by myself!

I can learn anything to do a presentation, so I want to try many times in our life. Thank you!

- Nothing.

- Mr. Hinkelman's advice.

Teacher

- I Very enjoyed this class!

- teaching me how to presentation is good for my future. Hope this class will keep going.

Tool

- Making a presentation repeatedly.

- I can show video of presentation by myself.

- I could have a presentation in front of international students.

- 私は紙を見ないで練習しながら、徐々にジェスチャー等を付け加えていく練習と、ほかの人の発表を見てポイントを絞るのがためになりました。

- All teachers are very kindly so that is very helped me.

Appendix C

Video Transcriptions of Real-Time Presentation Feedback by Teacher (2017)

Teacher A to Student A

Well done! Good gestures. Good slides. Very easy to understand. You just need one more slide with a line or a bar graph. And this pie graph is good but make it bigger, bigger letters. Good! And a good pace!

Teacher A to Student B

That was really good! For the first introduction slide say a few things about an outline, or this is what I will talk about. Maybe say a little less on the comparison slide.

Teacher A to Student C

Well done. I think you could have had more pictures. Maybe in the outline or in the front page here. Because Disneyland has some great pictures...maybe you could have a castle or mickey mouse etc. Also, your comparative table, you only had two things. Price and location. So maybe one more thing to compare would be good. Anything is OK...like most delicious food or something.

Teacher A to Student D

Good job! Good gestures, like using action gestures. Probably you need one more slide at the end here to finish off. Also, this is simple and good but it is Shiretoko! More photos... beautiful nature photos! You only have these three. Here, put another one to make it more pretty!

Teacher A to Student E

OK...good job. I think you could have some more pictures too, but you also need a comparative table. It is good to have some Japanese on the slides but you need to explain it in English for example say this is data from the 2011 earthquake and tsunami so you ...

Appendix D

Transcribed Extracts From Teacher Focus Group Discussion

Video reference <https://www.youtube.com/watch?v=VycXBLwJMUc&feature=youtu.be>

0:08 Q1 What did we do differently?

Time	Teacher	Quote
0:55	Matt	Assessment wise we didn't assess peers at all.
1:03	Don	Last year I did 5 peer assessments and this year zero, which I am fine with. We did more class assessments. We did more class assessments. Last year I did five self-assessments, and this year three presentation 2, 3 and 5.
4:40	Matt	I think last year I was relying more on them looking at the videos themselves, assessing themselves and getting peers to assess them and give comments and the teacher too. Because this year we weren't doing the peer one this year I spent more time after each presentation... asking the class what did they do well, what didn't they do well and then putting my comments on top of that... and I think that the students got comments at the start and afterwards when they did self assessing and they had my comments there again, it was like "ah yeah that's right" it was kind of more of a reminder. But if we didn't have the assessment after...it might have been gone. Kind of like a back up it was quite good.
5:58	Don	I changed the coaching. I doubled it. Especially at the beginning presentations as that affected the later ones.
6:22	Don	In fact, I stopped people on their posture. I got up there and I said, "OK move your body." They were facing the screen. This year I felt that in Presentation 5 almost everyone was facing the audience.
6:47	Don	Maybe we are spending too much time on the physical message and maybe we have gotten better, we can do less. They are mastering it.
7:09	Don	The only weak point of their physical message is their stress. A few of them got it but some...

8:01 Q2 What worked well?

Time	Teacher	Quote
8:04	Rob	I like how organized the classes was, no surprises, you were never caught off guard, you know what's coming... there is no panic.
8:24	Don	I think if we look at the numbers of the presentations, the scores were better this year.
8:32	Rob	They were better, my class, they were heads and shoulders above the students last year.
8:40	Matt	I agree.
9:13	Don	This is an example of a guy who has a low GPA but he did well in this class. Because he started with some confidence and it got refined.
10:02	Rob	I don't think we should worry about grades being too high as long as we can show what they did. We want to set them up for success.

10:20 Q3 What didn't work well?

Time	Teacher	Quote
11:02	Don	The quality of the videos was poor in presentation 5. And that was I set them on QVGA which helped with the uploading speed but hurt the visibility. It has got to be VGA.
12:12	Rob	We might want to think about using a microphone too as some of the audios were poor too.
12:44	Rob	My suggestion would be to be to check out the material from other places and see what they are doing...based on this rubric it would be possible to pass this speech without speaking a lot of English. Because you have too much weight on slides, too much weight on physical appearance, stress. It collapses under its own weight...it is too detailed. Some of the categories could be meshed together.
13:58	Rob	Also the slides.... I think we spending way too much time making slides.
14:45	Don	This is the only presentation skills class our students get in the university so there is a value to the non-English part their overall development.

Hinkelman & Cotter: *Balancing Real-Time vs. Postperformance Feedback for EFL Presentation Classes*

15:10	Rob	There are going to be times when you are going to have to make a presentation and you aren't going to have the luxury of three weeks of slide making time. You will have to speak in 5 minutes. The world won't wait for you to make a PowerPoint presentation.
15:20	Matt	Should we bring an impromptu one back in? (Agreed)
18:49	Rob	I think it would work if we did one every week and built from that and did it as a warm up
19:05	Matt	It would be good if are decreasing what they are getting good at with the physical message and decreasing slides as well we should have time for that (and add it as warm up)
19:22	Don	One thing I am unhappy with the class is that the verbal message gets caught in the last three weeks before the conference. We cut the transitions, the words they use to give evidence and things
19:50	Rob	I would limit the slides to 5-7 slides...have only one prep day for slides. The rest they have to do on their own.