

Using Collaborative Testing to Better Evaluate L2 Classroom Learning

Glenn Davies

Temple University, Japan Campus

Although collaborative learning activities are now common in contemporary second-language classrooms, they are often accompanied by traditional assessment methods, such as eliciting knowledge and skills from students individually. These assessment methods, however, are inadequate as they provide an incomplete picture of learner abilities and ignore some of the essential communication skills that teachers actually focus on in their teaching. This article looks at the L2 classroom use of collaborative testing. It argues for greater use of this alternative testing method and suggests ways in which it could be introduced to complement current approaches to classroom assessment.

現在の第二言語学習の授業では、協働学習が一般的になっているが、その際には、個々の学生の知識や技能を問うような、従来の評価方法が用いられることが多い。しかし、この評価方法では、学習者の能力を完全には把握できず、教師が実際に教える際に重視する本質的なコミュニケーション能力を測れないため、評価方法としては不十分なものである。本論では、L2クラスでの協働テスト実施について考察している。また、この評価方法を、これまでの方法に替わるものとしてさらに活用すべきであり、現在の授業評価をさらに補完するために導入するべきだと提案している。

<https://doi.org/10.37546/JALTTLT47.3-1>

Testing is an essential part of any educational program. When used before, during, or at the end of a course, tests “measure individual competence of students in their thinking skills and subject-matter knowledge and expertise” (Webb, 1995, p. 240). Despite the multitude of possible individual and institutional differences in the goals of testing, its basic purpose remains the same: the measurement of relevant knowledge and skills. However, it is important that the tests in use sample their entire domain. Applying this principle to the testing of L2 learning, our assessments must be connected to the communication skills currently being taught in our classes.

Research on teaching tasks and techniques shows not only that group work in the classroom has become commonplace, but that facilitating learning and working as part of a team is also now essential (Cheng & Warren, 2000). One common pedagogical approach to group work is the use of collaborative learning activities. Collaborative learning (CL) has

been defined as “group learning activity organized so that learning is dependent on the socially structured exchange of information between learners” (Candlin et al., 2003, p. 339). Perhaps directly born from Vygotsky’s ideas of learner development and cooperative learning (Cole et al., 1978), collaborative tasks enable students with different abilities to work together toward group goals or to engage in activities where individual success influences group success and learning occurs through interaction in a social context (Hassanien, 2006).

Once collaborative learning tasks are in place, the next logical step would be to implement collaborative testing to measure their associated outcomes. Also referred to as group testing, double testing, paired testing, cooperative testing, and dyad testing (Akiyamen et al., 2017), collaborative testing is a student-centered, active learning approach to assessment, in which students are tested in groups. The skills targeted by this approach (and those that later should be tested) include noticing, self-correction, rephrasing, increased exposure to the language being practiced, improved vocabulary growth, improved written or spoken output based on shared input, and an overall increase in student confidence with what is being learned. If we accept that testing in general should seek to measure acquisition of knowledge with some meaningful, real-world application beyond the classroom, individual student answering of multiple-choice questions alone seems unlikely to achieve this objective. By contrast, a shift in focus to L2 communicative production reveals the potential of group work, or collaborative tasks, as an authentic means of assessing learners’ capacity for later language use (Davies, 2009).

Literature Review

The use of collaborative learning and testing tasks has shown benefits in numerous subject areas, ranging from chemistry, biology, and math, to atmospheric and computer sciences, forestry, and food systems (Clarkston & Gilley, 2014). For instance, Heller and Hollabaugh (1992) utilized cooperative grouping in physics classes to teach and test problem-solving abilities. They found the use of group

tasks to be an effective means of solving problems, with 72% of the participants reporting that group discussions helped them understand the course material. Moreover, research by Heiner and Rieger (2014) concluded that collaborative exams were not only easy to set up, but that they also resulted in deeper learner engagement and promoted more effective support for and evaluation of learning.

Research on collaborative testing methods evidences how they can positively impact various aspects of learning. For instance, students have been shown to demonstrate significantly greater improvement on subsequent individual testing after having been tested in groups as opposed to having been initially tested only as individuals (Clarkston & Gilley, 2014). Moreover, in an investigation of problem-based learning (PBL) in which students collaboratively activated their knowledge to complete group tasks, Mennin (2007) concluded that “students [found] PBL to be challenging, satisfying, and motivating” (p. 305). Thus, having learners engage and perform in these types of activities elicits not only their knowledge, but also their actual communication skills.

Possible Approaches

One approach to collaborative testing is the use of two-stage exams in which students first complete and submit an exam individually and then work in small groups to answer the same exam questions again. During the group work stage, students receive immediate, targeted feedback on their solutions from their classmates while becoming aware of other ways of looking at the problems (Clarkston & Gilley, 2014). This approach is perhaps the most logical and balanced as it allows the teacher to get a sense of a student’s abilities both as an individual and as a social actor. Although developing two-stage exams creates an additional burden for teachers and institutions, the fuller picture of learner abilities that this type of exam potentially provides is arguably worth the extra effort. Moreover, the initial set-up and administration time may not be so different from most current practices, the main additional consideration perhaps being how to reliably elicit representative participation from each member during the group stage.

A second, and quite common, approach to collaborative testing is the use of pair-to-group-to-class activities, in which exam topics or questions are first shared and discussed in pairs, then in ever-increasingly numerous groups, from 4 to 6 to 8 and so on, until the entire class is involved in the same discussion. This type of activity allows students

to confirm or correct their understanding of the content on which they have just been tested. This approach can positively influence the motivation level of each group member through success with the activity as well as ensure that everyone has learned the material and is motivated to teach the others (Slavin, 1996).

A third alternative to traditional individual testing is the use of self, peer, and collaborative assessment tasks (Falchikov, 1986). These tasks allow for three stages of assessment, namely self-evaluation, then peer discussion, and finally small group discussion of the correct and incorrect answers to questions on tests taken individually. The idea here is to focus on learners’ ability to explain their positions or thought processes on answers given during individual testing. This approach is particularly useful in the analysis of incorrect multiple-choice answers. It also enables the teacher to better understand student reasoning for choosing specific answers and may even justify the awarding of partial test credit. Most importantly, this three-stage testing approach provides a clearer picture of learners’ overall abilities, the weighting of each depending on the priorities of each individual practitioner or institution. However, Davies (2009) suggests useful scoring techniques for each stage, ranging from evaluation packs including self-evaluation components and numerical scores to one-off peer evaluations with rankings of team members.

As a final suggested collaborative approach to classroom assessment, debate activities based on topics from exam readings or semester-specific tasks can be used to elicit a wide range of communication skills. Although not a new idea, the addition of these types of activities to the established assessment regimen provides a look at actual linguistic performance, as opposed to the mere knowledge that is typically measured by simple quizzes or multiple-choice tests. As with the previously suggested collaborative tasks, the benefits of these types of activities include “a potentially less stressful environment for new, international or less social students to interact with their peers, and a better quality of work that can be produced” (Davidson et al., 2014, p. 117). Furthermore, as Slavin (1996) notes, “students will learn from one another because ... cognitive conflicts will arise, inadequate reasoning will be exposed, disequilibrium will occur, and higher quality understandings will emerge” (p. 49). In short, even though the target of these types of assessment activities remains that of collectively testing current learner abilities, this approach entails the additional advantage of the social learning that takes place in the process.

Conclusion

As it is generally agreed that students learn well when they work together, collaborative learning is now a staple of many classrooms. Collaborative learning enables students to work together to achieve group goals and promotes individual success through interaction. However, traditional testing methods are often mismatched with these learning outcomes. As an alternative approach to assessment, collaborative testing more realistically mirrors not only what learners do in the classroom but also what they will be expected to do in the world beyond. Although care is necessary to ensure reliable measurement of individual learner abilities in group settings, by expanding our testing repertoire to include collaborative activities, we can hope to ensure that we are assessing our learners' abilities in a way that more accurately represents what it is that they are truly learning.

References

- Akiyamen, L. E., Cantwell, E. R., Jadotte, Y. T., Pierce, J., & Sousou, J. (2017). Collaborative testing for improving student learning outcomes and test-taking performance in higher education: A systematic review. *Campbell Systematic Reviews, 13*(1), 1–18. <https://doi.org/10.1002/CL2.186>
- Candlin, C. N., Rose, K. R., & Sachs, G. T. (2003). Developing cooperative learning in the EFL/ESL secondary classroom. *RELC Journal, 34*(3), 338–369. <https://doi.org/10.1177/003368820303400305>
- Cheng, W., & Warren, M. (2000). Making a difference: Using peers to assess individual students' contributions to a group project. *Teaching in Higher Education, 5*(2), 243–255. <https://doi.org/10.1080/135625100114885>
- Clarkston, B., & Gilley, B. H. (2014). Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students. *Journal of College Science Teaching, 43*(3), 83–91. https://www.sei.ubc.ca/bitstream/seima/2072/1/Gilley-Clarkston_2-Stage_Exam_Learning_JCST2014.pdf
- Cole, M., John-Steiner, V., Scribner, S., & Souberman, E. (Eds.) (1978). *Mind in society: The development of higher psychological processes*. L. S. Vygotsky. Harvard University Press.
- Davidson, P. M., Deek, H., Disler, R., Hickman, L. D., Jackson, D., Potgieter, I., & Power, T. (2014). Small group learning: Graduate health students' views of challenges and benefits. *Contemporary Nurse, 48*(1), 117–128. <https://doi.org/10.1080/10376178.2014.11081933>
- Davies, W. M. (2009). Groupwork as a form of assessment: Common problems and recommended solutions. *Higher Education, 58*(4), 563–584. <https://doi.org/10.1007/s10734-009-9216-y>
- Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self-assessments. *Assessment and Evaluation in Higher Education, 11*(2), 146–166. <https://doi.org/10.1080/0260293860110206>
- Hassanien, A. (2006). Student experience of group work and group assessment in higher education. *Journal of Teaching in Travel & Tourism, 6*(1), 17–9. https://doi.org/10.1300/J172v06n01_02
- Heiner, C. E., & Rieger, G. W. (2014). Examinations that support collaborative learning: The students' perspective. *Journal of College Science Teaching, 43*(4), 41–47. https://sei.ubc.ca/bitstream/seima/2222/1/Rieger-Heiner_2-stage-Exams_JCST2014.pdf
- Heller, P., & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups. *American Journal of Physics, 60*(7), 637–644. <https://doi.org/10.1119/1.17118>
- Mennin, S. (2007). Small-group problem-based learning as a complex adaptive system. *Teaching and Teacher Education, 23*(3), 303–313. <https://doi.org/10.1016/j.tate.2006.12.016>
- Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*(1), 43–69. <https://doi.org/10.1006/ceps.1996.0004>
- Webb, N. M. (1995). Group collaboration in assessment: Multiple objectives, processes, and outcomes. *Educational Evaluation and Policy Analysis, 17*(2), 239–261. <https://doi.org/10.3102/01623737017002239>

Glenn Davies has been teaching and working in the ESL field in Japan since 1997. He received his Master's degree in Science in Education (MS.Ed) from Temple University Japan Campus, where he also currently teaches in the Continuing Education program. His specialist interests are language testing methodology and collaborative learning. He can be contacted at <davies@tuj.temple.edu>



JALT2023 – Growth Mindset in Language Education

Tsukuba, IBARAKI
November 24~27, 2023
<https://jalt.org/conference/>