The Use of Distributed Consensus as a Method for Analyzing Peer Assessment

By
Brian Patrick THOMAS*

Abstract

In most classroom settings peer assessment as a method of formative assessment has proved beneficial for increasing student responsibility and autonomy, as well as lifting the role of the student from passive to active learner. However, because of cultural predilections towards conflict avoidance, neutrality, peer acceptance and face-saving dispositions, the effectiveness of peer assessment with Japanese students is less than optimal even under ideal conditions. While, under typical conditions it can be completely ineffective. In recent times, the application of distributed consensus, and the principals of trustless systems as outlined in "The Byzantine Generals' Problem" have received much attention in the realm of technology, for example the use of trustless systems. This paper intends to demonstrate that these principles of distributed consensus along with objective criteria can also be applied, in a practical way to a classroom setting as a method of increasing the effectiveness of peer assessment with Japanese students.

Key Words: formative assessment, distributed consensus, peer evaluation

1. Introduction

Formative assessment is an invaluable tool in contemporary pedagogy for providing feedback from the students to the classroom teacher. The feedback can be used both in real time to make adjustments to the course of an individual lesson, or be used to inform the planning of future lessons to maximize student progress. Of the many methods of obtaining this feedback from students, peer assessment provides a unique set of benefits for both students and teachers. Peer assessment allows students to develop self-evaluation skills by objectively assessing their peers while encouraging a more active role in the learning process. It also allows students to improve their understanding of assessment criteria and the task assigned. Teachers benefit from peer assessment in terms of timely feedback and reduced workload. In particular, it

* Senior Assistant Professor, Sojo International Learning Center

can alleviate some of the difficulties associated with collecting feedback from large classes.

The students in the context of this article are participants in mandatory English courses that take place in the first 4 semesters at university. Though there are some exceptions, it is fair to say that the majority of the students do not have specific English language goals. The passing of the required courses is their only goal. As such, they tend to be minimally invested in learning outcomes. This coupled with the Japanese student's tendency to avoid conflict and disagreement often results in peer assessments in which students give neutral to slight positive assessments of their peer's performance rather than an objective and accurate assessment, thus nullifying the potential benefits of such an activity.

By developing peer assessment activities, based on a specific set of criteria, that require students to collect objective, quantitative data from their peers' work or performance, then viewing the collected data with the very basic principles as outlined in the "The Byzantine Generals' Problem," it is possible to maximize the benefits of peer assessment for both students and teachers. This approach will minimize the student's ability and desire to offer inaccurate responses for the sake of conflict avoidance, as people endeavor to make their actions accepted by society by trying to avoid threatening face or face-threatening acts (FTA) (Brown & Levinson, 1987). This is perhaps further complicated by the complex social structure of Japanese society. By asking objective questions, we are soliciting objective answers. Teachers, rather than looking at individual responses from individual students as a form of feedback, view the data as a whole, looking for consensus or lack of consensus in student responses. This enables teachers to identify inconsistencies in student learning as well as underdevelopment of particular skills, so as to better target future learning. This paper will detail some of the difficulties in using peer assessment with typical Japanese students and offer a possible solution to increasing the effectiveness of such activities by applying a quantitative rather than qualitative means of data collection, followed by viewing these data through the prism of distributed consensus.

2. Peer assessment with Japanese students

One of the prominent difficulties in using peer assessment as a tool with Japanese students is that, at its core, peer assessment is an overt criticism of one's peers, played out with the students being more or less a captive audience to the criticism. Students of the same major, with relationships that extend beyond the language learning classroom into other spheres, are being asked to potentially cast a negative light upon each other. One must consider that people tend to avoid disagreement and seek agreement (Nozawa, 2015; Liddicoat, 2007) or in this case, avoid being disagreeable and seek to be agreeable by providing a positive assessment of their peer's performance.

Additionally, Japanese people tend to prefer saving the face of others over their own (Nozawa, 2015; Kiyama, Tamaoaka & Takiura 2012; Tanaka, 1988), which could lead to a student's willingness to assess his fellow students inaccurately, even at the expense of his own overall performance.

Furthermore, "groups play an important part in Japanese society. And not surprisingly, this is also found in the classroom" (Moxon 2009, para. 8). The power structure in the Japanese classroom is not a clear-cut dichotomy of teacher/student roles. The existence of the role of 'sempai' in the classroom is a clear example of this. 'Sempai' can be translated as 'group leader', but this does not accurately convey the extent and importance of this role in the classroom. The position of 'sempai' may be determined by age, or knowledge of a

subject. It is a position of trust and authority, often times the 'sempai' is the link between the students and the teacher (Moxon, 2009, para. 8). Would a typical student be willing to objectively assess the performance of his 'sempai' and present that assessment before the class and teacher? In the very least, one may surmise that is complicates the matter.

In a pilot study of self-evaluation and peer evaluation conducted at Waseda University "some students commented that they were not confident in rating other students' speeches. They also mentioned that they felt peer pressure" (Oi, 2012, p. 6).

3. Objective peer assessment based on criteria

Studies in Japan have indicated positive learner attitudes towards peer assessment (Oi, 2012; Baierschimidt, 2012), but as Baierschimidt points out there is a need to take precautions to ensure that students' cultural backgrounds are not interfering with the peer feedback process (2012, p. 109). Additionally, there has been some indication that "peers are prone to produce ratings based on uniformity, race and friendship if there is no extensive training in peer rating." (Dancer & Dancer 1992; Sluijsmans 1998, p. 301). Two of the most predominate limitations in peer assessment are friendship marking, which results in overly positive results, and collusive marking which may result in lack of differentiation among peers (Pond, Ul-Haq, & Wade, 1995). These limitations are by no means unique to Japanese students; however they may be exacerbated in the Japanese classroom for the reasons mentioned above.

In the development of peer assessment materials the inclusion of questions that seek to collect quantifiable data on a peer's performance are beneficial as the answers tend to resist the student's ability to engage in friendship or collusive marking. For example, if a student is asked to complete a questionnaire of a peer's writing assignment, a question such as "Was your classmate's writing long enough?" is perhaps not as useful as "Was your classmate's writing at least 100 words long?" Questions of quality, such as, "Was the writing interesting?" will not be as informative for the teacher and are likely to yield inaccurate answers. Instead, "How many examples were included in the body of the paragraph? What were they?" is more likely to yield accurate answers and provide useful data for the teacher. The most straightforward way to develop peer assessment materials, therefore, is to work directly from the established criteria, such as the rubric used by the teacher for the assignment, to form questions that are most likely to elicit objective responses. These data are more likely to better inform the teacher, increasing the effectiveness of peer assessment as a means of formative assessment. There is also the added benefit of ensuring that the students fully comprehend the criteria by which they are being assessed.

4. Application of distributed consensus in the analysis of peer assessment data

The Byzantine Generals Problem is an analogy used to express how distributed computer systems may tolerate inconsistencies within a closed system. Below is a brief summary from the abstract of the original paper.

Reliable computer systems must handle malfunctioning components that give conflicting information to different parts of the system. This situation can be expressed abstractly in terms of a group of generals of the Byzantine army camped with their troops around an enemy city. Communicating only by messenger, the generals must agree upon a common battle plan. However,

one or more of them may be traitors who will try to confuse the others. The problem is to find an algorithm to ensure that the loyal generals will reach agreement. (Lamport, Shostak, & Pease, 1982)

On the surface, the ideas expressed in this abstract seem worlds away from the classroom, unless we decide to view the classroom itself as a distributed system, with each student representing one component in that system. In this system various input is supplied primarily by the teacher, but also through student contributions. Likewise, feedback loops are generated by these interactions. Making sense of contradictory or inconsistent feedback from students can prove a daunting for a teacher. However, by applying some of the very basic concepts outlined in the Byzantine Generals Problem, a teacher may view the data more holistically and the inconsistencies themselves may aid the teacher in obtaining accurate feedback from the class.

In a typical peer assessment activity a teacher might supply a questionnaire or checklist for students to utilize while engaging with a peer's work. A student's writing or a performance during an oral presentation are common activities that lend themselves well to this type of peer assessment. The resulting data can be shared with the teacher and or student being assessed. The dilemma arises as to what to do or how to process or interpret the information.

Technology affords teachers the opportunity to collect data on class performance as well as have the data analyzed and presented in a form that can be easily understood by both teachers and students. Online services such as SurveyMonkey or Google Forms are great vehicles for collecting data and are often used as a way to obtain feedback from students, for example end of semester feedback on overall course quality. These same tools, however, are generally underutilized in relation to formative assessment.

What follows is one method of collecting and processing data from peer assessment activities so as to maximize the usefulness and authenticity of said data by mitigating the disruptive effects of phenomenon such as friendship and collusive marking.

Initially, a questionnaire is written by the teacher based on established criteria of the task being assessed. Special care should be given to write questions that elicit factual responses as these responses typically would not require coding by the teacher. The completed questionnaire is input to the program of choice and distributed to the students.

Students must have access to individual computers, tablets or smart phones. Students are then put into small groups where they perform (in the case of an oral presentation) or trade their completed assignments (in the case of a written assignment). Students complete the questionnaire based on their classmate's performance. This continues round-robin style until all group members have been assessed. The teacher then assigns new groups and the process is repeated. In each round care should be given to ensure that students see and assess as many individuals as possible. This will guarantee a large enough sample size in order to distinguish between accurate and outlying responses. It also has the added benefit, in the case of presentation type projects, of providing students with ample opportunity for rehearsal.

The resulting data can automatically be displayed as graphs or pie charts. Consensus or lack of consensus can easily be identified and outlying responses can be given less weight than the predominate responses. What results is a collective assessment of a student's performance that is more accurate than an individual peer

assessment. This data is better suited to inform a teacher's future planning and provides students with more desirable objective feedback on their individual performance.

Conclusion

As outlined in this paper, despite its benefits, the use of peer assessment with Japanese students presents its own set of challenges as a result of complex social influences. Rather than dismiss it as ineffective for Japanese students, it is worthwhile to seek out better methods of implementing peer assessment that take into account the nature of the classroom environment.

Developing criteria-based materials that elicit objective answers can help to mitigate issues such as friendship and collusive marking; however it cannot completely eliminate these issues. Therefore, further care must be taken to ensure accurate results.

Rather than relying on individual peer assessment results, by collecting multiple results in a distributed fashion then viewing the data collectively with readily available technological resources, one is able to obtain accurate feedback with which to better inform both teachers and students.

References

Baierschimidt, J. (2012). Japanese learner attitudes towards peer feedback. *Kanda University of International Studies Minutes*, 24, 101-114. Retrieved from http://id.nii.ac.jp/1092/00000605/.

Brown & Levinson (1987). Politeness; Some universals in language use. *Cambridge University Press*.

Dancer, W.T. & Dancer, J. (1992). Peer rating in higher education. *Journal of Education for Business*, 67, 306-309.

Kiyama, S., Tamaoaka, K., & Takiura, M. (2012). Applicability of Brown and Levinson's politeness theory to a non-western culture: Evidence from Japanese facework behaviours. *SAGE Open*, 1-15.

Lamport, L., Shostak, R. and Pease, M. (1982). The Byzantine Generals Problem. *ACM Transactions on Programming Languages and Systems*, 4(3), pp.382-401.

Liddicoat, A.J. (2007). *An introduction to conversation analysis*. London: Continum.

Moxon, A. (2009). Understanding your Japanese students. *IH Journal*, 27. Retrieved from http://ihjournal.com.

Nozawa, Y. (2015). The use of hedges in polite disagreement by Japanese advanced learners of English. *Waseda University Graduate School of Education*.

Oi, S.Y. (2012). A Pilot Study of self-evaluation and peer evaluation. *Selected Papers of the 17th Conference of Pan-Pacific Association of Applied Linguistics*. Retrieved from www.paaljapan.org.

Pond, K., Ul-Haq, P. & Wade, W. (1995). Peer review: a precursor to peer assessment. *Innovations in Education and Training International*, 32, 314-323.

Sluijsmans, D., Dochy, F., & Moerkerke, G. (1998). Creating a learning environment by using self-, peer-, and co-assessment. *Learning Environments Research*, 1, 293-319.

Tanaka, N. (1988). Politeness: Some problems for Japanese speakers of English. *JALT*, 9(2), 81-102.