

## Student Engagement and Assessment Modes: a Study

Anil Pathak

Centre for Communication, Teaching and Learning, Universiti Teknologi Brunei, Brunei Darussalam

---

### Article Info

#### Article history:

Received Jan 1, 2018

Revised Jan 31, 2018

Accepted Feb 27, 2018

---

#### Keyword:

Communication

Oral presentation

Peer assessment

---

### ABSTRACT

The aim of this project is to attempt a factorial analysis of the congruence amongst three layers of assessments of Oral Presentation: *Expert, Self, and Peer*. Participants included graduate and undergraduate students of Asian background studying Research Writing at a technological university. The research instrument consisted of a set of assessment checklists with weighted and specified criteria for Expert, Self and Peer assessment of research presentations. Broader criteria were based on the following four factors: Use of spoken media, Use of visuals, Interrelationships, and Communicative Effect. Participants received training and were involved in a practice session on self and peer assessment before engaging in the assessment tasks. While the pilot study is unable to conclusively establish congruence among the modes of assessment, it is expected that results from a wider database would indicate the nature of well-defined assessment criteria and the ways they can be effectively communicated to assessors. The present study suggests that the differences in ratings given by peers and experts are not due to a lack of capability on the part of students, but due to lack of orientation and training. It is necessary to evolve criteria that are well-documented to enable students to evaluate their own performance as well as the performance of their peers as effectively as done by experts.

Copyright © 2018 Institute of Advanced Engineering and Science.  
All rights reserved.

---

### Corresponding Author:

Anil Pathak,  
Centre for Communication, Teaching and Learning,  
Universiti Teknologi Brunei,  
Brunei Darussalam  
Email: anil.pathak@utb.edu.bn

---

## 1. INTRODUCTION

This paper outlines a research proposal to conduct an investigation into a specific aspect of the dynamics of oral presentations. It aims to attempt a factorial analysis of the congruence amongst three layers of assessments of Oral Presentation: Expert, Self, and Peer.

Competence related to Oral Presentation is usually considered to consist of the verbal, the non-verbal, and the visual. In professional contexts, however, four other components become relevant from the delivery aspects [1]. They are: 1) Professionalism: The way a presenter manages the technical content of the presentation; 2) Practical aspects: The way a demonstration is planned and executed; 3) Ideas: Presenting the ideas so as to highlight novelty, ingenuity, and creativity; 4) Collaborative effort: Co-operation among team members and integration of skills and abilities [2].

The project has specifically defined the three modes of assessment under discussion as follows. Self-assessment is defined following Klenowski (1995) as “the evaluation or judgment of ‘the worth’ of one’s performance and the identification of one’s strengths and weaknesses with a view to improving one’s learning outcomes” [3].

Peer assessment, in the context of this project has been defined as “an arrangement in which individuals consider the amount, level, value, worth, quality, or success of the products or outcomes of learning of peers of similar status [4].”

Similarly, Expert assessment has been defined as assessment usually carried out by the subject teacher/lecturer or a person identified as an expert in the field using pre-determined criteria communicated to the assessee well before the assessment [5].

## 2. EARLIER RESEARCH

A study by Luc et al. focused on the agreement between professional assessment and self- and peer assessment of oral presentation skills and investigates into student perceptions about peer assessment [6]. Since the main focus of this study is on the inter-rater reliability, it presents a comparison of the teacher and peer assessment scores and points towards a positive relationship. However, the data also indicates existence of some crucial differences. In general, the study suggests that peers and teachers interpret the criteria the rubric from fundamentally different perspectives. With regard to the comparison of self-assessment scores and teacher scores, the study concludes that there are significant differences between these scores, indicating an undesirable state of affairs. As is the general expectation, self-assessment scores are, in many cases, higher than the scores awarded by experts.

McGarr & Clifford reported on the use of peer learning and peer assessment with a cohort of four-year undergraduate physiotherapy students [7]. The study observed the students' engagement in the process based on a survey and focus group discussion. It was discovered that most of the students valued the experience of assessing their peers. However, when asked whether it was a fair mode of assessment, their responses were diverse. Similarly, when asked whether Peer Assessment should have a significant weight in the overall assessment, there was a general disagreement. The study highlighted the value of student engagement in peer learning and suggests that the experience does add pedagogical value to the overall educational experience. Panadero & Jonsson placed their focus on the use of rubrics to achieve congruence between self and peer assessment [8]. Their work highlights some important points regarding rubrics. While rubrics are scoring tools that need to be based on specific criteria, they can be used for summative or formative assessment purposes. It is generally seen that rubrics help improve transparency in the assessment procedure. Formative effects of rubrics are more lasting than the summative effects. The study concluded that effects of differences in personal attributes on the use of rubrics need further research.

While Self-Assessment and Peer Assessment is a well-researched area, its application to oral presentations has not been seen to be implemented in a rigorous manner. Secondly, only a handful of studies laid emphasis on the development of criteria, documentation, orientation, and training. The present study wishes to address some of these issues in greater detail. The significance of the present study is related to the greater weight being assigned to participatory evaluation. The area of the study is significant also due to the fact that by involving students in assessment, the dynamics of instruction can be remarkably improved. However, if the assessment administered by students is not systemic, such an exercise is likely to be futile.

The aim of the present study is to investigate the congruence (or lack thereof) among the three modes of assessment (Self, Peer, and Expert). The study also wishes to focus on the effects of assessor training on achieving such congruence among the three modes.

## 3. METHOD

Participants included graduate students of Asian background studying Research Writing at a technological university. The research instrument consisted of a set of assessment checklists with weighted and specified criteria for Expert, Self and Peer assessment of research presentations. Broader criteria have been based on the following four factors: Use of spoken media, Use of visuals, Interrelationships, and Communicative Effect. Participants received training and were involved in a practice session on self and peer assessment before engaging in the assessment tasks. Table 1 describes the criterias used for teacher (expert) assessment.

Table 1. Criteria used for Teacher (Expert) Assessment

Criterion	Weight (%)
Planning and Organisation	40
- Planning	
- Sequencing	
- Organisation	
Delivery	40
- Use of body language	
- Oral delivery	
Overall Impact	20
- Achievement of purpose	
- Impact on audience	

#### 4. RESULT

The following sample data analysis is based on the results of a Pilot Study carried out on 30 subjects (Table 2). The data was subjected to a Paired t-test ( $\alpha=0.05$ ).

Table 2. Comparison of Self-Assessment with Peer Assessment

	<i>Self</i>	<i>Peer (AVG)</i>
Mean	69.6	71.9
Variance	35.69	13.54
Observations	30	30
Pearson Correlation	0.40	
Hypothesized Mean Difference		0
Df		29
t Stat		2.24
P(T<=t) one-tail		0.016
t Critical one-tail		1.69
P(T<=t) two-tail		0.032
t Critical two-tail		2.045

Table 2 shows that the correlation between Self and Peer is 0.40 indicates a weak linear relationship among the two modes of assessment. This means that there is a general lack of agreement or congruence among the two sets of data indicating that the way students score themselves is significantly different from the way their peers score them.

Table 3 shows a Pearson Correlation Coefficient of 0.63 suggesting a strong positive relationship between Self-Assessment and Expert Assessment. This finding is significantly different from the findings reported in a number of earlier research studies.

Table 3. Comparison of Self-Assessment with Teacher (Expert) Assessment

	<i>Self</i>	<i>Expert</i>
Mean	69.6	60.63
Variance	35.69	14.72
Observations	30	30
Pearson Correlation	0.63	
Hypothesized Mean Difference		0
Df		29
t Stat		10.59
P(T<=t) one-tail		8.72
t Critical one-tail		1.69
P(T<=t) two-tail		1.74
t Critical two-tail		2.045

In terms of the relationship between Peer and Expert assessment, the situation is slightly better. The Pearson Correlation Coefficient here (Table 4) is 0.32 suggesting a very weak linear relationship. It can be said that students need further training regarding the use of criteria and rubric when it comes to assessing the performance of their peers.

Table 4. Comparison of Peer Assessment with Teacher (Expert) Assessment

	<i>Peer (AVG)</i>	<i>Expert</i>
Mean	71.9	60.63
Variance	13.54	14.72
Observations	30	30
Pearson Correlation		0.32
Hypothesized Mean Difference		0
Df		29
t Stat		14.06
P(T<=t) one-tail		8.65
t Critical one-tail		1.70
P(T<=t) two-tail		1.73
t Critical two-tail		2.045

Let us now consider other aspects of statistics as seen in the tables above. T is the calculated difference represented in units of standard error. The larger is the value of T (it can be either positive or negative), the stronger would be the support against the null hypothesis that there is no significant difference. A value of T closer to 0 indicates lack of significant difference. In the analysis presented in the tables above, for Self and Peer the T is 2.24. If this T value is compared to the value of T in other 2 sets (14.06 and 10.59), it is much lower and closer to 0. This means that there is a lack of significant difference in the two sets of data (Self and Peer). However, there is another interesting fact to observe here. In the case of the datasets related to Self and Peer (Table 2), since the p-value is very low ( $< \alpha$  level of 0.05), we need to reject the null hypothesis and conclude that there is a statistically significant difference in these two sets of data. Thus, the data (and, possibly, its interpretation) is anomalous and needs further investigation.

## 5. DISCUSSION

Considering the anomalies faced in the data analysis of the pilot study, a more systematic data collection over larger multiple samples is required. It is expected that results of such study would indicate what well-defined assessment criteria should include and how it can be effectively communicated to assessors. By implication, such criteria may enable students to evaluate their own performance as well as the performance of their peers as effectively as done by experts.

Validity in self-assessment usually implies congruence with the scores awarded by the expert or the teacher or an average of scores awarded by the peer group. The assumption here is that the teacher has an expertise that provides an absolute standard, and (with peer rankings), the mean of multiple rankings or scores is likely to be more reliable than the scores awarded by a single assessor.

It is usually seen that self-assessment scores have a higher congruence with peer judgments than with scores awarded by the teacher or the expert. An explanation of this phenomenon might be that peers interpret assessment criteria in a significantly different manner [9]. It is also possible that peers focus on features of the Oral Presentation which might be superficial [10],[11].

A number of studies have focused on the effects of training students towards self and peer assessment. The orientation and training usually consists of systematic explanation of each of the components that define the assessment. At the first stage, it might be worthwhile to involve assessors in defining and elaborating on the components of the assessment criteria. This stage would essentially involve development of assessment framework and discussion on the weight assigned to each of the component in the framework [5].

It would also be useful to have a discussion with the students regarding the application of the criteria to the actual performance. For self-assessment, perhaps it would be useful to ask students to consider using vide-recorded presentations. Model applications of the criteria and the rubric can be presented as examples of performance assessment. It is also worthwhile to provide feedback on self-assessments and peer assessments conducted by the students. This would necessitate engaging students in evidence-based discussions using the assessment data that particularly represents congruence (or lack thereof) among the three different modes of assessment [12].

Earlier research shows that interactive methods of teaching have a positive effect on student engagement [13]. A significant and desirable feature of any assessment is that the worth of an assessment is determined by its effects on the overall learning experience. This feature of assessment is usually referred to as Beneficial Backwash.

## 6. CONCLUSION

To conclude, the data related to t value is anomalous to our interpretations based on p value and Pearson correlation coefficient. Further investigation is needed to express this anomaly. While the pilot study is unable to conclusively establish congruence among the modes of assessment, it is expected that results from a wider database would indicate what well-defined assessment criteria should include and how it can be effectively communicated to assessors. By implication, such criteria may enable students to evaluate their own performance as well as the performance of their peers as effectively as done by experts.

## REFERENCES

- [1] A. Pathak, "Teaching and Assessing Multimedia-based Oral Presentations," *Business Communication Quarterly*, vol/issue: 64(4), pp. 63-71, 2001.
- [2] A. Pathak and M LeVasan, "Developing Oral Presentation Competence in Professional Contexts: A Design-Based Collaborative Approach," *International Journal of Evaluation and Research in Education (IJERE)*, vol/issue: 4(4), pp. 179-184, 2015.

- 
- [3] V. Klenowski, "Student self-evaluation processes in student-centred teaching and learning contexts of Australia and England," *Assessment in Education*, vol/issue: 2(2), pp. 145-163, 1995.
- [4] K. Topping, "Peer Assessment between Students in Colleges and Universities," *Review of Educational Research*, vol/issue: 68(3), pp. 249-276, 1998.
- [5] A. Pathak, "Presentation Media: Training the Technologists," *Teaching and Learning*, vol/issue: 21(1), pp. 45-52, 2000.
- [6] L. De Grez, *et al.*, "How effective are self- and peer assessment of oral presentation skills compared with teachers' assessments?" *Active Learning in Higher Education*, vol. 13, pp. 129-142, 2012.
- [7] O. McGarr and A. M. Clifford, "Just Enough to Make You Take it Seriously," *Exploring students' Attitudes Towards Peer assessment Higher Education*, vol. 65, pp. 677, 2013.
- [8] E. Panadero and A. Jonsson, "The use of scoring rubrics for formative assessment purposes revisited: A review," *Educational Research Review*, vol. 9, pp. 129-144, 2013.
- [9] A. Pathak and N. S. Chaudhari, "Medium is the Message: Web-based Lecture Presentations in Distance Education in Singapore," *Asian Journal of Distance Education*, vol/issue: 3(2), pp. 13-22, 2005.
- [10] A. Pathak and S. Kathpalia, "Discourse strategies for ensuring involvement in electronic lectures," *International journal of Web Based Communities*, vol/issue: 1(3), pp. 330-345, 2004.
- [11] A. Pathak, "Analysis of a Conversational Floor: Implications of Teaching and Assessing," *Online Discussion ELT Quarterly*, vol/issue: 12(3), pp. 1-10, 2011.
- [12] A. Pathak and S. B. Bee, "Identifying Learning Needs: Using Narrative Analysis," *Humanising Language Teaching*, vol/issue: 13(4), 2010.
- [13] A. Pathak and F. Cavallaro, "Teaching Conflict Resolution: A Study of Two Interactive Methods," *The Asian ESP Journal*, vol/issue: 2(2), pp. 5-11, 2006.