Practical Assessment, Research, and Evaluation

Volume 26 *Volume 26, 2021* Article 10

May 2021

Creating an Evaluation Factor for Group Work Assessment

Scott Mark
Bow Valley College

Recommended Citation

Mark, Scott (2021) "Creating an Evaluation Factor for Group Work Assessment," *Practical Assessment, Research, and Evaluation*: Vol. 26, Article 10.
DOI: https://doi.org/10.7275/szfa-da27
Available at: https://scholarworks.umass.edu/pare/vol26/iss1/10

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 26 Number 10, May 2021 ISSN 1531-7714

Creating an Evaluation Factor Group Work Assessment

Scott Mark., *Bow Valley College*

The idea of group work assessment is often seen as complicated, but necessary given the value of group work to learning. However, the traditional method involving a collective grade has been shown to be a less effective option, and raises questions about fairness for the students. Utilizing an evaluation factor provides a more inclusive assessment strategy, factoring in peer and self-assessment to enhance the reliability of the assessment. Through the evaluation factor, an instructor can maintain the product of the group work as the foundation of assessment, while allowing the process of the group work to be reflected in the evaluation factor. In this way, the evaluation factor provides a lens through which each individual in the group is more comprehensively assessed, and concerns for fairness are addressed.

In education, incorporating group work, in which students are directed to work together and collaborate on a single project, into lessons presents a significant challenge. With the growing popularity for group work at all levels of education, as well as research indicating scientific support for the benefits of group work in learning (Chiriac, 2014), the inclusion of group work has become a virtual necessity and a crucial aspect of effective instruction. However, to incorporate group work into a course, one must choose an assessment strategy, and therein lies a substantial aspect of the challenge. Teachers often describe group work assessment as complex and challenging, because there are concerns regarding fairness and individual assessment (Forsell et al., 2019). Commonly, at the end of a group work assignment, there is a final product or presentation, but a teacher can only infer the contributions of each group member. As a result, "the most common practice in higher education is for students to be graded solely on the basis of the quality of a submitted piece of work without consideration of the effort or input into that product" (Ko, 2013, p.302). As the prevalence of group work increases in higher education, many students have raised concerns and disagreed with grades being collectively awarded (Grammenos, et al., 2019). Stemming from this concern has been an increased focus on group assessment involving the use of peer and self-assessment to address individual concerns. As a whole, there is a general consensus that self-assessment provides benefits for students (Brown & Harris, 2012; Chin 2016), and peer feedback can strengthen the accuracy of a student's comprehension of the quality of their work (Ross, 2006). While the manner in which peer and self-assessment is integrated into the evaluation varies, I propose higher education specifically can use peer and self-assessment as a lens through which the teacher can observe more information and provide a more comprehensive assessment.

Evaluation Factor Method

As the final product of a group work assignment is commonly evaluated collectively, with little to no inquiry into individual effort, there is a clear concern for the lack of individual assessment. Due to the presence of others in a group work context, individuals may perceive less accountability (Garcia et al., 2002), and consequently withhold or reduce their efforts as they believe doing so will not affect the outcome (Karau & Williams, 1993). This type of reduced effort for some of the group members may be described as

free riding or social loafing, which ultimately reduces the morale and effectiveness of the group work context (Karau & Williams, 1993; Schippers, 2014; McArdle et al., 2005).

To address this concern, peer and self-assessment may be used to both combat the perception of lowered accountability, as well as allow the teacher to develop a lens, or Evaluation Factor (EF), through which each individual group member's contribution may be viewed, influencing their individual assessment. In this way, the overall product of the group work remains at the base of the assessment but the peer and self-assessments provides the lens for each individual assessment. This EF may be determined using a brief but simple calculation, as shown in figure 1, which begins by determining the Grade Factor (GF) for each group member. In this calculation the self-evaluation has as much weight as the combined average of all peer evaluations. Then each individual's GF will be measured against the average Figure 1 – GF and EF calculation. of the group's GF, and this will determine the individual EF for each group member. An example of this strategy is outlined in Table 1, where you can see how an instructor assigned a grade of 86% for the group's final submission. When the evaluation factor was applied as a lens for each individual, there has been a range of final grades from 81% to 90% awarded, and the change was influenced by peer and self-assessments, allowing the individual assessment to be more comprehensive.

Figure 1. GF and EF Calculation

$$\text{Grade Factor (GF)} = \frac{\text{Self-Assessment}}{\text{Average of Peer Assessments}}$$

$$\text{Evaluation Factor (EF)} = \frac{\text{Individual GF}}{\text{Average of all group GF's}}$$

Table 1 . An example of the Evaluation Factor (EF) in a group work scenario. This example was taken from a real student group where the assessment strategy was implemented, and a rubric-based assessment was employed. All calculations were easily completed in a Microsoft Excel spreadsheet.

	Assessment Type				**GF**	**EF**	**Group Product Grade 86%**
	Self	**Peer**	**Peer**	**Peer**	Self / Avg of Peer	GF / Avg Group GF	Individual Grade 86% x EF
Student #1	15/16	15/16	16/16	15/16	15.167	0.991	**85%**
Student #2	15/16	13/16	14/16	14/16	14.333	0.937	**81%**
Student #3	16/16	16/16	15/16	15/16	15.667	1.025	**88%**
Student #4	16/16	16/16	16/16	16/16	16	1.046	**90%**

Peer and Self-Assessment

As a key aspect of the Evaluation Factor (EF), the first step is to determine what type of assessment system may be most effective for the peer and self-assessment. A criterion-referenced assessment -- in which the assessor measures the subject against pre-set criteria -- is recommended (Burton, 2006). Criterion-referenced assessment provides greater reliability, validity, and transparency than norm-referenced or any other type of assessment, and strength of the reliability and validity of peer and self-assessment is dependent upon objective and clearly defined criteria being used (Ross, 2006; Ramdass & Zimmerman, 2008; Brown & Harris, 2012). Therefore, a rubric-based assessment was determined to be an effective method, and has been the only type of peer and self-assessment applied for this overall assessment strategy.

The Rubric

To develop an effective rubric, there are two primary aspects to begin with: the criteria and the scale the criteria is assessed on. Both of which were considered to develop the rubric seen in Appendix #1, which was used for the example seen in Table 1.

The Criteria

As Tierney and Simon, 2004, indicate the criteria chosen should reflect products or performances that are valued in the course being taught, and Ko, 2013, indicates group work may be broken down into two main aspects: the product, and the process. The process, which lies outside of the perception of the assessor, but within the perception of the group's members, should be the focus. By making the process the focus, the assessor is including a perspective from within, resulting in a more comprehensive evaluation (Forsell et al., 2019). Then, from the process, rather than the product, the concentration of the criteria may be on performances that are valued within the scope of the instruction. With consideration for the common objectives of group, which often include a range of skills, such as: effective team work; appreciation and respect for other views; as well as techniques and problem-solving methods (Sofroniou, & Poutos, 2016). The criteria selected should be connected to attributes or skills that fall within these processes of group work. That being said, the selection should not be made solely by the teacher, as it is best if the students are involved in the co-construction of the criteria selected (Ross, et al., 1998a). By involving the students in the creation, they can be taught to use explicitly detailed criteria, to pay attention to the rubric, and to develop the ability to justify their evaluation, all of which can enhance the accuracy and reliability of rubric-based assessments (Laveault & Miles, 2002; Dunning et al., 2004; Ramdass & Zimmerman, 2008).

The Scale

With criteria selected, the next consideration is for the scale they will be placed on. A key aspect is to ensure the scale progresses clearly from one level to the next, and is a positive, progressive scale, which does not create a negative tone for lower levels (Tierney & Simon, 2004). A positive progressive scale, such as: developing, capable, proficient, exceptional, can use progressive terms in the description such as: few, some, most, all, when proceeding from the lower levels to the upper levels of the criteria. In doing so, the scale is promoting learning as opposed to demonstrating little to no expectations when descriptors such as none or never are used at lower levels. In addition, the rubric in Appendix #1 has an extra level for rare situations in which a student evaluator may feel as though a criterion cannot be evaluated because no discernable effort was made to address it. In such an instance, it is recommended the student evaluator communicate directly with the instructor to discuss issues that may fall beyond the scope of the rubric.

Limitations of the EF

Like all assessment strategies, consideration must be given to the limitations of this strategy. The first is that this strategy may not be as optimal for in-class group work, where the assessor may informally include their observations of the group process, as they walk around and offer immediate formative assessment while they observe interactions. The EF is most effective when used to increase the information on which the assessment is based, when the bulk of the group work and interaction occurs outside the formal classroom environment.

Another potential limitation for the EF is based on the quality and quantity of the peer and self-assessments. In terms of quantity, the EF may be less effective when the groups are smaller than four persons, as any aberrations in the peer feedback from a single individual may have a detrimental effect. As well, the quality of the peer assessment, based upon the criteria selected, requires consideration. However, in cases where aberrations or irregularities appear in the peer or self-assessments, this is where the teacher's action, as described by Ross, 2006, may be a conversation with the assessor to explore the justification behind the assessment. In addition, research has demonstrated that peer and self-assessment can be more effective than the formative assessment provided by an instructor (De Sande & Godino-Llorente, 2014), and both peer and self-assessment have been demonstrated to be suitable assessment instruments (Sharma et. al, 2016; Alzaid, 2017).

Conclusions

For group work assessment, the common method of awarding a collective grade has been shown to be a less effective option, that raises questions regarding its fairness among the students. As result, a more inclusive assessment strategy that factors in peer and self-assessment is a virtual necessity. Through the use of the evaluation factor method, a teacher is able to leave the product of the group work as the foundation of assessment. While also allowing the process of the group work, which is reflected in evaluation factor, to act as a lens through which each individual is more comprehensively assessed. By combining peer and self-assessment, along with teacher assessment, the added features enhance the fairness and efficacy of group work assessment (Cheng & Warren, 2000; Farcell, 2019; Alzaid, 2017).

References

Alzaid, J. (2017). The effect of peer assessment on the evaluation process of students. *International Education Studies*. 10. 159. DOI: https://doi.org/10.5539/ies.v10n6p159.

Brown, G., & Harris, L. (2012) Student self-assessment. In McMillan, J. H. (Ed.). (2012). *Sage handbook of research on classroom assessment (pp.367-393)*. Sage Publications. Retrieved from https://ebookcentral-proquest-com.proxy.queensu.ca

Burton, K. (2006). Designing criterion-referenced assessment. *Journal of Learning Design*. 1(2), 73-82. Retrieved from: https://files.eric.ed.gov/fulltext/EJ1066490.pdf

Cheng, W., & Warren, M. (2000). Making a difference: Using peers to assess individual students' contributions to a Group Project. *Teaching in Higher Education*. 5. 243-255. DOI: https://doi.org/10.1080/135625100114885

Chin, P. (2016). Peer Assessment. *New directions in the teaching of physical sciences. 13-18*. DOI: https://doi.org/10.29311/ndtps.v0i3.410

Chiriac, E. (2014). Group work as an incentive for learning – students' experiences of group work. *Frontiers in Psychology. 5. 558*. DOI: https://doi.org/10.3389/fpsyg.2014.00558

De Sande, J. C., & Godino Llorente, J. (2014). Peer assessment and self-assessment: Effective learning tools in higher education. *International Journal of Engineering Education, 30*, 711-721. Retrieved from: https://www.researchgate.net/publication/264116440_Peer_Assessment_and_Self-assessment_Effective_Learning_Tools_in_Higher_Education

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*(3), 69–106. DOI: https://doi.org/10.1111%2Fj.1529-1006.2004.00018.x

Forsell, J., Forslund Frykedal, K., & Hammar Chiriac, E. (2019). Group work assessment: Assessing social skills at group level. *Small Group Research*, *51*(1), 87–124. https://doi.org/10.1177/1046496419878269

Garcia, S., Weaver, K., Moskowitz, G., & Darley, J. (2002). Crowded minds: The implicit bystander effect. *Journal of personality and social psychology. 83. 843-53*. DOI: https://doi.org/10.1037/0022-3514.83.4.843

Grammenos, R., Souto, M., Chester, I., & Gimeno, L. (2019). Peer assessment of individual contribution in group work: a student perspective. Retrieved from: https://www.researchgate.net/publication/337731418_Peer_assessment_of_individual_contribution_in_group_work_a_student_perspective

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology, 65: 681–706.* Retrieved from: https://pdfs.semanticscholar.org/dbfb/3c9153d3aa75d98460e83fa180bc9650d6fd.pdf

Ko, S.-S. (2013). Peer assessment in group projects accounting for assessor reliability by an iterative method. *Teaching in Higher Education*. 19. 301-314. DOI: https://doi.org/10.1080/13562517.2013.860110

Laveault, D., & Miles, C. (2002, April). *The study of individual differences in the utility and validity of rubrics in the learning of writing ability.* Paper presented at the annual meeting of the American Educational Research Association. Retrieved from: https://files.eric.ed.gov/fulltext/ED464925.pdf

McArdle, G., Clements, K., & Hutchinson-Lendi, K. (2005). *The Free Rider and Cooperative Learning Groups: Perspectives from Faculty Members*. Retrieved from: https://files.eric.ed.gov/fulltext/ED492459.pdf

Ramdass, D., & Zimmerman, B. J. (2008). Effects of self-correction strategy training on middle school students' self-efficacy, self-evaluation, and mathematics division learning. *Journal of Advanced Academics, 20*(1), 18–41. DOI: https://doi.org/10.4219%2Fjaa-2008-869

Ross, J. A., Rolheiser, C., & Hogaboam-Gray, A. (1998a). *Impact of self-evaluation training on mathematics achievement in a cooperative learning environment.* Paper presented at the annual meeting of the American Educational Research Association. Retrieved from: https://files.eric.ed.gov/fulltext/ED422381.pdf

Ross, John A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation. 11(10).* DOI: https://doi.org/10.7275/9wph-vv65

Schippers, M. (2014). Social loafing tendencies and team performance: The compensating effect of agreeableness and conscientiousness. *Academy of Management Learning & Education. 13. 62-81.* DOI: https://doi.org/10.5465/amle.2012.0191

Sharma, R., Jain, A., Gupta, N., Garg, S., Batta, M., & Dhir, S. (2016). Impact of self-assessment by students on their learning. *International Journal of Applied and Basic Medical Research*. 6. 226. DOI: https://doi.org/10.4103/2229-516X.186961

Sofroniou, A., & Poutos, K. (2016). Investigating the effectiveness of group work in mathematics. *Education Sciences.* 6 (30) DOI: https://doi.org/10.3390/educsci6030030

Citation:

Mark, S. (2021). Creating an Evaluation Factor for Group Work Assessment. *Practical Assessment, Research & Evaluation*, 26(10). Available online: https://scholarworks.umass.edu/pare/vol26/iss1/10 /

Corresponding Author

Scott Mark, M.Ed.
Bow Valley College
Calgary, Alberta, Canada

email: smark [at] bowvalleycollege.ca

Appendix 1

Example Peer and Self-Assessment

		Scale				
Criteria	**Score**	**Exceptional (4)**	**Proficient (3)**	**Capable (2)**	**Developing (1)**	**See Instructor (0)**
Attended Group Meetings or Discussions		Attended all group meetings or discussions and participated sufficiently to the groups expectations	Attended most group meetings or discussions and participated sufficiently to the groups expectations	Attended some group meetings or discussions and participated sufficiently to the groups expectations	Attended few group meetings or discussions and participated sufficiently to the group's expectations	This category is only for situations when no discernible attempt was made, in any way, to meet the criteria involved. Please see the instructor if this is the situation.
Completed Assigned Tasks		All assigned tasks were completed to the level of expectation within the group	Most assigned tasks were completed to the level of expectation within the group	Some assigned tasks were completed to the level of expectation within the group	Few of the assigned tasks were completed to the level of expectation within the group	This category is only for situations when no discernible attempt was made, in any way, to meet the criteria involved. Please see the instructor if this is the situation.
Tasks completed on time		All tasks completed at times agreed upon by group, allowing time for peer review or submission.	Most tasks completed at times agreed upon by group, allowing time for peer review or submission	Some tasks completed at times agreed upon by group, allowing time for peer review or submission	Few tasks were completed at times agreed upon by group, allowing time for peer review or submission	This category is only for situations when no discernible attempt was made, in any way, to meet the criteria involved. Please see the instructor if this is the situation.
Maintained Respectful and Professional Attitude		Always maintained a respectful and professional attitude with the group	Usually maintained a respectful and professional attitude with the group	Occasionally maintained a respectful and professional attitude with the group	Rarely maintained a respectful and professional attitude with the group	This category is only for situations when no discernible attempt was made, in any way, to meet the criteria involved. Please see the instructor if this is the situation.
Total out of 16		Additional Comments:				