

## Marking Schemes for an Authentic Group Project, Trial by Statistics - A Case Study

**Kim Bryceson**

School of Agriculture and Food Sciences, The University of Queensland, Australia.

---

### **Abstract**

*This study is an analysis of two different marking schemes for an ‘authentic’ Group Project worth 50% of a first year undergraduate university agribusiness course at The University of Queensland (UQ). A number of different marking schemes for the Group Project had been trialled over the last ten years in an effort to obtain an equitable method of marking individual students doing the Group Project. In 2019, a marking scheme for the Group Project that had been successfully used previously was advertised for 2019 prior to the commencement of semester. However, issues during the semester within some of the Groups meant that students requested a Peer Evaluation marking scheme be employed. Eventually, for a class of 105 students, both marking schemes were used in assessing students’ work and a Pearson Correlation coefficient was run on the results of the final project mark to determine how equivalent the two marking schemes were. A good correlation (0.75) between the two schemes was returned, which was also reflected in a good correlation in the comparison for the final overall mark for the whole course (0.87). These statistical results suggest that there is a good argument for the existing marking scheme to continue to be used rather than a peer evaluation, which can have behavioural issues associated with it that are difficult to resolve.*

**Keywords:** Authentic; Group Project; Marking Scheme; Peer Evaluation; Correlation coefficient.

---

## **1. Introduction**

When lecturers meet in the corridor or chat over a cup of coffee, inevitably they get around to discussing issues they are facing in their classes. Often, the discussion quickly moves towards the assignments being undertaken by students, how they are doing them (or not) and the marking of them. Much can be learned by colleagues and peers during such informal discussions (Nonaka and Takeuchi 1996). In fact many times people find that ‘their’ problem or issue is not theirs alone but has presented itself already to others who have managed to create a solution or workaround. The project presented in this paper came about because of such discussions.

This study resulted from three issues in teaching a first year undergraduate course about the use of ‘E’ technologies in the agrifood chain in the Bachelor of Agribusiness at The University of Queensland (UQ) in Semester 2 2019. The issues were that (i) the main piece of assessment for the course was worth 50%, and needed to be an ‘authentic’ assessment piece that engaged students in making use of content taught in the course which related strongly to industry practice; (ii) with 105 students in the course, this piece of assessment needed to be undertaken in small groups (no more than 4 students in a group); and (iii) the marking system used had to deliver a fair and equitable mark to individual students within each group, which was more difficult than it initially sounds, because while the course was ostensibly a 1<sup>st</sup> year course, other higher year level students were able to take it as an elective. This meant that quite apart from potential issues of students not pulling their weight in the group because of poor scholastic effort (for example), group members were quite diverse in their knowledge, maturity and abilities, which did impact on the overall quality of the final projects delivered by groups. This in turn meant that a high performing individual in a low performing project could be penalised as a result of the make up of the members of the group (Chang and Brickman, 2018).

The project described here looks at the three issues in the context of contemporary educational literature and provides a statistical analysis of the validity of two criterion based marking (grading) schemes used in order to deliver equity to individual students in terms of their final marks for the Group Project in the case described.

### ***1.1. Group Projects as Authentic Assessment***

A group project can be defined as “a graded assignment requiring students to work collaboratively across multiple class periods and involving some time outside the normal class meeting” (Ettington & Camp, 2002, p. 357). Group projects as part of university course assessments are a widely used teaching tool (Wilson et al, 2018), and there is a well established pedagogical literature on the topic showing a number of benefits of group work for students, including: (i) learning teamwork skills (Davis and Miller 1996; Michaelsen et al. 2014) - a skill often requested by employers (Graduate Outlook Survey 2010); (ii)

improving critical thinking skills (Anderson et al. 2001, MacGuiness, 2005); (iii) improving communication and collaborative skills (Slavin, 2014); and (iv) gaining insight into a particular topic at a deeper level than individual research (Burke 2011). Furthermore, as large organizations have become increasingly dependent on small groups or teams to achieve their goals, it has become increasingly important for employees to have the ability to work together collaboratively in today's business world (Aggarwal and O'Brien 2008). As an 'authentic' piece of assessment (Frey et al. 2012), a group project, if designed properly, can often tick all the boxes. The definition of authentic assessment is commonly agreed to be an assessment with real world applicability and one that students can employ what they have learnt during the course to perform real-world tasks (Mueller, 2018).

Despite being a good teaching and learning tool, there are challenges associated with group projects, which if not managed can prevent effective learning and result in poor-quality outputs, unequal distribution of workload, and conflict among team members (Chang and Brickman 2018). Indeed, the concept of 'social loafing' or scholastic laziness (Aggarwal and O'Brien 2008, Pandeira and Aseng 2017) creates an imbalance of effort, such that 'free riders' are able to benefit from the contributions of others, which is received badly by other students. Lecturers also need to be aware that group projects introduce their own grading complexities – and it is the grading complexity of a group project that arose during the course of the semester that forms the focus of the study described here.

### ***1.2. Marking/ Grading of Group Projects***

Marking (or grading) a group project is a complex task. Koshy (2009) and Brookhard (2018) give good overviews of the literature and a multiplicity of university 'how to' webpages give advice. Essentially, however, what has to be determined is whether the lecturer assesses the product (overall piece of assessment or project) or assesses the process (evaluating an individual's work within the project), or both. Once this has been decided, the actual marking schemes to be used needs to be developed.

Barnes, (1997) describes two main schemes (i) criterion based reference frameworks, where assessment of an assignment is made on the basis of performance defined by pre-specified criteria; and (ii) norm referenced approaches, where assessment is made on the basis of performance relative to that of other members of the class or cohort. Criterion based frameworks have become more popular over time and requirements for more transparent schemes with better learning outcomes for the student have ramped up in recent years (Koshy 2009). Thus rubrics that 'articulate expectations for student work by listing criteria for the work and performance level descriptions across a continuum of quality' (Brookhard, 2018) within criterion based schemes have become more favoured compared to norm-based schemes.

## **2. The Case Study**

This case study is based on the marking of an authentic piece of assessment for a course entitled “ETechnologies in the Food & Fibre Industry” where students were asked to create an innovative “E” product within an electronically enabled Agribusiness and to ‘sell’ this product in as innovative a manner as possible, documenting how the new product could add value to the business. The project had to be undertaken in small groups of four students

(Burke, 2011 indicates that small groups generally realise better learning than larger groups), and was worth 50% of the overall course assessment marks. The groups were put together by the lecturer to prevent groups of friends or year levels ‘ganging up’. The project when completed was delivered via a new online platform at UQ called CIRRUS, which also had a facility to create a PDF document to be uploaded into the Course Blackboard site by each student. The Project was scaffolded with an instruction document with clear guidelines of requirements and Marking Criteria (Table 1) and was assessed on two main areas: The “Idea” (20 marks); and an ‘E’ Product/Agribusiness Business Plan (30 marks).

**Table 1 Marking Criteria and Ratings for the Group Project**

| <b>CRITERIA</b>   | <b>MARK</b> |
|---|-------------|
| <b>EAgricusiness IDEA</b>   | <b>(10)</b> |
| Creativity / Idea   |             |
| Logic in terms of an E Agribusiness and how it fits                       | <b>(15)</b> |
| Presentation method within CIRRUS (Innovative / Interesting)              | <b>(20)</b> |
| <b>TOTAL x/45 = x/20</b>  |             |
| <b>REPORT</b>   |             |
| <b>Business Snap</b> - Overview of the Business Mission, Vision, Goals,   | <b>(20)</b> |
| Products & Services Evidence of reading and desktop research              |             |
| (What, Why, How)  |             |
| <b>Key EAg issues</b> for the agribusiness industry sector                | <b>(15)</b> |
| you are dealing with in general.  |             |
| <b>Product Canvas Model</b> - Diagram correctly filled in                 | <b>(20)</b> |
| <b>IT &amp; Info Sys Plan</b> - What do you need                          | <b>(10)</b> |
| <b>Budget</b> - Income and basic costs                                    | <b>(10)</b> |
| <b>Marketing Strategy</b> - What is it                                    | <b>(5)</b>  |
| <b>Cyberdefense Strategy</b> - What is it                                 | <b>(5)</b>  |
| <b>Report Presentation</b> - Appropriate use of CIRRUS as per guidelines, | <b>(20)</b> |
| Grammar, spelling, writing skills, quality of diagrams, flow charts,      |             |
| URLs etc,   |             |
| <b>TOTAL (x/105) = (x/ 30)</b>  |             |
| <b>TOTAL (Idea + Business Plan Report) (x/50)</b>                         |             |

A marking scheme (the ECP Weighting system (ECPW)) that has been successfully used since 2013 for evaluating an individual's mark for the group project was used and advertised prior to commencement of semester (*success here defined in the sense that students have reported that the ECPW reduces 'free riding' that can occur in group projects*). Table 2 shows a worked example of how an individual's mark was calculated using the overall Project mark as basis. A weighting for each student in a Group was assigned by averaging their marks over the previous assignments in the course equivocated to projected final grade.

**Table 2. Worked example of Individual's mark calculation example is for a 5 person Group with an overall Project mark of 85%.**

| Weight/Grade<br>1,2,3 | Weight/Grade<br>4                | Weight/Grade<br>5 | Weight/Grade<br>6,7 |
|-----------------------|----------------------------------|-------------------|---------------------|
| 0.7                   | 0.8                              | 0.9               | 1                   |
| Persons 1 &2          | (average grades 7&6 respectivly) |                   | $85*1 = 85$         |
| Person 3              | (average grade 5)                |                   | $85*0.9 = 76.5$     |
| Person 4              | (average grade 4)                |                   | $85*0.8 = 68$       |
| Person 5              | (average grade 2)                |                   | $85*0.7 = 59.5$     |

However, issues during the semester within some of the groups (social loafing and free riding, not logging onto the CIRRUS platform to deliver any physical input into the project, poor or no communication, plagiarised work from websites), meant that several groups of students requested a Peer Evaluation marking scheme (Dyrud, 2001).

Consultation with the class to verify that this was what was collectively wanted and that everyone was familiar with the process from previous classes, took place. With an affirmative answer, a formal evaluation scheme with rubrics was developed (Table 3) which was mandatory for all students to use – uploading their evaluations with the CIRRUS linked PDF of their project into the Course Blackboard site Project assignment link for marking.

**Table 3. Peer Evaluation (PE) Marking Scheme and rubric. 7 Assessment Criteria and 5 levels of performance 1 (no input) - 5 (high input). Mark converted to PE weighting (PEW) factor 0-1.**

| Assessment Criteria   | Poss<br>Mark<br>5             | 0-1<br>No<br>input                  | 2<br>Low/<br>poor<br>input | 3<br>Average<br>input | 4<br>Substant<br>input   | 5<br>Frequent<br>input |
|---|-------------------------------|-------------------------------------|----------------------------|-----------------------|--------------------------|------------------------|
| <b>1. Level of input in the dev of the concept (<i>The IDEA</i>)</b>          |                               |                                     |                            |                       |                          |                        |
| <b>2. Level of input in <i>designing the Project</i></b>                      |                               |                                     |                            |                       |                          |                        |
| <b>3. Level of input in <i>collecting required background information</i></b> |                               |                                     |                            |                       |                          |                        |
| <b>4. Level of input in <i>report writing (<i>The REPORT</i>)</i></b>         |                               |                                     |                            |                       |                          |                        |
| <b>5. Level of input in finding Graphics and Graphics design</b>              |                               |                                     |                            |                       |                          |                        |
| <b>6. Group work &amp; Communication</b>                                      | No work in group              | Barely worked with & little effort. | Adequate comms & effort.   | Good Ave team member  | Excel & put into report. | Excel work, & comms.   |
| <b>• Collaboration</b>  |                               |                                     |                            |                       |                          |                        |
| <b>• Understanding</b>  |                               |                                     |                            |                       |                          |                        |
| <b>• Editing</b>  | effort Plays no part in Group | & min effort shown                  |                            |                       | Vital team Imp team memb | memb                   |
| <b>7. Quality of Work</b>   | Poor                          | Low                                 | Accept                     | Good                  | Exel                     |                        |

Total Peer Evaluation Mark based on above criteria = x/35

### 3. Marking Process for the Group Projects and Results

Eventually, for a class of 105 students, both the ECP weighting scheme (ECPW) and Peer Evaluation weighting scheme (PEW) were used in assessing students' work. A Pearson Correlation coefficient was run on the results of each group's overall project mark obtained using each marking scheme in order to determine how equivalent the schemes were. The process of marking the group projects was an exhaustive practice to ensure equity, given the concerns that students had voiced.

The project was first assessed overall (to obtain 'product' mark) and a group mark was assigned based on the Marking Criteria shown in Table 1. Individual student marks within each group were then calculated using their group's overall project mark and the ECPW shown in Table 2 (Individual Project Mark 1). An individual's project mark was also calculated using the overall project mark and the PEW scheme shown in Table 3 (Individual's Project Mark 2).

A Regression Analysis on each individual's marks for their *project* was calculated in Excel using the ECPW vs the PEW system and also on their ultimate overall *course* mark using the ECPW project mark vs the PEW project mark. An  $R^2$  of 0.7469415 and 0.876877 respectively was obtained meaning the two marking schemes were strongly correlated, and that neither would penalise students in terms of their overall final grade for the course.

There were some actual mark differences for individuals between the two marking systems (Figure 1). Most were less than 1 or 2 marks but eight student marks were significantly different ( $>5$  marks). On review these were due mainly to individuals either 'opting out' of being in group ( $n=2$ , with one student being given a PEW of 0.1 but having a much higher ECPW of 0.8, and the other a PEW of 0.35 and an ECPW of 1.0), or simply not participating at all and not submitting anything (eg PDF from CIRRUS or PEW document) on the Blackboard site ( $n=2$ ), or they were in a group in which all students gave each other full marks on the PEW ( $n=4$ ) such that poor early assignments giving a low ECPW, were negated.

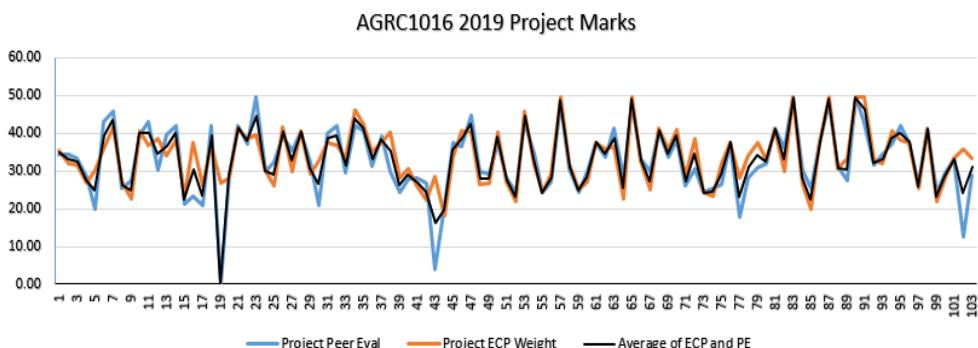


Figure 1. AGRC106 Group Project marks for 105 students.

My own observations and discussion with students indicated that the PEWs were fairly accurate in most cases, although it is acknowledged that scores were skewed in others. This reflects the literature where in a number of studies, PEW schemes have been shown to be both positive and negative – positive in that it gives students a chance to participate in the marking of their projects which gives ownership and interest and to some degree prevents 'free riding' but negative in that occasionally some students may not give honest evaluations or evaluations that really reflect other students efforts (Strong & Anderson, 1990, Dyrud, 2001). This, plus the fact that the PEW was introduced late in the semester and students were not thus 'prepped' for its use, made it seem fairest to calculate an individual's final project mark and thus final course mark using an *average* of the PEW calculated marks and the ECPW marks. As expected - the average had a strong correlation to both systems = PEW = 0.968871 and ECP = 0.968589 respectively.

## **4. Conclusion**

The results of this small statistical analysis between two criterion based marking schemes suggest that there is a good argument for the established and advertised ECPW marking scheme to remain in place for the Group Project in this course because it accurately reflects how students go on to work in their groups as assessed using the PEW scheme. This result, despite the previous years' of success in using the ECPW for marking the Group Project in this course, also allays a minor procedural concern that the ECPW - through its structural adjustment using the past quality of assignments to predict a future mark in the Group Project - could be seen as not facilitating a student to improve their final course mark over the semester. This potential issue which aligns with setting deep learning goals, is discussed in Hermida, (2015:Ch 1) and will be tested during a further iteration of the marking systems' comparison when the course is run again in 2020.

## **References**

- Anderson, T., Howe, C., Soden, R., Halliday, J., and Low, J. (2001). Peer interaction and the learning of critical thinking skills in further education students. *Instructional Science*, 29(1), 1–32.
- Aggarwal, P. and O'Brien C. (2008). Social Loafing on Group Projects Structural Antecedents and Effect on Student Satisfaction. *Journal of Marketing Education*, 30(3), 255-264.
- Barnes, L.L.B. (1997). Development of the 'Faculty Beliefs about Grades Inventory', *Educational and Psychological Measurement* 57(3), 459–469
- Brookhart S.M. (2018). Appropriate Criteria: Key to Effective Rubrics. *Frontier Education*, 3(22). doi: 10.3389/feduc.2018.00022.
- Burke, A. (2011). Group Work: How to Use Groups Effectively The *Journal of Effective Teaching*, 11(2), 87-95.
- Chang Y. and Brickman, P. (2018). When Group Work Doesn't Work: Insights from Students. *CBE—Life Sciences Education* • 17:ar52, 1–17.
- Davis, B.D. and Miller, T.R. (1996). Job Preparation for the 21st Century: A Group Project Learning Model to Teach Basic Workplace Skills, *Journal of Education for Business*, 72(2), 69-73. doi 10.1080/08832323.1996.10116828.
- Davies, W. M. (2009). Groupwork as a form of assessment: Common problems and recommended solutions. *Higher Education*, 58(4), 563–584.
- Dyrud, M. (2001). Group Projects and Peer Review. *Business Communication Quarterly*, 64, 106. doi10.1177/108056990106400412
- Ettington, D. and Camp, R. (2002). Facilitating transfer of skills between group projects and work teams. *J. Management Education*, 26, 356-379.
- Frey, B. B., Schmitt, V. L. and Allen, J. P. (2012). Defining Authentic Classroom Assessment, Practical Assessment, *Research, and Evaluation*: 17 (2).

- Graduate Outlook Survey (2010). University of Canterbury.  
<http://content.yudu.com/A1qpzf/GoAustralia2011/resources/index.htm?referrerUrl=http%3A%2F%2Fwww.graduateopportunities.com%2F>
- Hermida, J. (2015). *Facilitating Deep Learning: Pathways to Success for University and College Teachers*. Ch1:15. Apple Academic Press Inc, Toronto
- Koshy, S. (2009). Using marking criteria to improve learning: an evaluation of student perceptions', *Journal of Systemics, Cybernetics and Informatics*, 7(1) 72-76.
- McGuinness, C. (2005). Teaching thinking;theory and practice. *British Journal of Educational Psychology Monograph Series II, 3 Pedagogy-Teaching for Learning*, 1, 107–126.
- Michaelsen, L. K., Davidson, N., and Major, C. H. (2014). Team-based learning practices and principles in comparison with cooperative learning and problem-based learning. *Journal on Excellence in College Teaching*, 25(3&4), 57-84.
- Nonaka, I. and Takeuchi, H., (1996). The Theory of Organizational Knowledge Creation. *International Journal of Technology Management*, 11(7/8), 833 – 845.
- Pandeirot, L. and Aseng, A. (2017). Social Loafing and Group Performance: A Literature Review. *Abstract Proceedings International Scholars Conference*, 5(1), 106. <https://doi.org/10.35974/isc.v5i1.1655>
- Slavin, R. E. (2014). Cooperative learning and academic achievement: Why does groupwork work? *Anales de Psicología/Annals of Psychology*, 30(3), 785–791.
- Strong , J.T. and Anderson, R.E. (1990) Free riding in group projects: Control mechanisms and preliminary data. *Journal of Marketing Education*, 12, 61-67.
- Wilson, K.J., Brickman, P., Brame, C. J. (2018). Group Work. *CBE Life Science Education*. Spring; 17(1): fe1. doi: 10.1187/cbe.17-12-0258.